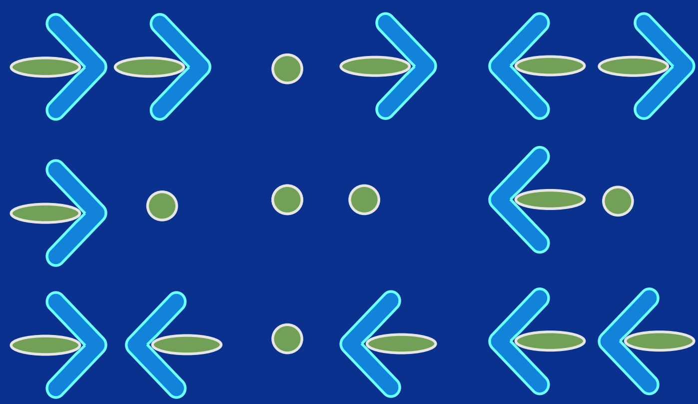
New Directions in the Analysis of Movement Patterns in Space and Time

Seyed Hossein Chavoshi



UNIVERSITEIT GENT

New Directions in the Analysis of Movement Patterns in Space and Time

Seyed Hossein Chavoshi

To the memory of Ali Mohammad Nasr Azadani

The research reported in this dissertation was conducted at the CartoGIS research unit, Department of Geography, Faculty of Sciences, Ghent University, and funded by the Special Research Fund (BOF).

UNIVERSITEIT GENT

New Directions in the Analysis of Movement Patterns in Space and Time

Dissertation submitted in accordance with the requirements for the degree of Doctor of Sciences: Geomatics and Surveying

Nieuwe Richtingen in de Analyse van Bewegingspatronen in Ruimte en Tijd

Proefschrift aangeboden tot het behalen van de graad van Doctor in de Wetenschappen: Geomatica en Landmeetkunde

by/door Seyed Hossein Chavoshi

Supervisors

Prof. dr. Nico Van de Weghe
Department of Geography, Ghent University

Prof. dr. Bernard De Baets
Department of Mathematical Modelling, Statistics and Bioinformatics,
Ghent University

Prof. dr. Guy De Tré
Department of Telecommunications and Information Processing,
Ghent University

Members of the Reading Committee

Prof. dr. Christophe Claramunt
Department of Computer Sciences, The French Naval Academy Research Institute

Prof. dr. Roland Billen
Department of Geography, University of Liege

Prof. dr. Peter Bogaert
Department of Geography, Ghent University

Dr. Tijs Neutens
Department of Geography, Ghent University

Remaining members of the Examination Committee

Prof. dr. Ben Derudder (Chairman)
Department of Geography, Ghent University

Prof. dr. Nico Van de Weghe
Department of Geography, Ghent University

Prof. dr. Bernard De Baets
Department of Mathematical Modelling, Statistics and Bioinformatics,
Ghent University

Prof. dr. Guy De Tré
Department of Telecommunications and Information Processing,
Ghent University

Prof. dr. Antoon Bronselaer
Department of Telecommunications and Information Processing,
Ghent University

PREFACE

This thesis was carried out during five years of my life as a research assistant at the Department of Geography, Ghent University.

This research was financially supported by the Special Research Fund (BOF) of Ghent University, for which I would like to express my sincere gratitude. This thesis would not have been possible without the collaboration of several people who in one way or another contributed and extended their valuable assistance. Here, I name some of them and I inwardly thank the others.

To start with, I would like to sincerely thank my supervisors Prof. dr. Nico Van de Weghe, Prof. dr. Bernard De Baets, and Prof. dr. Guy De Tré, who offered me the opportunity to conduct a PhD. Nico has learned me about the Qualitative Trajectory Calculus (QTC) and introduced me to the world of 'Qualitativeness' upon which the basis of my PhD thesis is formed. Beyond his support, patience and encouragement, he has amazed me time and time again as a fount of original ideas. We have had many productive discussions both on and off the topic of this research, which helped to shape the ideas presented in this research. I am grateful to him for his admirable academic guidance and for his willingness to be always available for me.

I am also indebted to Bernard who enthusiastically supported me throughout the entire project. His discerning and constructive comments have guided and challenged my thinking and helped me to improve my work substantially. I consider myself very fortunate to have been advised by such a devoted and exceptionally bright scientist with an astonishing knowledge of the scientific literature.

I would further like to express my sincere gratitude to Guy for giving me the confidence and liberty to explore various research fields and for his invaluable aid and the helpful suggestions along the way. I would also like to thank Prof. dr. Philippe De Maeyer, head of the Department of Geography, who has helped me in many aspects during my work in this department.

I am particularly indebted to the other members of my PhD committee: Prof. dr Peter Bogaert, Prof. dr. Ben Derudder, Prof. dr. Christophe Claramunt, Prof. dr. Roland Billen, and Prof. dr. Antoon Bronselaer. I am very grateful for their time for discussion, their precious suggestions and inputs in this thesis.

Many thanks go to the colleagues of the Department of Geography, in particular to Dr. Tijs Neutens, who has helped me in many aspects during my work in this department. I will never forget his support during the last two years of my PhD when I was overwhelmed with many tasks and he was giving me a lot of motivations to manage that difficult situation.

I would like to express my gratitude to my (ex)officemates Dr. Matthias Delafontaine, Dr. Yi Qiang, Dr. Mathias Versichele, Manuel Claeys Bouuaert, and Roel Huybrechts for giving advices to my research and publications.

An extra word of gratitude goes to Dr. Luiz Naveda, Dr. Wim Kellens, Dr. Kristien Ooms, Dr. Els Verfaillie, Bart De Wit, Ilke Werbrouck, Bart Dewulf the colleagues who helped me a lot either scientifically or practically. I also acknowledge the excellent administrative assistance of Helga Vermeulen who has liberated me from numerous and trivial administrative matters.

This study has also benefited from the fruitful collaboration with Prof. dr. Hyowon Ban who has given me a warm welcome at the Department of Geography at the California State University, Long Beach. My visit to Long Beach City has been an instructive and enjoyable highlight. Thanks, Hyowon! I would also like to thank Prof. dr. Kathleen Stewart for the productive cooperation and the useful brainstorming session in the University of Iowa. I have also benefited from conversations with Prof. dr. Ola Ahlqvist from the Department of Geography, Ohio State University.

During my PhD study, I benefited immensely from two European projects, namely FP7/ICT project MODAP "Mobility, Data Mining, and Privacy" (www.modap.org) and the COST Action IC0903 MOVE "Knowledge Discovery from Moving Objects" (http://move-cost.info/). I was lucky that the period of my PhD research coincided with the time period of these two projects. I could develop and improve my research work and strengthen my research network by attending different workshops, conferences, and summer schools organised/funded by them. Here, I would like to thank Dr. Francesca Cagnacci from the Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige (TN), Italy. I had a nice collaboration with her and her group during my Short Term Scientific Missions (STSMs) funded by COST Action IC0903 MOVE.

Beyond the academic context, I have been strongly supported by my friends and family. Thanks to all my friends in Ghent who provided me with the necessary diversion specially Mohammad Masoudi and his wife, Amir Fardis, and Amir Jahanshahi.

I would also like to give special thanks to my parents Forough and MohammadReza, my parents-in-law Ehteram and Adel, my brothers and sisters who emotionally and financially have been supporting me a lot up to now.

Finally, there is one person left without whom this thesis would never have been completed, my wife. Maryam, thanks awfully for taking care of me and accompanying me through the past five years. I am glad that you are my wife and also my colleague. I remember those days and nights that we were discussing about my research. I am indebted to you for your genuine interest, your constructive and discerning comments, and your excellent academic advice. Anita, our new family member, you have changed our life and have added lots of pleasure.

I would like to end with Albert Szent-Györgyi's quote: "Research is to see what everybody else has seen and to think what nobody else has thought". I sincerely hope that after this, I can continue the work on the untackled issues within this vast and multidisciplinary domain.

Seyed Hossein Chavoshi
Ghent University
December 2014

Table of Contents

LIST OF FIGURES

List of Tables

List of Abbreviations

1D	One-Dimension/Dimensional
2D	Two-Dimensions/Dimensional
3D	Three-Dimensions/Dimensional
CTM	Continuous Triangular Model
DTW	Dynamic Time Warping
GIS	Geographic Information System
GIScience	Geographic Information Science
GISystem	Geographic Information System
GKD	Geographic Knowledge Discovery
GPS	Global Positioning System
H	Head
JEPD	Jointly Exhaustive and Pairwise Disjoint
LF	Left Finger
LT	Left Toe
MDWP	Minimum Dynamic Warp Path
MoCap	Motion Capture
MPOs	Moving Point Objects
QTC	Qualitative Trajectory Calculus
QTC_B	Qualitative Trajectory Calculus – Basic
QTC_C	Qualitative Trajectory Calculus – Double-Cross
QTC_N	Qualitative Trajectory Calculus – Network
QTC_S	Qualitative Trajectory Calculus – Shape
R	Root
RCC	Region Connection Calculus
REMO	RElative MOtion
RF	Right Finger
RFID	Radio Frequency Identification
RT	Right Toe
SAM	Sequence Alignment Method
SESI	Sequence Signature
TM	Triangular Model

1

Introduction

I have no special talent. I am only passionately curious. Albert Einstein

1 INTRODUCTION

1.1 Background and State of the Art

The act of changing location from one place to another is known as movement. This concept has existed since the advent of the world and can be seen from different perspectives. The movement of the Earth, for example, causes day and night, the change of the seasons, and the tides. Human migration can be seen as another angle of the broad concept of movement. It relates to the physical/spiritual movement of a human being from one area to another with a different motivation. Movement has been the source of many revolutions such as tribal, agricultural, industrial, economic, social, political, etc. Therefore, it is not unreasonable that many prospective scientists from different areas of science and technology, such as ecology, meteorology, sociology, behavioural studies, transportation planning, surveillance, and intelligence services conduct research to fully comprehend different aspects of movement and unravel the knowledge hidden in movement data.

The contribution of this work, with regard to previous literature, is threefold: (i) to explain the concept of movement in a qualitative model and describe its position in GIScience; (ii) to explain the procedure of knowledge discovery in movement; (iii) to introduce the similarity analysis problem in movement and report on the relevant literature. In each of the following subsections, we discuss our contribution in more detail.

1.1.1 Qualitative Reasoning about Movement

GIScience is "the subset of information science that is about geographic information" (Goodchild et al., 1999, p. 737). Based on the definition by Worboys & Duckham, GISystems are computer-based information systems which enable us to capture, model, manipulate, retrieve, analyse, and present geographically referenced data (Worboys and Duckham, 2004). GISystems can also handle data regarding dynamic phenomena. Dynamic phenomena include a wide range of applications in many domains of scientific research and engineering such as animal ecology, disaster management (e.g. fire propagation, plate tectonics, and volcanic eruptions), urban planning, and traffic management. Because of the dynamic nature of most spatial objects, much effort has been devoted to extend the potential of spatio-temporal visualisation and reasoning in GISystems. Development of technologies that capture the spatio-temporal nature of objects outpaces the capability of GISystems to cover the full range of space and time reasoning, and still requires fundamental research.

In previous decades, valuable work has been done in the areas of spatial and temporal reasoning (for example, see (Rajagopalan, 1995; Cohn, 1997, Muller, 1998; Cristani et al., 2000; Wolter and Zakharyaschev, 2000; Hornsby and Egenhofer, 2002; Merz, 2003; Héas, 2005; Gottfried, 2006; Guesgen, 2010; Petitjean, 2012; Menon, 2013; Basiri, 2014). Among others, the Interval Calculus (Allen, 1983), the Semi-Interval Calculus (Freksa, 1992a), the Region Connection Calculus (Cohn et al., 1997), and the 9-Intersection Model (Egenhofer & Franzosa, 1991) express qualitative spatial or temporal relationships between entities. In this regard, moderate attention has been paid to consider both spatial and temporal dimensions simultaneously to establish spatio-temporal relationships between entities. In recent years, considerable attention has been paid to spatio-temporal relations between entities, i.e. relative motion (see, for example, (Claramunt & Jiang, 2001; Noyon et al., 2007)), but less to effectively describe motion within a qualitative framework. The Qualitative Trajectory Calculus (QTC) has been proposed by Van de Weghe (2004) as a qualitative spatio-temporal representation for objects that are moving based on the characterisation of their trajectories with respect to each other. Van de Weghe started from the idea that interactions between objects in the real-world can be described by relationships between pairs of moving point objects (MPOs) moving in a one-dimensional space (Van de Weghe, 2004). In many cases, the enormous complexity of real-world dynamic phenomena can be described by MPOs, being constantly disjoint. For example, an increasing number of moving vehicles has caused many traffic experts to focus on research for safe and efficient traffic flow. They often treat vehicles as MPOs in their analysis regardless of the size and shape of the vehicles. This way building quantitative and qualitative models becomes much more appropriate.

In fact, the strength of QTC, compared to other approaches in qualitative spatial reasoning (e.g. Region Connection Calculus (Cohn et al., 1997)), is that QTC starts from a dynamic distance change between a pair of objects over time whereas others pay attention to static relationships between objects. Therefore, we believe that QTC is perfectly suitable to explore the movement of objects in a qualitative framework.

This thesis seeks to contribute to research on the movement of tracked individuals/groups of objects that can be treated as MPOs. For this purpose, we make extensive use of QTC which has been studied to formulate interrelations between moving objects. We employ the concept of this calculus to show its usefulness, applicability, and power in reasoning about the movement of objects. Below, we describe why we employ qualitative models, and particularly QTC, instead of many quantitative/numerical models for handling movement.

We would like to reduce the complexity of the analysis of movement while preserving its important features and obtain desirable results. High dimensionality, geometry, topology, spatial dependency, and spatial heterogeneity are among those properties that add complexity to geographic systems and particularity GISystems (Miller & Han, 2009). Spatio-temporal data including moving objects are even more complicated to process than other geographic data (Dodge, 2011). Therefore, appropriate analytical methods need to be developed to capture different aspects of spatio-temporal data (Imfeld, 2000). To decrease the complexity of movement analysis, QTC makes four simplifications, namely relational simplification, object simplification, topological simplification, and temporal simplification. In brief, QTC considers the relation between two objects, i.e. binary relations, as relational simplification. Additionally, in this calculus, moving objects are spatially simplified into MPOs known as object simplification. Looking at the real world, we discern that many moving objects around us, such as people, airplanes, and celestial objects, have disjoint relations. Since there are only two topological relationships between moving objects, namely disjoint and equal, QTC only considers disjoint relations as topological simplification. QTC relations hold at a particular time point. This is helpful to know what happens at one time point and understand the temporal dimension in depth (temporal simplification). These simplifications have been extensively expressed in (Delafontaine, 2012).

Van de Weghe, in his doctoral dissertation, stated that humans usually prefer to communicate in qualitative categories supporting their intuition and not in quantitative categories (2004). He has also mentioned some examples in this regard. For example, “the first car is moving faster than the second one, and not: the first car is moving at a speed of 119 kilometers per hour and the second car at a speed of 116 kilometers per hour” (Van de Weghe, 2004, p. 2). A number of approaches have been developed in this regard. For example, Zadeh (2002) proposed a methodology to compute with words in contrast to numbers. "Computing with words is a methodology in which the objects of computation are words and propositions drawn from a natural language" (Zadeh, 2002). Thus, information systems, including GIS, which analyse spatio-temporal phenomena, should employ such qualitative reasoning methods in their systems.

Many researchers have investigated various aspects of moving objects, from generating spatiotemporal datasets that simulate real-world behaviour (Brinkhoff, 2002; Pfoser & Theodoridis, 2003) and indexing (Agarwal et al., 2003; Šaltenis et al., 2000), to modelling and querying (Erwig et al., 1999; Prasad Sistla et al., 1997). However, the reasoning on the relations between MPOs has attracted less attention, especially in a qualitative framework (for example, see (Bogaert, 2008; Cohn & Renz, 2008; Van de Weghe, 2004)). We believe that QTC constitutes a basis to adequately study the movement of MPOs. Some works have already been published to show the applicability of QTC in analysing the movement of MPOs (for example, see (Bogaert et al., 2007; Delafontaine et al., 2011a; Delafontaine et al., 2011b; Van de Weghe et al., 2005a)). As for our contribution, we add or enhance the practical usefulness of QTC, and close the gaps between theory and application. In addition, we intend to disclose the links that interconnect the spatio-temporal qualitative reasoning such as QTC and knowledge discovery from movement.

1.1.2 Knowledge Discovery from Movement

One major area of concern is how to effectively integrate qualitative reasoning in the process of knowledge discovery. Basically, dynamic behaviour, especially movement data, are relatively complex to study and analyse compared to other types of data. Therefore, as mentioned earlier, qualitative reasoning can be involved in this procedure to greatly reduce the complexity.

Knowledge discovery from movement databases was introduced in (Giannotti & Pedreschi, 2008), and then further extended in (Dodge, 2011). It includes three main steps, namely trajectory reconstruction, knowledge extraction, and knowledge delivery, as shown in Figure 1-1.

In Figure 1-1, the components of the knowledge discovery process, which will be covered throughout this thesis, are highlighted. The darker the highlight, the greater the stress on the component in this work will be. We only marginally touch upon the first component of the knowledge discovery process, i.e. trajectory reconstruction and preprocessing. We often, instead of static geometric shape of trajectories, transform them into qualitative relationships between MPOs. This step is considered as a preprocessing step.

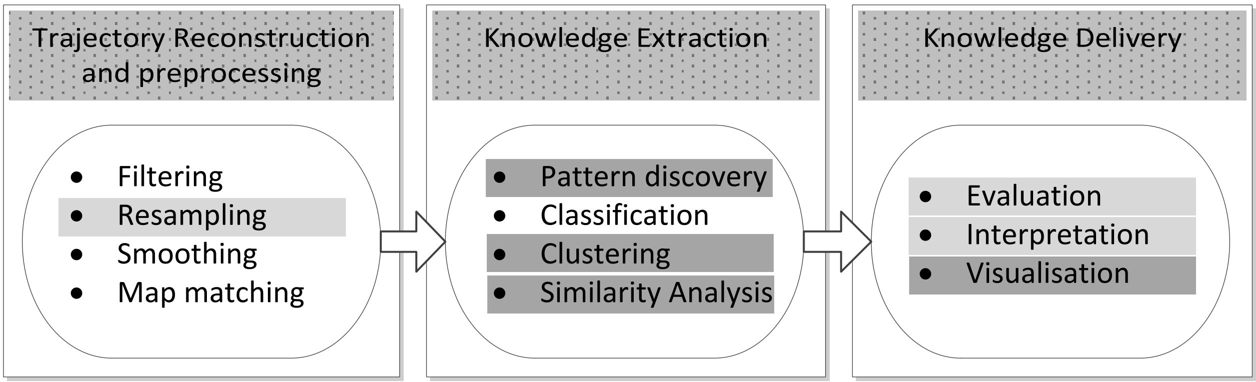


Figure 1-1: Knowledge discovery process in movement databases (modified based on (Dodge, 2011))

The second component of the knowledge discovery process presented in Figure 1-1 is knowledge extraction. It refers to the extraction of useful knowledge from the movement of MPOs using appropriate techniques in knowledge discovery and data mining such as pattern discovery, classification, clustering, and similarity analysis (Dodge, 2011; Giannotti & Pedreschi, 2008; Miller & Han, 2009). Advantages of knowledge discovery techniques from moving objects are demonstrated in many studies, including contributions on pattern discovery (Demšar & Virrantaus, 2010; Dodge et al., 2008; Du Mouza & Rigaux, 2005; Gudmundsson et al., 2007; Laube et al., 2008; Laube et al., 2005; Wilson, 2008), extraction of clusters (Buzan et al., 2004; Jensen et al., 2007; Li et al., 2004; Nanni & Pedreschi, 2006; Rinzivillo et al., 2008; Zhang & Lin, 2004), and similarity within moving object data (Buchin et al., 2011b; Dodge et al., 2012; Dodge et al., 2009; Lin & Su, 2008; Pelekis et al., 2007). Terms such as pattern discovery, similarity analysis, and clustering are used very often in this thesis and reflect our attention to this stage in the knowledge discovery process. The next subsection (1. 1. 3) is dedicated to the literature about similarity analysis in movement data. Similarity analysis in movement data forms a major part of this thesis.

The third component of the knowledge discovery process, i.e. knowledge delivery, has been extensively studied in literature (Andrienko & Andrienko, 2008; Andrienko et al., 2013; Andrienko & Andrienko, 2012; Pelekis et al., 2012). At this stage of the procedure, we need to appropriately assess the outcomes. In this thesis, the proposed visualisation techniques support us to evaluate and interpret extracted knowledge.

1.1.3 Similarity Analysis in Movement Data

Based on the definition given by Alt and Guibas (1999), the purpose of similarity analysis is to define how much two objects resemble (are similar to) each other. For example, in movement data, trajectories, patterns, and the behavioural properties of moving objects are among those objects to be analysed for similarity. Many studies have been conducted to address various aspects of similarity analysis problems in movement-related disciplines. In this subsection, we review only some of them. Dodge (2011), has elaborateda comprehensive literature study on similarity analysis in movement data. According to what she described, similarity search efforts in movement data have been mainly inspired from two research fields, namely time series analysis and computational geometry (Dodge, 2011). This is reasonable because most of the similarity analysis methods developed for time series analysis or geometric shapes can be easily adapted to data mining applications for movement.

Basically, there are five groups of similarity measures defined for time series data, namely Minkowski distance, edit distance, dynamic time warping (DTW), longest common subsequence (LCSS), and distances based on local features (Dodge, 2011). Research along these lines constitutes a part of our focus in this dissertation. With the consideration that a trajectory is defined as a sequence of consecutive locations of the moving object over a period of time, it can accordingly be considered as a time series of spatial data in data mining tasks (Spaccapietra et al., 2008).

In addition to the analysis of the geometric shape of trajectories and point-related attributes such as speed, the evaluation of interaction between moving objects becomes more important. This thesis is centered on the development of similarity measures to address relationships between moving objects.

The complexity of analysis of movement data is considerably increasing with the growing number of repositories of movement data. Therefore, novel approaches to extract knowledge from this type of data are needed. One of the primary techniques to understand movement data sets is clustering of moving objects in terms of trajectories, patterns, etc. Clustering is defined as the procedure of grouping a collection of objects into subsets whose in-class members share a similarity in some sense. There are several criteria playing a role in the clustering of moving objects. In (Kisilevich et al., 2010), an overview of the state-of-the-art approaches and methods of spatio-temporal clustering in different application domains is given. The following example, taken from (Kisilevich et al., 2010), clearly shows that clustering is mostly application dependent and depends on many factors. “When tracking pedestrians, for example, two geographically close sample points co-occurring within a minute interval could belong to the same cluster, whereas two sample points at near distance within a time interval of a few nanoseconds in a physics experiment might belong to different clusters” (Kisilevich et al., 2010, p. 855). Until now, based on the main classes of spatio-temporal data types introduced in (Kisilevich et al., 2010), different clustering techniques have been defined to cluster spatio-temporal events and geo-referenced times series of moving objects. Most of the existing trajectory clustering techniques rely on the similarity of the geometric shapes (Buchin et al., 2011a; Fu et al., 2005; Giannotti & Pedreschi, 2008; Lee et al., 2007; Li et al., 2010b; McArdle et al., 2013; Miller & Han, 2009; Nanni & Pedreschi, 2006; Rinzivillo et al., 2008; Yanagisawa & Satoh, 2006), and non-spatio-temporal attributes of the trajectories (e.g. transport mode). Less attention has been paid to developing techniques to cluster interrelations between moving objects. In this dissertation, the latter is fully addressed. Different similarity analysis techniques are developed and their outcomes are substantially exploited in the clustering of interrelations between moving objects. In this respect, additional information is required to cluster trajectory data according to the interaction of moving objects. As discussed earlier, QTC can perfectly describe motion, and more specifically the interactions between moving objects (Van de Weghe, 2004). Therefore, we consider this formalism to form the basis for defining novel similarity measures and, consequently, for developing clustering techniques for trajectories of moving objects. Unlike various geometric clustering techniques of trajectories, which do not necessarily capture spatio-temporal similarity between the movements of objects (Dodge, 2011), the approach presented in this research can effectively reflect similarities in movement behaviour of interacting moving objects.

1.2 Case Study

Human movement analysis is receiving increasing attention from experts of different disciplines such as athletic performance analysis and rehabilitation (Grassi et al., 2005; Mirabella et al., 2011), mass event management (Versichele et al., 2012), surveillance (Haritaoglu et al., 2000), games and animation (Menache, 2000), and art (Leman &

Naveda, 2010). Recently, in GISciences, various basic and applied research have been pursued to increase the understanding of human movements and activities. In this thesis, we also narrow our focus to human movement analysis and employ data from some potential applications to better understand the techniques presented in each chapter.

The proposed approaches in this thesis are mostly applied in a case study of Samba dance. The rhythmic movement in dance attracted our attention for several reasons; (i) The movement of each body part of a dancer can be represented by a trajectory in a three-dimensional space, as well as by a time series of different motion attributes to be used to reach the objectives outlined in this thesis. (ii) In any dance genre, the quality of the performance highly depends on the movement of each individual body part and also its interactions with other body parts. Since one of our goals in this research is to examine interactions between multiple MPOs, the case of dance is an appropriate example for this purpose. (iii) In this study, depending on the purpose of each chapter, different visualisation techniques are developed and/or employed. Unlike disorderly movements, dancers perform certain movements regularly during their performances and these regularities are perfectly reflected in the representations. Understanding the types of patterns which may exist in the underlying phenomena facilitates the detection of patterns (Andrienko & Andrienko, 2007). Displayed patterns allow us to understand, explore, and interpret the movements of dancers.

In general, the use of a dance data set allows us to look at intuitive aspects of dance, such as staying on the beat and to complete particular dance moves, and use them as analogies for object movement patterns of interest to a geographic analysis.

Fast technological improvements in positioning and tracking systems have made it possible to capture massive amounts of movement of animals, aircrafts, ships, humans, etc. Unlike the conventional methods of collecting indoor / outdoor human motion data, such as GPS, Bluetooth, ByteLight (using light to beam information to your phone), Barometer (using atmospheric pressure), Ultrasonic, and Wi-Fi, there are many other approaches to capture dance movements available for further analysis. Choosing the right technique highly depends on the purpose of research and the degree of accuracy expected from the results. In the following, some examples are provided. Sensing floors have been used to detect footsteps (Griffith & Fernstrom, 1998; Johnstone, 1991; McElligott et al., 2002; Paradiso et al., 1997; Srinivasan et al., 2005). For example, the *Magic Carpet system* (Paradiso et al., 1997) was made of a grid of wires that were insulated with a piezoelectric material to sense the footsteps. Sensing shoes, equipped with pressure sensors, allow measurement of the pressure exerted by the toes and the heel of dancers (Paradiso et al., 1999; Yoonji et al., 2008). There have been many systems that use cameras to capture the movement of dancers (Bevilacqua et al., 2001; Castellano et al., 2007; Ng, 2004). Camera-based systems are used in many domains and not only in dance to record the changes in movement of objects over time such as in traffic. On the other hand, various sensing advances have been used to create wearable sensors to record dance movements (El-Nasr & Vasilakos, 2008; Flety, 2005; Hromin et al., 2003). For example, Hromin et al. (2003) used accelerometers, flex sensors, temperature sensors, photoresistors and pressure sensors, each of which provides a certain type of movement data. The Samba dance data in this research has been collected with a Motion Capture (MoCap) system that records the position of objects over time by means of reflective markers attached to the objects in combination with infrared cameras. Only a very basic dataset including three-dimensional coordinates of five body parts, namely right finger, left finger, right toe, left toe, and head of each Samba dancer is considered. We may use right hand|right finger, left hand|left finger, right foot|right toe, left foot|left toe interchangeably throughout this thesis. More information with regard to data preparation and customization is given in the following chapters.

1.3 Research Objectives

The previous section elaborated several problems with respect to the analysis of movement. We mainly aim to extend our understanding of the movement behaviour of single or multiple moving objects. In this regard, we seek to accomplish several objectives throughout this thesis. This section presents the general research questions (RQ) that this thesis intends to address and the corresponding chapters in which they are addressed. To cover all the objectives below, an integration of diverse techniques in geographic information systems, data mining, and visualisation is utilised.

RQ 1: How do we enhance the practical usefulness of QTC?

The first research question is addressed in all chapters of this thesis except in Chapter 8, which claims that results from qualitative reasoning are comparable with those from quantitative ones.

In this study, we make extensive use of QTC, which has been studied in the field of diagrammatic representation and reasoning of MPOs (Van de Weghe, 2004). We are convinced that this calculus, despite its uncomplicated concept, is appropriate to describe and reason about the enormous complexly-interacting objects of the real-world. Throughout this thesis, we contribute to the enhancement of the practical usefulness of QTC. We take advantage of this existing theoretical contribution and apply it to potential applications. In this light, Chapter 2 reviews the theoretical cornerstones of the fundamental types of QTC and demonstrates how the calculus can be implemented and extended in order to represent and reason about raw moving object data. In Chapter 3, the concept of the continuous triangular model (CTM) is explained in detail. Chapters 4-7 further elaborate on a specific type of QTC: the *Qualitative Trajectory Calculus-Basic* (QTC_B). QTC_B defines a binary relation between two MPOs on the basis of the Euclidean distance in an unconstrained n-dimensional space. All other types of QTC including QTC Double-Cross (QTC_C) (Van de Weghe et al., 2005b), QTC-Shape (QTC_S) (Van de Weghe et al., 2004), and QTC-Network (QTC_N) (Bogaerts et al., 2004; Bogaert et al., 2006; Bogaert et al., 2007; Delafontaine et al., 2011b; Van de Weghe et al., 2004) were formed on the basis of QTC_B. In line with previous efforts, we incorporate the concept of QTC into knowledge discovery techniques for the detection of movement patterns of MPOs and for the exploration of similarities in movement patterns of MPOs. For this purpose, in Chapters 4-6, we attempt to employ the concept of *QTC conceptual animation*, a sequence of successive QTC relations, in order to obtain insight into the movement behaviour of MPOs and identify movement patterns, particularly repetitive ones. In each of these chapters, a similarity analysis technique is proposed in order to extract knowledge from the QTC conceptual animations. Unlike Chapters 4-6, which focus more on the movement behaviour of MPOs over durations of time, Chapters 7 and 8 investigate the movement behaviour of objects at time stamps.

When we develop algorithms in order to identify patterns or make predictions, it is of crucial importance to know what sorts of structures are likely to exist within the dataset under consideration (Mountain, 2005). The general meaning of a pattern (i.e. a regular occurrence in time) can be applied to a wide variety of cyclical natural phenomena having a periodicity or frequency of anything from microseconds to millions of years. As stated in the previous subsection, the effectiveness of different approaches in this thesis is tested using a data set of Samba dance representing some basic rhythmical movements of dancers. Pre-knowledge on the structure of dancer movements improves our ability to design techniques to extract more significant information. The question then arises, "what are the advantages of qualitative reasoning in this case?" Dance itself is a complicated process. If we also add the complexity of the collected dataset, then learning to digest this dynamic behaviour would be a demanding job. In qualitative reasoning (QR), precise numerical values or quantities are avoided, and qualitative values are used instead (Daintith, 2004). Qualitative spatio-temporal reasoning, a subset of QR, is based on a qualitative abstraction of the spatial and temporal aspects of the common-sense knowledge on which the human perspective of physical reality is based. As discussed earlier, QTC provides a context to qualitatively abstract the movement of MPOs and to express the relationships between them. And dance is the case that can satisfactorily illustrate the power of qualitative spatio-temporal reasoning of QTC. In dance, keeping control over the movements of interacting body parts determines the quality of the performances of dancers. This can be formulated through QTC and based on qualitative data (QTC binary relations between body parts). In this thesis, Chapters 4-7 deliver insights into the movement of dancers and quickly identify potential problems that warrant more detailed quantitative analysis (Chapter 8).

RQ 2: Is it possible to use QTC in the context of knowledge discovery from movement?

There are many examples in which detecting patterns are useful such as finding motifs in DNA sequences which have a biological significance or extracting specific patterns in speech used in speech processing. Given the importance of this issue, which has led many researchers to attempt to find patterns from different research domains like bioinformatics, speech processing, and image processing (for example see (Gilbert & Viksna, 1999; Hsu et al., 2001; Huang & Yu, 1999; Kovar & Gleicher, 2004; Laptev et al., 2005; Li & Holstein, 2002; Park & Glass, 2005; Qu et al., 1998; Wu et al., 2004)), we pursue this goal to discover patterns in the movement of MPOs. Among various types of patterns which may be observed in the movement of MPOs (a comprehensive classification of movement patterns can be found in (Dodge et al., 2008)), we focus on repetitive behaviour (i.e. frequent patterns) of moving objects. Frequent patterns reflecting regular behaviour of moving objects is an intensively studied topic (e.g. (Cao et al., 2007; Giannotti et al., 2007; Laube & Imfeld, 2002; Li, 2013; Li et al., 2010a)). Unlike most previous works that examine the behaviour of moving objects individually to detect frequent movement patterns, we aim to discover frequent patterns from the interactions of moving objects. We detect repetition in the behaviour of interacting moving objects based on low-level qualitative data. In most previous studies, the most accurate information is provided as an input to be considered in the detection of patterns. However, in this study we show how low-level qualitative data can be appropriately employed in finding frequent patterns. In response to this research question, we employ the qualitative trajectory calculus (QTC) to detect repetitive patterns on the basis of some mathematical functions that calculate similarity in QTC relations of moving objects. We focus more on frequent patterns manifested in a number of ways, ranging from simple repeated movements of the hands of a dancer to complex movements of football players during a match. In each chapter of this thesis (4-8), a section is devoted to introduce a specific similarity measure used to retrieve and detect periodic patterns from the interactions of moving objects.

RQ 3: How do we appropriately employ visualisation techniques in the analysis of movement data?

Data visualisation techniques allow us to form multi-dimensional representations of data that can be easily interpreted to gain knowledge and insights into those data sets (Soukup & Davidson, 2002). Ware (2004) pointed out some of the advantages of data visualisation such as its power in comprehension of enormous and complex data associated with errors and artifacts. The process of data visualisation includes data collection and storage, preprocessing and transformation, displaying, and perception (Ware, 2004). Accordingly, “raw” data collections need to undergo preprocessing, in the form of data cleaning and data transformation (Larose, 2005). Then, they are transformed into something that can be displayed and understood. With data visualisation and visual data mining tools and techniques, one can identify the interesting (nontrivial, implicit, perhaps previously unknown and potentially useful) information or patterns (Soukup & Davidson, 2002).

As a branch of this research area, visualisation of movement data has been an active topic for many years. Some of the key works exploring the visualisation of movement data include (Andrienko et al., 2012; Andrienko et al., 2007; Andrienko et al., 2008; Enguehard et al., 2013; Hägerstraand, 1970; Kwan, 2000; Randell et al., 1992; Ren & Kwan, 2007; Rinzivillo et al., 2008; Shamoun-Baranes et al., 2012; Willems et al., 2009; Zeng et al., 2013). Given incremental improvement of visual exploration tools and techniques, the level of understanding of movement datasets is increasing over time. Many of these advancements rely on direct depiction of data rather than summarisation or pattern extraction (Andrienko et al., 2008; Enguehard et al., 2013). Direct depiction of movement data, in most cases, provides only a general overview of the entire dataset. This limitation calls for other complementary techniques such as computer-aided synthesis (e.g. categorisation or pattern extraction) and filtering (either interactively or automatically) (Enguehard et al., 2013).

The current study emphasises the importance of visualisation to represent and reason about movement patterns in movement data. In this regard, visual analysis has been identified as an essential tool for data exploration, utilising human vision to interpret patterns in data (MacEachren & Kraak, 2001). A well-structured survey of the state of the art in visual analysis regarding the analysis of movement data is given in (Andrienko & Andrienko, 2013). In agreement with (Andrienko & Andrienko, 2007), visualisation techniques alone are not adequate to conduct a broad analysis of dynamic phenomena and, therefore, a compromise between computational methods and visualisation techniques are suitable. We offer some appropriate visualisation techniques to visualise and straightforwardly detect the repetitive movement patterns of multiple objects. The proposed visualisation techniques are either close enough to human perception or effectively fit with computational analysis methods.

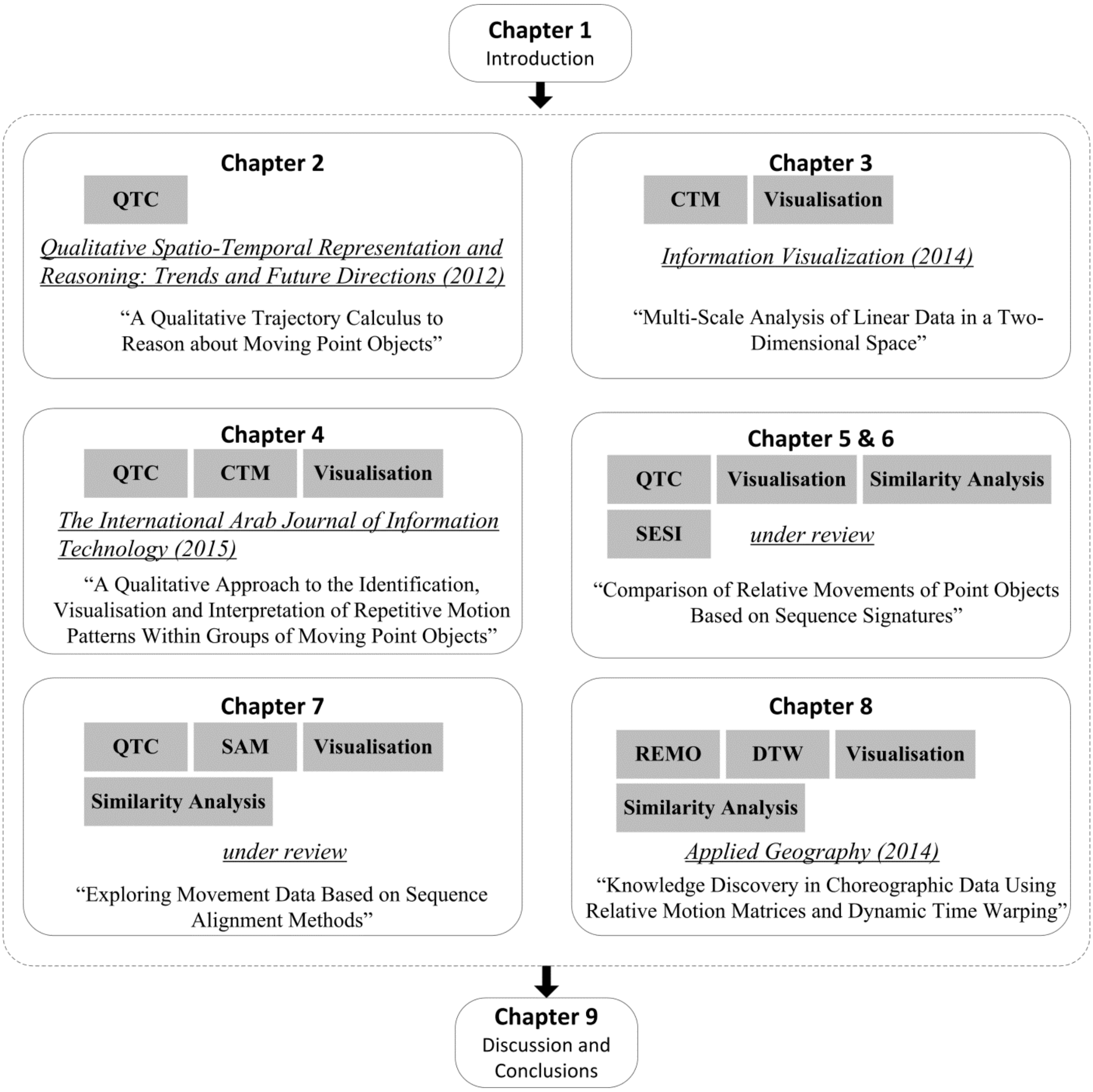
In Chapter 4, we aim to contribute to the development of the concept of the Continuous Triangular Model (CTM). In previous efforts, the use of CTM was demonstrated in reasoning about imperfect intervals and visual analytics (i.e. analytical reasoning facilitated by interactive visual interfaces) (Qiang et al., 2010; Qiang et al., 2012). We apply CTM to display the degrees of similarity between movement patterns. In fact, we employ CTM as an exploratory method to delve into the movement data (i.e. dance dataset), identifying interesting relations and interactions between moving objects, and detecting regular and non-regular movement patterns. In Chapter 5, we propose a visualisation technique called *sequence signature* to map patterns of the interactions between two moving objects in an indexed raster space. We further develop this concept to visualise patterns of multiple objects in Chapter 6. Although the focus of Chapters 7 and 8 is more on examining similarity in movement patterns, the employed methods for this purpose are integrated with visualisation techniques. In Chapter 7, the Sequence Alignment Method (SAM) is applied to discover matched and mismatched patterns of dancers. Visualised aligned patterns provide added value to interpret the results. Chapter 8 proposes two approaches, namely REMO (RElative MOtion) and

DTW (Dynamic Time Warping) to study motion attributes of moving objects such as speed and motion azimuth. The visual exploration of these methods is the most powerful incentive to investigate movement patterns of moving objects.

1.4 Thesis Outline

The thesis consists of nine chapters. Chapters 2 to 8 are the substantial part of the thesis, comprising a collection of academic papers that have been published in international peer-reviewed journals, submitted or are in preparation for submission. The papers on which Chapters 2 and 3 are base, are mainly written by the first author. Based on the previously mentioned research questions, each chapter of this manuscript is organised to answer them. The chapters are not grouped because some of them may partly answer all research questions, whereas other chapters respond to a specific question. There are strong links between the chapters presented in this manuscript. The general outline of the thesis is presented in Table 1-1. In order to allow someone to read these papers smoothly and independently, there are some unavoidable overlaps in the individual chapters with regard to the literature reviews and the description of the research design and methodology. The last chapter summarises the main findings of our study and returns to the three research questions to reflect on our contributions and proposes avenues for future research.

Table 1-1: General outline of the thesis



References

Agarwal, P. K., Arge, L., & Erickson, J. (2003). Indexing moving points. *Journal of Computer and System Sciences, 66* (1), 207-243.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Alt, H., & Guibas, L. J. (1999). Discrete geometric shapes: Matching, interpolation, and approximation. In: J.-R. Sack & J. Urrutia (Eds.), *Handbook of Computational Geometry* (pp. 121-153).

Andrienko, G., & Andrienko, N. (2008). Spatio-temporal aggregation for visual analysis of movements. *Proceedings* of *IEEE Symposium on Visual Analytics Science and Technology (VAST'08)* (pp. 51-58).

Andrienko, G., Andrienko, N., Bak, P., Keim, D., & Wrobel, S. (2013). *Visual Analytics of Movement*. Berlin Heidelberg: Springer Verlag.

Andrienko, G., Andrienko, N., Burch, M., & Weiskopf, D. (2012). Visual analytics methodology for eye movement studies. *IEEE Transactions on Visualization and Computer Graphics, 18* (12), 2889-2898.

Andrienko, G., Andrienko, N., & Wrobel, S. (2007). Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter, 9* (2), 38-46.

Andrienko, G. L., Andrienko, N. V., Dykes, J., Fabrikant, S. I., & Wachowicz, M. (2008). Geovisualization of dynamics, movement and change: Key issues and developing approaches in visualization research. *Information visualization, 7* (3-4), 173-180.

Andrienko, N., & Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. *Cartographica: The International Journal for Geographic Information and Geovisualization, 42* (2), 117-138.

Andrienko, N., & Andrienko, G. (2012). A visual analytics framework for spatio-temporal analysis and modelling. *Data Mining and Knowledge Discovery*, 1-29.

Andrienko, N., & Andrienko, G. (2013). Visual analytics of movement: An overview of methods, tools and procedures. *Information visualization, 12* (1), 3-24.

Basiri, A., & Amirian, P. (2014). Automatic point of interests detection using spatio-temporal data mining techniques over anonymous trajectories. *Proceedings of Computational Science and Its Applications (ICCSA)* (pp. 185-198).

Bevilacqua, F., Naugle, L., & Valverde, I. (2001). Virtual dance and music environment using motion capture. *Proceedings of IEEE Multimedia Technology and Applications Conference (MTAC)*. Irvine, CA.

Bogaert, P. (2008). *A Qualitative Calculus for Moving Point Objects Constrained by Networks*. Ghent University, Ghent.

Bogaert, P., Van de Weghe, N., and De Maeyer, P. (2004). Description, definition and proof of a qualitative state change of moving objects along a road network. In: M. Raubal, A. Sliwinski, and W. Kuhn (Eds.). *Geoinformation and Mobility - Geoinformation and Mobility (GI-days 2004)*, (pp. 239-248).

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., and De Maeyer, P. (2006). Reasoning about moving point objects on networks. In: M. Raubal, H. J. Miller, A. U. Frank, and F. Goodchild (Eds). *Proceedings of the 4th International Conference on Geographic Information Science (GIScience)*(pp. 29-32).

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., & De Maeyer, P. (2007). The qualitative trajectory calculus on networks. *Spatial Cognition V: Reasoning, Action, Interaction, 4387*, 20-38.

Brinkhoff, T. (2002). A framework for generating network-based moving objects. *Geoinformatica, 6* (2), 153-180.

Buchin, K., Buchin, M., Gudmundsson, J., Löffler, M., & Luo, J. (2011a). Detecting commuting patterns by clustering subtrajectories. *International Journal of Computational Geometry & Applications, 21* (03), 253-282.

Buchin, K., Buchin, M., van Kreveld, M., & Luo, J. (2011b). Finding long and similar parts of trajectories. *Computational Geometry, 44* (9), 465-476.

Buzan, D., Sclaroff, S., & Kollios, G. (2004). Extraction and clustering of motion trajectories in video. *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)* (Vol. 2, pp. 521-524).

Cao, H., Mamoulis, N., & Cheung, D. W. (2007). Discovery of periodic patterns in spatiotemporal sequences. *IEEE Transactions on Knowledge and Data Engineering, 19* (4), 453-467.

Castellano, G., Bresin, R., Camurri, A., & Volpe, G. (2007). Expressive control of music and visual media by full-body movement. *Proceedings of the 7th International Conference on New Interfaces for Musical Expression* (pp. 390-391).

Claramunt, C,. & Jiang, B. (2001), An integrated representation of spatial and temporal relationships, *Geographical Systems*, *3*(4), 154-159.

Cohn, A. G., Bennett, B., Gooday, J., & Gotts, N. M. (1997). Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica, 1* (3), 275-316.

Cohn, A. G., & Renz, J. (2008). Qualitative spatial representation and reasoning. *Foundations of Artificial Intelligence, 3*, 551-596.

Cristani, M., Cohn, A. G., & Bennett, B. (2000). Spatial locations via morpho-mereology. In: A. G. Cohn, F. Giunchiglia & B. Selman (Eds.). *Proceedings of the 7th Conference on Principles of Knowledge Representation and Reasoning (KR)* (pp. 15-25).

Daintith, J. (2004). *A Dictionary of Computing*. Oxford University Press.

Delafontaine, M. (2012). *Modelling and Analysing Moving Objects and Travelling Subjects: Bridging Theory and Practice.* Ghent University, Ghent.

Delafontaine, M., Bogaert, P., Cohn, A. G., Witlox, F., De Maeyer, P., & Van de Weghe, N. (2011a). Inferring additional knowledge from QTC_N relations. *Information Sciences, 181* (9), 1573-1590.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011b). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Demšar, U., & Virrantaus, K. (2010). Space–time density of trajectories: Exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science, 24* (10), 1527-1542.

Dodge, S. (2011). *Exploring Movement Using Similarity Analysis.* University of Zurich, Zurich.

Dodge, S., Laube, P., & Weibel, R. (2012). Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science, 26* (9), 1563-1588.

Dodge, S., Weibel, R., & Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems, 33* (6), 419-434.

Dodge, S., Weibel, R., & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information visualization, 7* (3-4), 240-252.

Du Mouza, C., & Rigaux, P. (2005). Mobility patterns. *Geoinformatica, 9* (4), 297-319.

Egenhofer, M. J., & Franzosa, R. D. (1991). Point-set topological spatial relations. *International Journal of Geographical Information System, 5* (2), 161-174.

El-Nasr, M. S., & Vasilakos, A. V. (2008). DigitalBeing - using the environment as an expressive medium for dance. *Information Sciences, 178* (3), 663-678.

Enguehard, R. A., Hoeber, O., & Devillers, R. (2013). Interactive exploration of movement data: A case study of geovisual analytics for fishing vessel analysis. *Information visualization, 12* (1), 65-84.

Erwig, M., Gu¨ ting, R. H., Schneider, M., & Vazirgiannis, M. (1999). Spatio-temporal data types: An approach to modelling and querying moving objects in databases. *Geoinformatica, 3* (3), 269-296.

Flety, E. (2005). The wise box: a multi-performer wireless sensor interface using WiFi and OSC. *Proceedings of the 2005 conference on New interfaces for musical expression* (pp. 266-267). National University of Singapore, Vancouver, Canada.

Freksa, C. (1992a). Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54* (1), 199-227.

Freksa, C. (1992b). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Fu, Z., Hu, W., & Tan, T. (2005). Similarity based vehicle trajectory clustering and anomaly detection. *Proceedings of IEEE International Conference on Image Processing (ICIP'05)* (Vol. 2, pp. 602-605).

Giannotti, F., Nanni, M., Pinelli, F., & Pedreschi, D. (2007). Trajectory pattern mining. *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 330-339). San Jose, CA.

Giannotti, G., & Pedreschi, D. (2008). *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*. Berlin Heidelberg: Springer-Verlag.

Gilbert, D., & Viksna, J. (1999). Pattern discovery methods for protein topology diagrams. *Proceedings of German Conference on Bioinformatics* (pp. 194-196).

Goodchild, M. F., Egenhofer, M. J., Kemp, K. K., Mark, D. M., & Sheppard, E. (1999). Introduction to the Varenius project. *International Journal of Geographical Information Science, 13* (8), 731-745.

Gottfried, B., Hans, W. G., & Sebastian, H. (2006). Spatiotemporal reasoning for smart homes. *Designing Smart Homes* (pp. 16-34).

Grassi, G. P., Santini, T., Lovecchio, N., Turci, M., Ferrario, V. F., & Sforza, C. (2005). Spatiotemporal consistency of trajectories in gymnastics: A three-dimensional analysis of Flic-Flac. *International Journal of Sports Medicine, 26* (2), 134-138.

Griffith, N., & Fernstrom, M. (1998). LiteFoot: A floor space for recording dance and controlling media. *Proceedings of International Computer Music Conference* (pp. 475-481).

Gudmundsson, J., van Kreveld, M., & Speckmann, B. (2007). Efficient detection of patterns in 2D trajectories of moving points. *Geoinformatica, 11* (2), 195-215.

Guesgen, H. W., & Marsland, S. (2010). Spatio-temporal reasoning and context awareness. *Handbook of Ambient Intelligence and Smart Environments* (pp. 609-634).

Hägerstraand, T. (1970). What about people in regional science? *Papers in Regional Science, 24* (1), 7-24.

Haritaoglu, I., Harwood, D., & Davis, L. S. (2000). W^4: Real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22* (8), 809-830.

Héas, P., & Mihai, D. (2005). Modelling trajectory of dynamic clusters in image time-series for spatio-temporal reasoning. *IEEE Transactions on Geoscience and Remote Sensing, 43*(7), 1635-1647.

Hornsby, K., & Egenhofer, M. (2002). Modelling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence*, *36* (1-2), 177-194.

Hromin, D., Chladil, M., Vanatta, N., Naumann, D., Wetzel, S., Anjum, F., & Jain, R. (2003). CodeBLUE: A Bluetooth interactive dance club system. *Proceedings of the Global Telecommunications Conference (GLOBECOM ’03)* (Vol. 5, pp. 2814-2818).

Hsu, J.-L., Liu, C.-C., & Chen, A. L. (2001). Discovering nontrivial repeating patterns in music data. *IEEE Transactions on Multimedia, 3* (3), 311-325.

Huang, Y.-W., & Yu, P. S. (1999). Adaptive query processing for time-series data. *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 282-286).

Imfeld, S. (2000). *Time, Point and Space - Towards a Better Analysis of Wildlife Data in GIS.* University of Zurich, Zurich.

Jensen, C. S., Lin, D., & Ooi, B. C. (2007). Continuous clustering of moving objects. *IEEE Transactions on Knowledge and Data Engineering, 19* (9), 1161-1174.

Johnstone, E. (1991). A MIDI foot controller - The PodoBoard. *Proceedings of the International Computer Music Conference* (pp. 123-126).

Kisilevich, S., Mansmann, F., Nanni, M., & Rinzivillo, S. (2010). Spatio-temporal clustering. *Data Mining and Knowledge Discovery Handbook*, 855-874.

Kovar, L., & Gleicher, M. (2004). Automated extraction and parameterization of motions in large data sets. *Proceedings of the ACM Transactions on Graphics (TOG)* (Vol. 23, pp. 559-568).

Kwan, M.-P. (2000). Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. *Transportation Research Part C: Emerging Technologies, 8* (1), 185-203.

Laptev, I., Belongie, S. J., Perez, P., & Wills, J. (2005). Periodic motion detection and segmentation via approximate sequence alignment. *Proccedings of the 10th IEEE International Conference on Computer Vision (ICCV 2005)* (Vol. 1, pp. 816-823).

Larose, D. T. (2005). *Discovering knowledge in data: An introduction to data mining* (1st ed.). Wiley.

Laube, P., Duckham, M., & Wolle, T. (2008). Decentralized movement pattern detection amongst mobile geosensor nodes. In: T. Cova, H. J. Miller, K. Beard, A. Frank & M. Goodchild (Eds.), *Proceedings of the 5th International Conference on Geographic Information Science* (pp. 199-216).

Laube, P., & Imfeld, S. (2002). Analysing relative motion within groups of trackable moving point objects. In: M. J. Egenhofer & D. M. Mark (Eds.), *Geographic Information Science*. (pp. 132-144). Boulder: Springer.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Lee, J.-G., Han, J., & Whang, K.-Y. (2007). Trajectory clustering: A partition and group framework. *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data* (pp. 593-604).

Leman, M., & Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in Samba and Charleston. *Music Perception: An Interdisciplinary Journal, 28* (1), 71-91.

Li, B., & Holstein, H. (2002). Recognition of human periodic motion: A frequency domain approach. *Proceedings of the 16th International Conference on Pattern Recognition* (Vol. 1, pp. 311-314).

Li, Y., Han, J., & Yang, J. (2004). Clustering moving objects. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 617-622).

Li, Z. (2013). *Mining Periodicity and Object Relationship in Spatial and Temporal Data.* University of Illinois, Urbana-Champaign.

Li, Z., Ding, B., Han, J., Kays, R., & Nye, P. (2010a). Mining periodic behaviour for moving objects. *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1099-1108).

Li, Z., Lee, J.-G., Li, X., & Han, J. (2010b). Incremental clustering for trajectories. *Database Systems for Advanced Applications* (pp. 32-46). Berlin Heidelberg: Springer.

Lin, B., & Su, J. (2008). One way distance: For shape based similarity search of moving object trajectories. *Geoinformatica, 12* (2), 117-142.

MacEachren, A. M., & Kraak, M.-J. (2001). Research challenges in geovisualization. *Cartography and Geographic Information Science, 28* (1), 3-12.

McArdle, G., Demšar, U., van der Spek, S., & McLoone, S. (2013). Interpreting pedestrian behaviour by visualising and clustering movement data. In: S. H. L. Liang, X. Wang & C. Claramunt (Eds.), *Web and Wireless Geographical Information Systems*. (pp. 64-81). Berlin Heidelberg: Springer.

McElligott, L., Dillon, M., Leydon, K., Richardson, B., Fernström, M., & Paradiso, J. (2002). 'ForSe FIElds'-force sensors for interactive environments. *UbiComp 2002: Ubiquitous Computing*, (pp. 168-175). Berlin Heidelberg: Springer.

Menache, A. (2000). *Understanding motion capture for computer animation and video games*. Elsevier.

Menon, V., Jayaraman, B., & Govindaraju, V. (2013). Enhancing biometric recognition with spatio-temporal reasoning in smart environments. *Personal and Ubiquitous Computing*, *17*(5), 987-998.

Merz, S., Wirsing, M., Zappe, J. (2003). A spatio-temporal logic for the specification and refinement of mobile systems. In: M. Pezzé (Ed.). *FASE 2003* (Vol. 2621, pp. 87-101).

Miller, H. J., & Han, J. (2009). *Geographic Data Mining and Knowledge Discovery* (2nd ed.). CRC Press.

Mirabella, O., Raucea, A., Fisichella, F., & Gentile, L. (2011). A motion capture system for sport training and rehabilitation. *Proceedings of the 4th International Conference on Human System Interactions (HSI)* (pp. 52-59).

Mountain, D. M. (2005). *An Investigation of Individual Spatial Behaviour and Geographic Filters for Information Retrieval.* City University, London.

Muller, Ph. (1998). A qualitative theory of motion based on spatiotemporal primitives. In: A. G. Cohn, L. Schubert & S. Shapiro (Eds.), *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR)* (pp. 131-142).

Nanni, M., & Pedreschi, D. (2006). Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems, 27* (3), 267-289.

Ng, K. C. (2004). Music via motion: Transdomain mapping of motion and sound for interactive performances. *Proceedings of the IEEE 92*(4), 645-655.

Noyon, V., Claramunt, C., & Devogele, D., (2007), A relative representation of trajectories in geographical spaces, *Geoinformatica*, *4*(11), 479-496.

Paradiso, J., Abler, C., Hsiao, K., & Reynolds, M. (1997). The magic carpet: Physical sensing for immersive environments. *Proceedings of the Human Factors in Computing Systems* (pp. 277-278).

Paradiso, J. A., Hsiao, K., & Hu, E. (1999). Interactive music for instrumented dancing shoes. *Proceedings of the International Computer Music Conference*, Beijing, China.

Park, A., & Glass, J. R. (2005). Towards unsupervised pattern discovery in speech. *Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding* (pp. 53-58).

Pelekis, N., Andrienko, G., Andrienko, N., Kopanakis, I., Marketos, G., & Theodoridis, Y. (2012). Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems, 38* (2), 343-391.

Pelekis, N., Kopanakis, I., Marketos, G., Ntoutsi, I., Andrienko, G., & Theodoridis, Y. (2007). Similarity search in trajectory databases. *Proceedings of the 14th International Symposium on Temporal Representation and Reasoning (TIME 2007)* (pp. 129-140).

Petitjean, F., Kurtz, C., Passat, N., & Gançarski, P. (2012). Spatio-temporal reasoning for the classification of satellite image time series. *Pattern Recognition Letters 33*(13), 1805-1815.

Pfoser, D., & Theodoridis, Y. (2003). Generating semantics-based trajectories of moving objects. *Computers, Environment and Urban Systems, 27* (3), 243-263.

Prasad Sistla, A., Wolfson, O., Chamberlain, S., & Dao, S. (1997). Modelling and querying moving objects. *Proceedings of the 13th International Conference on Data Engineering* (pp. 422-432).

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P., & Van de Weghe, N. (2010). Modelling imperfect time intervals in a two-dimensional space. *Control and Cybernetics, 39* (4), 983-1010.

Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P., & Van de Weghe, N. (2012). Interactive analysis of time intervals in a two-dimensional space. *Information visualization, 11* (4), 255-272.

Qu, Y., Wang, C., & Wang, X. S. (1998). Supporting fast search in time series for movement patterns in multiple scales. *Proceedings of the 7th International Conference on Information and Knowledge Management* (pp. 251-258).

Rajagopalan, R. (1995). *Qualitative reasoning about dynamic change in the spatial properties of a physical system.* University of Texas, Austin.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In: B. Nebel, W. Swartout & C. Rich (Eds.), *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR)* (Vol. 92, pp. 165-176).

Ren, F., & Kwan, M. P. (2007). Geovisualization of human hybrid activity-travel patterns. *Transactions in GIS, 11* (5), 721-744.

Rinzivillo, S., Pedreschi, D., Nanni, M., Giannotti, F., Andrienko, N., & Andrienko, G. (2008). Visually driven analysis of movement data by progressive clustering. *Information Visualization, 7* (3-4), 225-239.

Šaltenis, S., Jensen, C. S., Leutenegger, S. T., & Lopez, M. A. (1999). Indexing the positions of continuously moving objects. *Proceedings of the ACM SIGMOD International Conference on Management of data* (pp. 331-342).

Shamoun-Baranes, J., van Loon, E. E., Purves, R. S., Speckmann, B., Weiskopf, D., & Camphuysen, C. (2012). Analysis and visualization of animal movement. *Biology Letters, 8* (1), 6-9.

Soukup, T., & Davidson, I. (2002). *Visual data mining: Techniques and tools for data visualization and mining*. Wiley.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data & Knowledge Engineering, 65* (1), 126-146.

Srinivasan, P., Birchfield, D., Qian, G., & Kidan, A. (2005). A pressure sensing floor for interactive media applications. *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (Vol. 265, pp. 278-281).

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Van de Weghe, N., Cohn, A. G., Bogaert, P., and De Maeyer, P. (2004). Representation of moving objects along a road network. In: S. A. Brandt (Ed.). *Proceedings of 12th International Conference on Geoinformatics Geospatial Information Research: Bridging the Pacific and Atlantic (*pp. 187-197).

Van de Weghe, N., Cohn, A. G., Maeyer, P. D., & Witlox, F. (2005a). Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems with Applications, 29* (4), 977-983.

Van de Weghe, N., Kuijpers, B., Bogaert, P., & De Maeyer, P. (2005b). A qualitative trajectory calculus and the composition of its relations. In: M. A. Rodriquez, I. F. Cruz, M. J. Egenhofer & S. Levashkin (Eds.), *Geospatial Semantics*. (pp. 60-76) Berlin Heidelberg: Springer-Verlag.

Van de Weghe, N., Maddens, R., Bogaert, P., Brondeel, M., & De Maeyer, P. (2004). Qualitative analysis of polygon shape-changes. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 7, pp. 4157 - 4159).

Versichele, M., Neutens, T., Delafontaine, M., & Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography, 32* (2), 208-220.

Ware, C. (2004). *Information Visualization: Perception for Design* (2nd ed.). Elsevier.

Willems, N., Van De Wetering, H., & Van Wijk, J. J. (2009). Visualization of vessel movements. *Computer Graphics Forum, 28* (3), 959-966.

Wilson, C. (2008). Activity patterns in space and time: calculating representative Hagerstrand trajectories. *Transportation, 35* (4), 485-499.

Wolter, F., & Zakharyaschev, M. (2000). Spatio-temporal representation and reasoning based on RCC-8. In: A. G. Cohn, F. Giunchiglia & B. Selman (Eds.). *Proceedings of the 7th Conference on Principles of Knowledge Representation and Reasoning (KR)* (pp. 3-14).

Worboys, M., & Duckham, M. (2004). *GIS: A Computing Perspective*. CRC Press.

Wu, H., Salzberg, B., & Zhang, D. (2004). Online event-driven subsequence matching over financial data streams. *Proceedings of the ACM SIGMOD International Conference on Management of Data* (pp. 23-34).

Yanagisawa, Y., & Satoh, T. (2006). Clustering multidimensional trajectories based on shape and velocity. *Proceedings of the 22nd International Conference on Data Engineering Workshops*.

Yoonji, K., Donggi, J., Sehwi, P., Jumin, C., Taewoo, K., & Seny, L. (2008). The shadow dancer: A new dance interface with interactive shoes. *Proceedings of the International Conference on Cyberworlds* (pp. 745-748).

Zadeh, L.A. (2002). From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. *International Journal of Applied Mathematics and Computer Science, 12* (3), 307-324.

Zeng, W., Fu, C. W., Arisona, S. M., & Qu, H. (2013). Visualising interchange patterns in massive movement data. *Computer Graphics Forum, 32* (3), 271-280.

Zhang, Q., & Lin, X. (2004). Clustering moving objects for spatio-temporal selectivity estimation. *Proceedings of the 15th Australasian Database Conference* (Vol. 27, pp. 123-130).

2

The Qualitative Trajectory Calculus

The reason why we reason seems reasonable. Nico Van de Weghe

Modified From: Matthias Delafontaine, Seyed Hossein Chavoshi, Anthony G. Cohn, and Nico Van de Weghe (2012). A Qualitative Trajectory Calculus to Reason about Moving Point Objects. In: S. M. Hazarika (Ed.), Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions, Hershey, PA, USA., pp. 147-167.

2 The Qualitative Trajectory Calculus

Abstract: A number of qualitative calculi have been developed that reason about space and time. A recent trend has been the emergence of integrated spatio-temporal calculi in order to deal with dynamic phenomena such as motion. In 2004, Van de Weghe introduced the Qualitative Trajectory Calculus (QTC) as a qualitative calculus to represent and reason about moving objects. This chapter presents a general overview of the principal theoretical aspects of QTC. It shows how QTC deals with important reasoning concepts and how the calculus can be employed in order to represent raw moving object data.

2.1 Introduction

Reasoning about spatial and temporal information takes a central place in human daily life. A number of qualitative calculi have been developed to represent and reason about spatial or temporal configurations. Most of them focus on one of the two domains, whereas a few are true spatio-temporal calculi that deal with spatio-temporal phenomena. One such calculus is the Qualitative Trajectory Calculus, which will be referred to as QTC. QTC is a qualitative calculus to reason about a specific spatio-temporal phenomenon: moving objects.

The remainder of this chapter is structured as follows. First, relevant background issues are discussed. Second, some general characteristics of QTC are explained and a brief overview of all QTC calculi that have been elaborated so far is given. The most fundamental QTC calculus, QTC_B, is then presented in detail. The following sections discuss representing and reasoning with QTC, as well as how QTC can be extended. An application section follows in order to highlight the potential of implementing QTC in information systems. The final section draws conclusions.

2.2 Background

In Artificial Intelligence, several qualitative calculi exist to reason about either spatial or temporal information, the most well-known being Allen's Interval Calculus (Allen, 1983). According to Wolter & Zakharyaschev (2000), an apparent and natural step is to combine both spatial and temporal formalisms in order to reason about spatio-temporal phenomena. Motion is a key research area in GISciences. Note that motion is an inherently spatio-temporal phenomenon (Peuquet, 2001). Dealing with motion is essential to spatial and geographical information systems, where an evolution from static to dynamic formalisms and representations has been made. A specific type of motion is associated with moving objects, i.e. objects whose position moves through space in time.

In the past decade, the modelling of moving objects has been a hot topic in fields such as GIScience, Artificial Intelligence and Information Systems (Bitterlich et al., 2008). In qualitative reasoning, however, considerable work has focused on the formalisation of motion, or moving objects in particular. Some examples are Muller (2002), Ibrahim (2007), Hallot & Billen (2008), and Kurata & Egenhofer (2009). These approaches have in common that they rely on topological models such as the Region Connection Calculus (Randell et al., 1992) or the 9-Intersection model (Egenhofer & Franzosa, 1991). However, a general shortfall of topological models is their inability to further differentiate between *disjoint* relations. This makes their applicability to represent and reason about continuously moving objects questionable, as in many cases moving objects remain disjoint for most of the time. For instance, cars in a traffic situation are usually disjoint, apart from the exceptional case of an accident.

In order to overcome this inability, the Qualitative Trajectory Calculus (QTC), was proposed by Van de Weghe (2004). QTC provides a qualitative framework to represent and reason about moving objects which enables the differentiation of groups of disconnected objects. The development of QTC has been inspired by some major qualitative calculi: the Region Connection Calculus (Randell et al., 1992), the temporal Semi-Interval Calculus (Freksa, 1992a), and the spatial Double-Cross Calculus (Freksa, 1992b; Zimmerman & Freksa, 1996).

2.3 The Qualitative Trajectory Calculus

2.3.1 Simplifications

Information systems usually represent knowledge according to an underlying model of the real-world. To this end, QTC makes four simplifications (Figure 2-1). First and foremost, QTC considers the relation between only two objects at the same time, i.e. binary relations (*relational simplification*, Figure 2-1b), as is common in spatio-temporal reasoning (Cohn & Renz, 2007). Second, moving objects are spatially simplified into moving point objects or MPOs (*object simplification*, Figure 2-1c), as is common in GIScience and geoinformatics (Gudmundsson et al., 2004; Guting et al., 2000; Laube, 2005; Noyon et al., 2007). There are only two topological relations (*disjoint* and *equal*) between two MPOs. Since the relation between two *equal* MPOs is trivial, the third simplification in QTC is the restriction to *disjoint* MPOs (*topological simplification*). Finally, in order to understand the temporal dimension in depth, it is important to find out what happens at one time point. Hence, QTC relations are relations that hold at a particular time point (*temporal simplification*, Figure 2-1d).

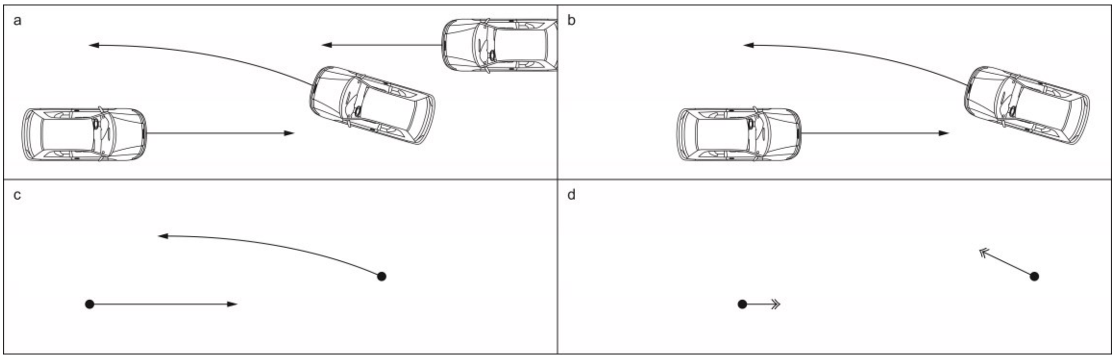


Figure 2-1: Simplification in QTC of a real-life situation (a) by taking cumulatively account of the relational simplification (b), the object simplification (c), and the temporal simplification (d) (simple arrows for trajectories, double arrows for instantaneous velocity vectors)

2.3.2 Continuity, Conceptual Neighbours, and Transitions

QTC assumes space and time, and thus the motion of objects, to be continuous. As a consequence, QTC relations change in time according to the laws of continuity. Along with continuity comes the important concept of conceptual neighbourhood as introduced by Freksa (1992b). Two QTC relations between the same pair of MPOs are *conceptual neighbours* if and only if these relations can directly follow each other through continuous motion of the MPOs, without the necessity for a third relation to hold at an intermediate point in time. A *transition* then denotes the continuous change of one relation into a conceptual neighbouring relation. Each transition thereby happens at a certain instant or point in time, which we will term a *transition instant*. A conceptual neighbourhood can be represented by a *conceptual neighbourhood diagram* (CND), i.e. a visualisation of a graph which nodes represent relations, and where two nodes are connected if they are conceptual neighbours of each other.

All QTC calculi are associated with a set of jointly exhaustive and pairwise disjoint (JEPD) base relations. Consequently, there is one and only one relation for each pair of coexisting MPOs at each time instant. In addition, due to continuity, the concurrent movement of two MPOs over a given time interval is uniquely mapped to a sequence of conceptually neighbouring base relations.

All QTC relations are formed by a tuple of labels (representing different primitive qualitative relations) that all have the same three-valued qualitative domain {–, 0, +}, which we will denote as *U* in the remainder of this chapter. A '0' symbol corresponds to a landmark value, and as Galton (2001) points out, this value always *dominates* both '–' and '+' values. Hence:

- A '0' must always last over a closed time interval (of which a time instant is a special case);
- A '–' / '+' must always last over an open time interval;
- Only transitions to or from '0' are possible (transitions from '–' / '+' to '+' / '–' are impossible) and transition instants always correspond with a '0' value.

Based on the notion of topological distance introduced by Egenhofer & Al-Taha (1992), the *conceptual distance* can be defined as a measure for the closeness of QTC relations (Van de Weghe & De Maeyer, 2005). We take the conceptual distance between '0' and another symbol to be one. This is the smallest conceptual distance, apart from zero, i.e. the distance between a symbol and itself. Since a direct transition is impossible, the conceptual distance between '–' and '+' is equal to two (one for '–' to '0' and one for '0' to '+'). The overall conceptual distance between two QTC relations can then be calculated by summing the conceptual distance over all relation symbols. For instance, for two QTC relations consisting of four symbols, the conceptual distance ranges from zero to eight.

2.4 Types of QTC

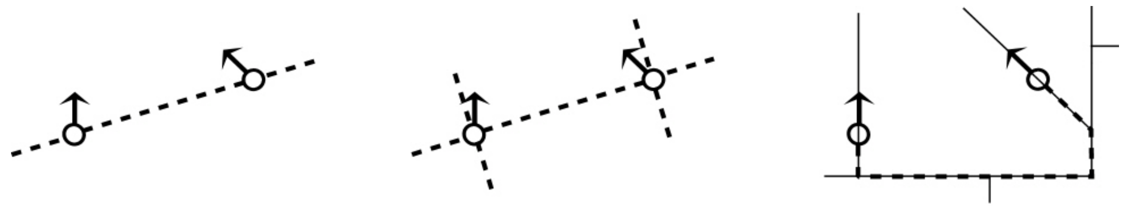
Due to the consideration of different spaces and frames of reference, the following types of QTC have been elaborated:

- Basic type – QTC_B (Van de Weghe et al., 2006), Figure 2-2a
- Double-Cross type – QTC_C (Van de Weghe et al., 2005), Figure 2-2b
- Network type – QTC_N (Bogaert et al., 2006), Figure 2-2c
- Shape type – QTC_S (Van de Weghe et al., 2005)

The QTC Basic (QTC_B) and the Double-Cross (QTC_C) types both deal with MPOs that have a free trajectory in an *n*-dimensional space. QTC_B relations are determined by referring to the Euclidian distance between two MPOs (Figure 2-2a). QTC_C relations on the other hand rely on the double cross, a concept introduced by Zimmerman & Freksa (1996), as a spatial reference frame (Figure 2-2b).

QTC_N (Network) focuses on the special case of MPOs which trajectories are constrained by a network, such as cars in a city. Since both the Euclidean distance and the double cross concepts ignore the spatial configuration of a potential underlying network, they are not well suited for QTC_N. Therefore, QTC_N relations rely on the shortest paths in the network between the considered MPOs (Figure 2-2c). In essence, QTC_N employs the philosophy of QTC_B in the context of a space constrained by a network.

Finally, QTC_S (Shape) employs the double cross concept in order to describe trajectory shapes or even arbitrary undirected polylines in a qualitative way. Thus, QTC_S deals with the relative configuration of a trajectory, rather than with the relation between MPOs. QTC_B will be only discussed in detail in this chapter.



(a) (b) (c)

Figure 2-2: Two MPOs represented in a typical two-dimensional QTC_B (a), QTC_C (b), and QTC_N (c) setting. The frame of spatial reference is represented by the dashed line.

2.5 QTC Basic (QTC_B)

A MPO is always characterised by an origin and a destination, whether explicit or implicit. Hence, a basic dichotomy concerning MPOs, perhaps the most fundamental one, is the distinction between *towards* and *away from* relations. This very generic idea underlies QTC_B where this binary relation is evaluated on the basis of Euclidean distance in an unconstrained n-dimensional space. In addition, also the relative speed between both objects can be taken into account. As mentioned earlier, QTC relations consist of qualitative symbols that share the threefold domain $U = \{-, 0, +\}$. QTC_B relations are constructed from the following relationships:

Assume: MPOs k and l and time point t

$k|t$ denotes the position of an MPO k at t

$d(u,v)$ denotes the Euclidean distance between two positions u and v

$\overrightarrow{v_k^t}$ denotes the velocity vector of k at t

$t_1 \prec t_2$ denotes that t_1 is temporally before t_2

A. Movement of k with respect to l at t (distance constraint):

−: k is moving towards l:

$$\exists t_1 \left(t_1 \prec t \wedge \forall t^- \left(t_1 \prec t^- \prec t \rightarrow d(k|t^-, l|t) > d(k|t, l|t)\right)\right) \wedge \exists t_2 \left(t \prec t_2 \wedge \forall t^+ \left(t \prec t^+ \prec t_2 \rightarrow d(k|t, l|t) > d(k|t^+, l|t)\right)\right) \quad \text{Eq. 2-1}$$

+: k is moving away from l:

$$\exists t_1 \left(t_1 \prec t \wedge \forall t^- \left(t_1 \prec t^- \prec t \rightarrow d(k|t^-, l|t) < d(k|t, l|t)\right)\right) \wedge \exists t_2 \left(t \prec t_2 \wedge \forall t^+ \left(t \prec t^+ \prec t_2 \rightarrow d(k|t, l|t) < d(k|t^+, l|t)\right)\right) \quad \text{Eq. 2-2}$$

0: k is stable with respect to l (all other cases)

B. Movement of l with respect to k at t (distance constraint), can be described as in A with k and l interchanged, and hence:

−: l is moving towards k Eq. 2-3

+: l is moving away from k Eq. 2-4

0: l is stable with respect to k (all other cases)

C. Relative speed of k with respect to l at t (speed constraint):

−: k is moving slower than l

$$\left|\overrightarrow{v_k^t}\right| < \left|\overrightarrow{v_l^t}\right| \quad \text{Eq. 2-5}$$

+: k is moving faster than l

$$\left|\overrightarrow{v_k^t}\right| > \left|\overrightarrow{v_l^t}\right| \quad \text{Eq. 2-6}$$

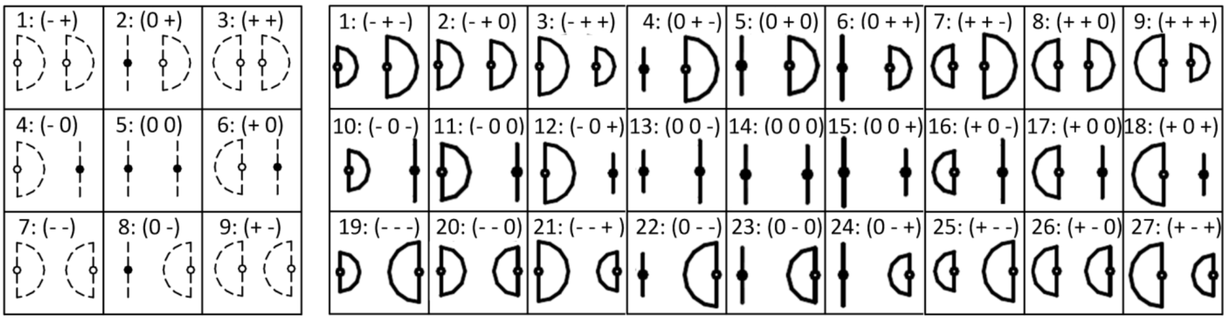
0: k and l are moving equally fast

$$\left|\overrightarrow{v_k^t}\right| = \left|\overrightarrow{v_l^t}\right| \quad \text{Eq. 2-7}$$

Two levels of QTC_B relations have been proposed: a first level QTC_{B1} that only considers the distance constraints (relationships A and B), and a second level QTC_{B2} taking account of the speed constraint (relationship C) as well. The resulting relation syntaxes are respectively the tuples $(A\ B)_{B1}$ and $(A\ B\ C)_{B2}$. Note that relationship C dually represents the relative speed of l with respect to k, and hence trivialises a fourth relationship. Relation icons for QTC_B are shown in Figure 2-3, where k is always on the left side, and l on the right side. The line segments and crescents represent potential motion areas. Note that their boundaries are open, and, for the crescents, the straight boundaries correspond to elements of another relation. A filled dot indicates that an

MPO might be stationary, whereas an open dot means that it must be moving. Dashed lines represent uncertain boundaries that follow from the ignorance of relative speed.

There are 9 (3^2) base relations in QTC_{B1} (Figure 2-3a). All these relations are possible in a one- or higher-dimensional space. QTC_{B2} on the other hand has 27 (3^3) base relations (Figure 2-3b), which are all possible in two- or higher-dimensional spaces. However, in a one-dimensional space, only 17 (63.0%) QTC_{B2} relations can occur. This reduction follows from a dependency between the distance constraints and the speed constraint in the case of a 1D space. In a 1D space, the direction of movement is always collinear with the direction of Euclidean distance, and hence a '0' in the distance constraints always corresponds to a stationary MPO. As a consequence, it is impossible for an MPO to be stationary and to have a higher speed than another MPO. In a two- or higher-dimensional space on the other hand, a '0' distance constraint does not necessarily indicate a stationary object, e.g. in the case of 'tangential motion' such as when one MPO is circling around the other MPO.



(a) (b)

Figure 2-3: QTC_{B1} relation icons (a) and QTC_{B2} relation icons (b)

2.6 Representing and Reasoning with QTC

QTC has been confronted with key concepts in qualitative reasoning. In this section, we will discuss two of these issues, respectively conceptual neighbourhood diagrams (CNDs) and incomplete knowledge.

2.6.1 Conceptual Neighbourhood Diagrams

As mentioned earlier, the construction of CNDs for QTC is based on the concepts of dominance (Galton, 2001) and conceptual distance. For an in depth description, we refer to Van de Weghe & De Maeyer (2005). CNDs for the QTC Basic calculi in 2D space are shown in Figure 2-4. For each link between conceptual neighbours the conceptual distance between the adjacent relations has been indicated.

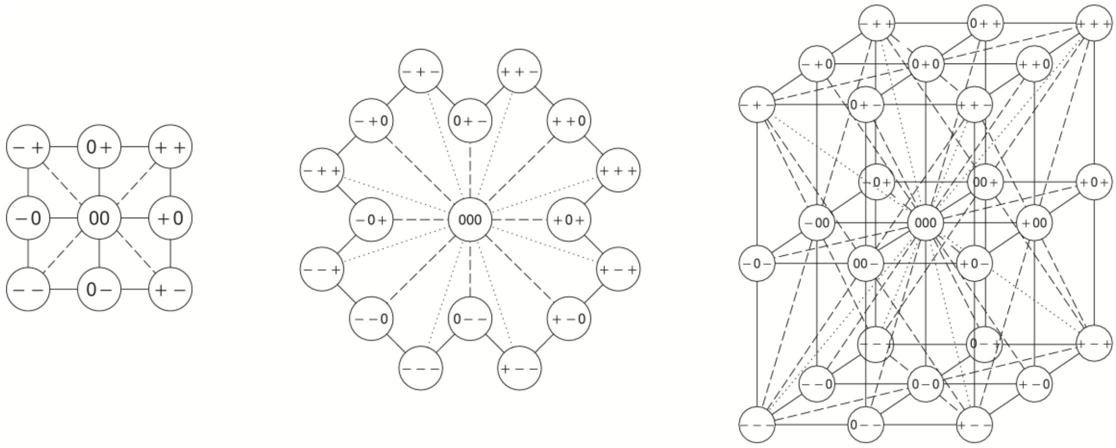
From the CNDs, we learn that, due to the laws of continuity, the conceptual neighbours of each particular relation constitute only a subset of base relations. This set comprises the candidate relations that may directly precede or follow the relation at hand in time, i.e. the set of possible transitions from/to this relation. This set of candidates is thereby highly limited when compared to the set of theoretical possibilities, as can be seen from Table 2-1. Note that each pair of conceptual neighbours R_1 and R_2 is associated with two transitions, i.e. a transition from R_1 to R_2, and its converse from R_2 to R_1. Similarly to a CND, a transition graph can be constructed with directed links to represent existing transitions.

Another notable finding is that all CNDs are completely symmetric with respect to the relation consisting solely of ‘0’ values. We call this symmetric and reflexive relation the *zero-relation*. Symmetry with respect to the zero-relation is due to the central position of ‘0’ in the qualitative set $U = \{-, 0, +\}$, as well as to the symmetry of conceptual neighbourhood for converse QTC relations.

Furthermore, every relation is a conceptual neighbour of the zero-relation (and vice versa), as is consistent with our intuition. For instance, it is highly reasonable that, whatever the relation between two MPOs at a certain moment, they may always become stationary the next moment, in which case their relation turns into the zero-relation.

Table 2-1: The number of base relations, transitions, theoretical combinations of base relations, and the ratio transitions / theoretical combinations for QTC Basic calculus.

QTC calculus	# spatial dimensions	# base relations	# transitions	# combinations	ratio
B_1	1+	9	32	72	44,4%
B_2	1	17	64	272	23,5%
	2+	27	196	702	27,9%



(a) (b) (c)

Figure 2-4: CNDs for QTC_{B1} in *n*-dimensional space (a), for QTC_{B2} in a one-dimensional space (b), and for QTC_{B2} in a two- or higher-dimensional space (c). The straight, dashed and dotted lines respectively represent the conceptual distances one, two and three.

2.6.2 Incomplete Knowledge

Not always everything has to be known about a situation to make inferences which are important for the issue at hand (Frank, 1996). Obviously, in these situations information may lack for offering complete answers to queries. However, '*a partial answer may be better than no answer at all.*' as Freksa (1992a, p. 203) argues. By abstracting away from the mass of metrical details, qualitative representations are much more appropriate for handling such incomplete knowledge, rather than quantitative approaches (Cristani et al., 2000).

As mentioned before, the development of the QTC has been inspired by some major QR calculi, especially the temporal Semi-Interval Calculus (Freksa, 1992a) and the spatial Double-Cross Calculus (Freksa, 1992b; Zimmermann & Freksa, 1996). Central in these theories is the specific attention to incomplete knowledge, and hence, one might expect QTC to be able to handle incomplete knowledge as well.

One kind of incomplete knowledge results from natural language expressions. Consider the expression "*k* is moving towards *l*, which is not slower than *k*". This expression can be represented in QTC, for instance by $(-\ U\ U_{+})_{B2}$ with $U_{+} = U \setminus \{+\}$. Hence, we obtain a union of six solutions. Interestingly, these solutions constitute a conceptual neighbourhood, i.e. they are mutually path-connected through conceptual neighbour relations when isolated from the complete CND of base relations (see Figure 2-4). According to Freksa (1992a), we achieve *coarse knowledge*, i.e. a kind of incomplete knowledge that allows to be represented by a conceptual neighbourhood of relations at a certain level of granularity. When relations between MPOs are perceived or described incompletely through natural language, the resulting knowledge will typically be coarse.

Whenever one expression may lead us to incomplete knowledge, multiple expressions can be combined in order to deduce finer knowledge. Table 2-2 gives an example of four expressions, each of which has a coarse result, for which the intersection results in complete knowledge. In addition, composition offers an appropriate inference mechanism to integrate expressions about three or more objects.

Table 2-2: Intersection of coarse solutions to obtain fine knowledge, with U_0 = U \(G\ et al.) (for more explanation, see (Van de Weghe 2004))

natural language expression	QTC_{B2} solution	integrated solution
“ *k* is moving towards *l* ”	$(-\ U\ U)_{B2}$	$(-\ U\ U)_{B2} \cap (U_0\ U_0\ U)_{B2} \cap (U + U)_{B2} \cap (U\ U\ 0)_{B2} = (- + 0)_{B2}$
“ *k* and *l* are moving along the same straight line ”	$(U_0\ U_0\ U)_{B2}$	
“ *l* is moving away from *k* ”	$(U + U)_{B2}$	
“ *l* is moving equally fast as *k* ”	$(U\ U\ 0)_{B2}$	

2.7 Extending QTC

Complex real-life motions go far beyond the earlier described simplifications applied in QTC. Can we relax these constraints? Obviously, not all simplifications can be ignored. Therefore, we now focus on how QTC can be extended, whilst still accepting the object simplification, i.e. the abstraction of moving objects to MPOs. In the remainder of this section we will discuss the respective and cumulative releases of the relational, temporal, and topological simplifications.

2.7.1 Multiple MPOs

The relations between multiple MPOs can be represented by means of a QTC cross table or matrix (Table 2-3). An element (i, j) in this matrix represents the QTC relation between MPOs i and j. A QTC matrix can be computed at each time point. The following compression rules and techniques can be used in order to reduce its size:

- The diagonal of the matrix can be excluded, as it is empty due to the topological constraint.
- Only the upper right (or lower left half) of the matrix has to be considered, as is gray shaded in bold in Table 2-3. The lower part of the matrix holds the converse relations of the upper part and vice versa and is therefore redundant.

Hence, for *n* objects, the number of elements can be reduced from n^2 to $(n^2-n)/2$. Note that research has been done in order to further simplify topological relations (Rodríguez et al., 2003) and simplifying temporal relations (Rodríguez, et al., 2004) over multiple elements. It could be interesting to combine both in order to simplify spatio-temporal relations, such as QTC relations. Thus, the number of elements in a QTC matrix could be further reduced so that it only contains relevant information, i.e. no redundancies.

Table 2-3: QTC_{B1} matrix for four MPOs *k*, *l*, *m*, and *n* at time *t*

t	*k*	*l*	*m*	*n*
k		--	-+	-0
l	--		++	+0
m	+-	++		-0
n	0-	0+	0-	

2.7.2 Multiple Time Points and Intervals

What if we consider QTC matrices at different time moments? According to the philosophy of qualitative reasoning, new relations only need to be calculated whenever transitions occur. As a consequence, it will be the most efficient to compute one initial matrix and to store only relations which have transitioned in all subsequent matrices.

2.7.3 Multiple Topological Relations

QTC does not distinguish topological relations, and might hence be complemented by topological calculi. As mentioned earlier, point objects only have two topological relations: *disjoint* and *equal*. Though QTC is developed to reason about *disjoint* objects, this constraint might be relaxed. Note that in case of *equal* MPOs, we will always obtain zero-relations.

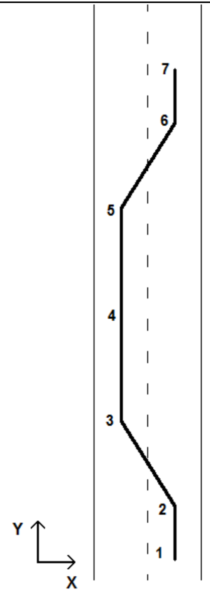
2.8 Example Case

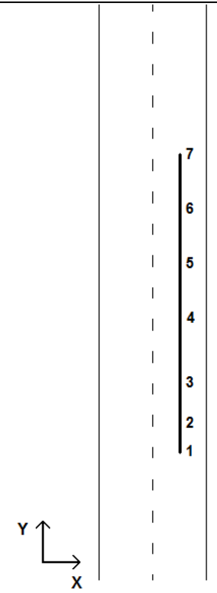
This section discusses an example application of QTC in one of the major domains of applied science that in essence deals with objects moving in a geographical space, namely transportation research. Ever since their invention, cars have been a focus of research for numerous traffic engineers that have tried to represent and understand their complex physics. A typical example is the case of an overtake event (André et al., 1989; Fernyhough et al., 2000). In this section, we will analyse this case in QTC starting from raw trajectory sample points as received from position aware devices. As the *left* / *right* distinction is crucial in overtake events, we will utilise QTC_{B1}.

Let us consider two cars *k* and *l*. Table 2-4 gives their two-dimensional sample coordinates during an overtake event at regular time steps of one second. As QTC assumes continuity, such a discrete set of sample points has to be interpolated in order to obtain continuous trajectories. Although several approaches are possible, we will, for this example case, rely on simple linear interpolation in space and time. This is also shown in Table 2-4.

Table 2-4: Trajectory sample points of two cars *k* and *l* during an overtake event

sample point	car *k*			car *l*		
	x (m)	*y* (m)	*t* (s)	*x* (m)	*y* (m)	*t* (s)
1	15	0	0	15	10	0
2	15	5	1	15	13	1
3	10	13	2	15	17	2
4	10	23	3	15	23	3
5	10	33	4	15	28	4
6	15	41	5	15	33	5
7	15	46	6	15	38	6





We find the following QTC_{B1} relation pattern: $(-\,+)_{B1} \rightarrow (-\,+)_{B1} \rightarrow (0\ 0)_{B1} \rightarrow (+\,-)_{B1} \rightarrow (+\,-)_{B1}$. Since this is a pattern of subsequent conceptual neighbours, we call it a *conceptual animation* (Van de Weghe, Cohn et al., 2005). It consists of five relations, four of which hold over a time interval, whereas $(0\ 0)_{B1}$ occurs instantaneously at 3 s. Although all others last over intervals, continuity theory induces some subtle differences between them. As pointed out earlier, a '0' value must always last over a closed time interval (of which a time instant is a special case), whereas '–' and '+' must always hold over an open time interval. Therefore, it follows that $(-\,+)_{B1}$ and $(+\,-)_{B1}$ persist over open time intervals, whereas $(0\ 0)_{B1}$ occurs at an instantaneous closed time interval. Note that, as Table 2-4 does not provide a preceding and following sample point for respectively the first and the seventh sample point, the change in movement direction is unknown at these instants. Consequently, the beginning of $(-\,+)_{B1}$ and the end of $(+\,-)_{B1}$ are unknown. With this knowledge, a more complete description of the complete conceptual animation would be:

$$]0, 1]:(-\ +)_{B1} \rightarrow\ ]1, 3[:(-\ +)_{B1} \rightarrow [3]:(0\ 0)_{B1} \rightarrow\ ]3, 5[:(+\ -)_{B1} \rightarrow [5, 6[:(+\ -)_{B1}.$$

Similarly to the overtake event, many other dynamic phenomena can be modelled by means of conceptual animations. Some of them are studied in this research. A qualitative framework can then be composed of such QTC patterns in order to reason about, recognise or simulate traffic events.

2.9 Conclusions

This chapter has presented the Qualitative Trajectory Calculus as a qualitative spatio-temporal calculus to handle the relations between moving objects adequately. The development of QTC and which spatial and temporal calculi inspired QTC has been discussed. The chapter has focused on the most general and fundamental QTC calculi, i.e. the Basic type, as it constitutes the basis of all other types. The principal reasoning mechanisms such as conceptual neighbourhoodness have been considered in some detail, as well as the ability for QTC to deal with incomplete knowledge. The usefulness and applicability of QTC has been illustrated in a simple case where, starting from raw trajectory data, a conceptual QTC animation is obtained.

References

Allen, J., F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26*(11), 832-843.

André, E., Herzog, G., & Rist, T. (1989). *Natural language access to visual data: Dealing with space and movement*. Universität des Saarlandes, Saarbrücken.

Bennett, B. (1997). *Logical Representations for Automated Reasoning about Spatial Relationships.* University of Leeds, Leeds.

Bitterlich, W., Sack, J. R., Sester, M., & Weibel, R. (2011). Representation, analysis and visualization of moving objects. *Proceedings of the Representation, Analysis and Visualization of Moving Objects*, Dagstuhl, Germany.

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., & De Maeyer, P. (2006). *The qualitative trajectory calculus on networks. Proceedings of the Spatial Cognition V Reasoning, Action, Interaction*, Bremen, Germany.

Bogaert, P., Van der Zee, E., Maddens, R., Van de Weghe, N., & De Maeyer, P. (2008). Cognitive and linguistic adequacy of the qualitative trajectory calculus. *Proceedings of the International Worshop on Moving Objects: From Natural to Formal Language, Utah, USA.*

Cohn, A. G., & Renz, J. (2007). Qualitative spatial reasoning. In: F. Van Hermelen, V. Lifschitz & B. Porter (Eds.), *Handbook of Knowledge Representation*. Elsevier.

Cristani, M., Cohn, A. G., & Bennett, B. (2000). Spatial locations via morpho-mereology. *Proceedings of the Principles of Knowledge Representation and Reasoning (KR'2000)* (pp. 15-25). Colorado, USA.

Egenhofer, M. J., & Al-Taha, K. K. (1992). Reasoning about gradual changes of topological relationships. In: A. U. Frank, I. Campari & U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. (pp. 196-219).

Egenhofer, M. J., & Franzosa, R. D. (1991). Point-set topological spatial relations. *International Journal of Geographical Information Systems, 5*(2), 161-174.

Fernyhough, J., Cohn, A. G., & Hogg, D. C. (2000). Constructing qualitative event models automatically from video input. *Image and Vision Computing, 18*(2), 81-103.

Frank, A. U. (1996). Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Systems, 10*(3), 269-290.

Freksa, C. (1992a). Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54*, 199-127.

Freksa, C. (1992b). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Galton, A. (2001). Dominance Diagrams: a Tool for Qualitative Reasoning about Continuous Systems. *Fundamenta Informaticae, 46*(1-2), 55-70.

Gudmundsson, J., van Kreveld, M., & Speckmann, B. (2004). Efficient detection of motion patterns in spatio-temporal data sets. *Proceedings of the 12th annual ACM International Workshop on Geographic Information Systems* (pp. 250-257).

Guting, R. H., Almeida, T. d., & Ding, Z. (2006). Modelling and querying moving objects in networks. *The VLDB Journal, 15* (2), 165-190.

Guting, R. H., Bohlen, M. H., Erwig, M., Jensen, C. S., Lorentzos, N. A., Schneider, M., et al. (2000). A foundation for representing and querying moving objects. *ACM Transactions on Database Systems, 25*(1), 1-42.

Hallot, P., & Billen, R. (2008). Generalized life and motion configurations reasoning model. *Proceedings of the International Worshop on Moving Objects: From Natural to Formal Language*. *Utah, USA.*

Ibrahim, Z., & Tawfik, A. (2007). An abstract theory and ontology of motion based on the regions connection calculus. *Proceedings of the Abstraction, Reformulation, and Approximation* (pp. 230-242).

Kurata, Y., & Egenhofer, M. J. (2009). Interpretation of behaviours from a viewpoint of topology. In: B. Gottfried & H. Aghajan (Eds.), *Behaviour Monitoring and*

Interpretation - BMI, Ambient Intelligence and Smart Environments (Vol. 3). Amsterdam: IOS Press.

Laube, P. (2005). *Analysing Point Motion: Spatio-Temporal Data Mining of Geospatial Lifelines*. University of Zurich, Zurich.

Muller, P. (2002). Topological spatio-temporal reasoning and representation. *Computational Intelligence, 18*(3), 420-450.

Noyon, V., Claramunt, C., & Devogele, T. (2007). A Relative representation of trajectories in geogaphical spaces. *Geoinformatica, 11*(4), 479-496.

Peuquet, D. J. (2001). Making space for time: Issues in space-time data representation. *Geoinformatica, 5*(1), 11-32.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In: B. Nebel, W. Swartout & C. Rich (Eds.), *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR)* (Vol. 92, pp. 165-176).

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach*. Ghent University, Ghent.

Van de Weghe, N., Cohn, A. G., De Maeyer, P., & Witlox, F. (2005). Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems With Applications, 29*(4), 977-983.

Van de Weghe, N., Cohn, A. G., De Tré, G., & De Maeyer, P. (2006). A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems. *Control and Cybernetics, 35*(1), 97-119.

Van de Weghe, N., & De Maeyer, P. (2005). Conceptual neighbourhood diagrams for representing moving objects. *Perspectives in Conceptual Modelling* (pp. 228-238).

Van de Weghe, N., De Tré, G., Kuijpers, B., & De Maeyer, P. (2005). The double-cross and the generalization concept as a basis for representing and comparing shapes of polylines. *Proceedings of the 1st International Workshop on Semantic-based Geographical Information Systems (SeBGIS'05)*. Agia Napa, Cyprus.

Van de Weghe, N., Kuijpers, B., Bogaert, P., & De Maeyer, P. (2005b). A qualitative trajectory calculus and the composition of its relations. In: M. A. Rodriquez, I. F. Cruz, M. J. Egenhofer & S. Levashkin (Eds.), *Geospatial Semantics*. (pp. 60-76) Berlin Heidelberg: Springer-Verlag.

Wolter, F., & Zakharyaschev, M. (2000, 2000). Spatio-temporal representation and reasoning based on RCC-8. *Proceedings of the 7th Conference on Principles of Knowledge Representation and Reasoning (KR2000)* (pp. 3-14).

Zimmermann, K., & Freksa, C. (1996). Qualitative spatial reasoning using orientation, distance, and path knowledge. *Applied Intelligence, 6*(1), 49-58.

Guting R. H., Almeida, T. d., & Ding, Z. (2006). Modelling and querying moving objects in networks. *The VLDB Journal, 15* (2), 165-190.

3

The Continuous Triangular Model

Each problem that I solved became a rule, which served afterwards to solve other problems.
Rene Descartes

Modified From: Yi Qiang, Seyed Hossein Chavoshi, Steven Logghe, Philippe De Maeyer, and Nico Van de Weghe (2014). Multi-Scale Analysis of Linear Data in a Two-Dimensional Space. Information Visualization, 13(3), 248-265.

3 THE CONTINUOUS TRIANGULAR MODEL

Abstract: Many disciplines are faced with the problem of handling time-series data. We introduce an innovative visual representation for time series, namely the continuous triangular model. In the continuous triangular model, all subintervals of a time series can be represented in a two-dimensional continuous field, where every point represents a subinterval of the time series, and the value at the point is derived through a certain function (e.g. average or summation) of the time series within the subinterval. The continuous triangular model thus provides an explicit overview of time series at all different scales. In addition to time series, the continuous triangular model can be applied to a broader sense of linear data. We also show how the coordinate interval space in the continuous triangular model can support the analysis of multiple time series through spatial analysis methods such as map algebra. A real-world dataset is employed to demonstrate the usefulness of this approach.

3.1 Introduction

Many disciplines are faced with the problem of handling time series data, which lead to considerable efforts dedicated to the research of time series (Enders, 2008; Hamilton, 1994; Liao, 2005). The temporal scale is one of the most important issues in time series analysis. Analogous to the well-known modifiable areal unit problem in spatial analysis, the way of aggregating temporal data may also significantly affect analysis results. Sometimes patterns or relationships detectible in a certain scale cannot be detected in other scales. Even in the same scale, different partitions of intervals may result in different patterns being revealed. On the other hand, a question can be answered in different scales. For example, the answers to the question when there are a lot of traffic jams in Belgium may include 'between 7:00am and 9:00am', 'during the days it snows' and 'in the months of school semester'. All these answers make sense because they may guide people to take actions in corresponding scales. Therefore, an appropriate choice of the temporal scale should take account of the characteristics of phenomena under study, the level of questions being asked, and the scale of actions to be taken. This choice is not easy, especially in the phase of exploratory analysis when there is not much known about the data and when the objective of the analysis is not accurately specified. In addition to specifying an appropriate scale for analysis, the hierarchy of phenomena in different scales can also be important in certain analytical tasks (Andrienko et al., 2010). Analysts may be interested in how long-term patterns are composed or influenced by short-term patterns within them. As a result, multi-scale analysis is of critical importance for analysing temporal data. Due to the complexity of this issue, the solution requires considerable human intelligence to be involved.

Visualisation has been proven to be an effective analytical approach for time series data (Aigner et al., 2007). An explicit visualisation can effectively combine the insight of humans and processing ability of computers (Andrienko & Andrienko, 2006; Keim et al., 2006a) to tackle analysis tasks. While a number of approaches have been developed to visualise time series (Havre et al., 2000; Hochheiser & Shneiderman, 2004; Lin et al., 2004; Weber et al., 2001), the line chart remains the most frequently used. In a line chart, the horizontal dimension indicates positions in the time line, and the vertical dimension indicates the values at the positions. The time series is represented as a curve, offering a direct view of the variation of time series along the linear space. Line diagrams usually only display time series in a certain scale. Displaying time series in different scales would require drawing more curves, which makes the data display matted. Manipulating a sliding bar to shift the scales to be displayed is an alternative approach. However, with the slider, one still cannot obtain an overall picture of time series in all different scales.

The Continuous Triangular Model (CTM) introduced in this chapter provides an alternative approach to represent time series and overcomes the difficulty of traditional approaches in visualising time series in multiple scales. The CTM is based on a diagrammatic representation of time intervals initially proposed by Kulpa (Kulpa, 1997b; Kulpa, 2006). Later, Van de Weghe named it the Triangular Model (TM) and applied it to archaeological use cases (Van de Weghe et al., 2007). More recently, Qiang investigated its use in reasoning about imperfect intervals (Qiang et al., 2010) and visual analytics (Qiang et al., 2012a; Qiang et al., 2012b). The basic idea of the TM is representing time intervals as points in a coordinated two-dimensional (2D) space. Evolved from the TM, the CTM adds the third dimension to the interval space of the TM and forms a continuous field, which can display time series in all different intervals. In the continuous field, every point represents a specific interval and is referenced to a certain value of the interval, such as the summation, average or standard deviation etc. On the one hand, the CTM can provide an overview of linear time series in all different scales. On the other hand, as the CTM is based on a 2D coordinate space, the glossary of spatial analysis methods in geographical information science (GIScience) are now open to be employed to manipulate and analyse the CTM data (Goodchild et al., 2007; Smith et al., 2007). In addition to time series, the CTM can also be applied to a broader sense of linear data, which refers to data sequences ordered in a one-dimensional (1D) space. Linear data can be derived from a linear geographical space, such as traffic speed along a road and runoff along a river, or objects with a linear structure, such as texts and DNA sequences.

In the remainder of Chapter 3, we first review the representative approaches in temporal visualisation. Next, the basic concept of the TM is introduced, followed by its extension the CTM. We then demonstrate how the CTM can be applied to visualise time series of soccer players. Afterwards, we show how map algebra can be applied to analyse time series represented in the CTM. The final section draws conclusions.

3.2 Related Work

Extensive reviews of time series visualisation can be found in (Aigner et al., 2007; Aigner et al., 2011; Muller & Schumann, 2003). A common weakness of existing approaches is that time series can only be displayed in one or a few pre-set temporal scales, which may miss interesting patterns in other scales. This also prohibits the observation of the complete hierarchy of scales, where smaller phenomena are nested within larger phenomena. Moreover, in a certain scale, time series is usually visualised in equal-length time granules, e.g. hour, day, month or year. The patterns in intervals that partially overlap granules cannot be displayed. To the best of our knowledge, there is no approach that can break away the barriers between scales and visualise time series in all intervals within the considered time frame. This problem also exists in a broader sense of linear data. The CTM presented in this chapter can be considered as a solution to this problem.

An idea similar to the CTM is the Growth Matrix introduced by Keim (Keim et al., 2006b), which visualises stock price changes in a 2D space. In the Growth Matrix, the horizontal axis indicates the time when the fund is purchased, and the vertical axis indicates when the fund is sold. Every point in the matrix is referenced to the price difference between the purchasing and selling times. Beyond Keim's research, this work demonstrates how other formulas (i.e. average and summations) can be applied to calculate the values of intervals, and how this representation can be useful for analysing different types of linear data. Moreover, it shows the use of map algebra in comparing multiple time series, and how cartographic modelling (Tomlin, 1990) can be applied to the CTM to solve multi-criteria decision-making problems based on time series.

3.3 Basic Concepts

3.3.1 Triangular Model

In the classical linear representation, a time interval I is represented as a linear segment bounded by a start point I^- and end point I^+. The properties of an interval are expressed by the location and extent of the linear segment in a 1D space.

The basic idea of the TM is mapping the linear segment in the 1D space into points in a 2D space. Given an arbitrary time interval I, two straight lines (L_1 and L_2) are projected from the two extremes (I^- and I^+), with L_1 passing through I^-; L_2 through I^+ (Figure 3-1). α_1 is the angle between L_1 and the horizontal axis, while α_2 is the angle between L_2 and the horizontal axis, where $\alpha_1 = -\alpha_2 = \alpha$. The intersection point of L_1 and L_2 is called the interval point, which expresses the properties of the time interval I. The horizontal position indicates the midpoint of I, i.e. $mid(I) = (I^- + I^+)/2$, while the vertical position indicates the duration of I, i.e. $dur(I) = \tan\alpha \cdot (I^+ - I^-)/2$. The start of the interval I^-, the end of the interval I^+ and interval point I form an isosceles triangle. Therefore, this representation of time intervals is called the Triangular Model (TM). The angle α is a pre-defined constant that is identical to the construction of all interval points. Here, we set $\alpha = 45°$ to be consistent with previous work (Kulpa, 1997a; Kulpa, 2006; Qiang et al., 2010), though α can be set to any value between 0° and 90° for specific purposes. In the TM, every time interval can be represented as a unique point in the 2D space. The 2D space where interval points are located in is called the Interval Space ($I\mathbb{R}$).

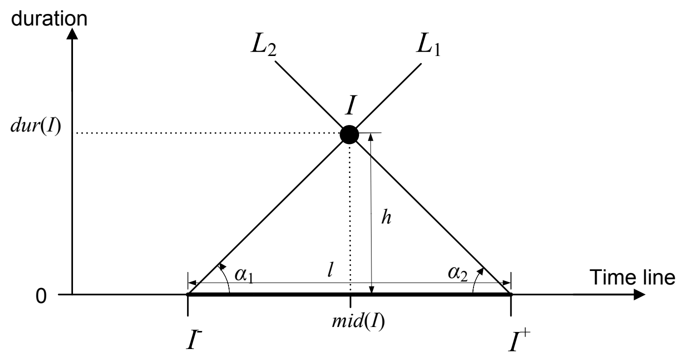


Figure 3-1. The configuration of the Triangular Model (TM).

3.3.2 Continuous Triangular Model

In addition to discrete time intervals, the TM can be extended to represent continuous temporal data. Given a time interval I, all intervals *during I* are enclosed in a triangular zone below it. In other words, every interval I_n *during I* corresponds to a specific point in this triangular zone. Let us consider a linear dataset arranged within I. Every point in the triangular zone represents a sub-interval I_n of the linear data. If every point is assigned a certain value, i.e. $f(I_n)$, of the interval it represents, then the triangular area can be filled and becomes a continuous field. $f(I_n)$ is a certain formula dependent on I_n, such as the average, summation or standard deviation of the linear data in I_n. Figure 3-2 illustrates how the CTM is built from a linear data sequence consisting of seven numbers. It shows that every point in the triangular area represents a certain subinterval of the sequence, and assigned a value that is calculated from the numbers within the subinterval. Here the granularity of the CTM is consistent to that of the linear data sequence. Finer granularity can be obtained through interpolation. Figure 3-3 gives an example of the implementation of the CTM in a raster space. Through colour-coding, the CTM can be displayed as an image.

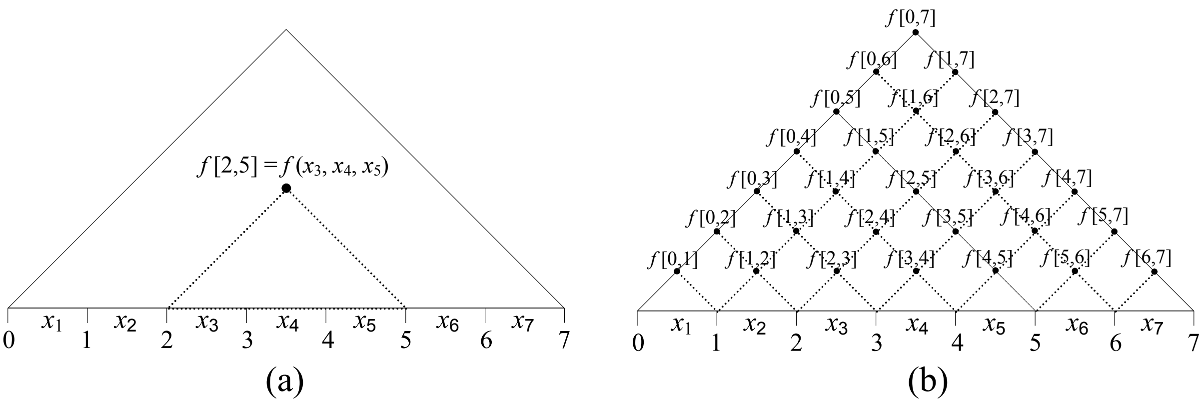
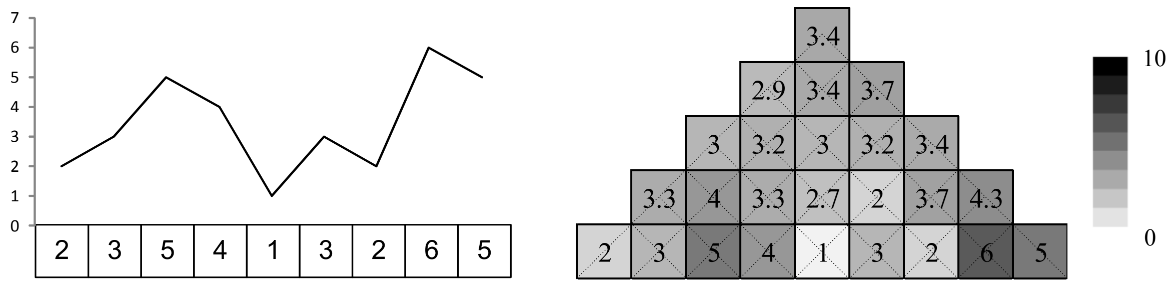


Figure 3-2: Representing a linear data sequence with seven numbers in the CTM. (a): A point is assigned a number calculated from the numbers within a subinterval. (b): Every point in the triangular space is assigned a number of a specific subinterval.



(a) (b)

Figure 3-3: The implementation of the CTM in a raster space. (a): A linear data sequence and its representation in a line chart. (b): The CTM representation of the linear data sequence in (a), with the average formula applied.

3.4 Visualisation of Linear Data

3.4.1 Visualising Time Series

This subsection demonstrates how the time series of the moving speed of a soccer player can be represented in the CTM. The movement of the soccer players is obtained through digitalisation of the game video. Here we study an indoor soccer game, in which each team has 5 players during a 13.33-minute portion of a game. The trajectories of the mini-soccer players are sampled from video frames taken at regular time stamps, with a temporal granularity of one second, by automatic computer vision-based tracking with field expert supervision. The time series is a player's speed in every second (i.e. meter/second or m/s) during a study interval of 800 s. As indoor soccer is rather intensive and fast-paced, the line chart (i.e. Figure 3-4a) exhibits dramatic changes of speed from second to second. However, variations in longer intervals (e.g. one minute or two minutes) are hard to observe. In Figure 3-4b the time series of the player's speed is represented by the CTM, where $f(I_n)$ is the average of the player's average speed during I_n. In the CTM, short-term fluctuations can be observed in lower levels, while the long-term patterns can be observed in higher levels. Moreover, it explicitly displays a hierarchy of the time series in all different scales, in which one can observe the relationship between the short-term variations and long-term variations. From this diagram, one can identify intervals of sprint from the red areas on the bottom of the CTM, e.g. I_1, I_2, and I_3. On a larger scale, it is clear that the player had a high average speed from 1:00 to 6:15 (i.e. I_4). However, during the next 3.5 minutes (i.e. I_5), he experienced a less active period, although there are still several sprints during it. Compared to the Growth Matrix of Keim (Keim et al., 2006b), in which I^- and I^+ are respectively coordinated along the vertical and horizontal axes, the coordinate space of the CTM preserves the linear nature of time that flows from left to right. Mapping longer intervals in higher positions is also somehow more intuitive than the Growth Matrix.

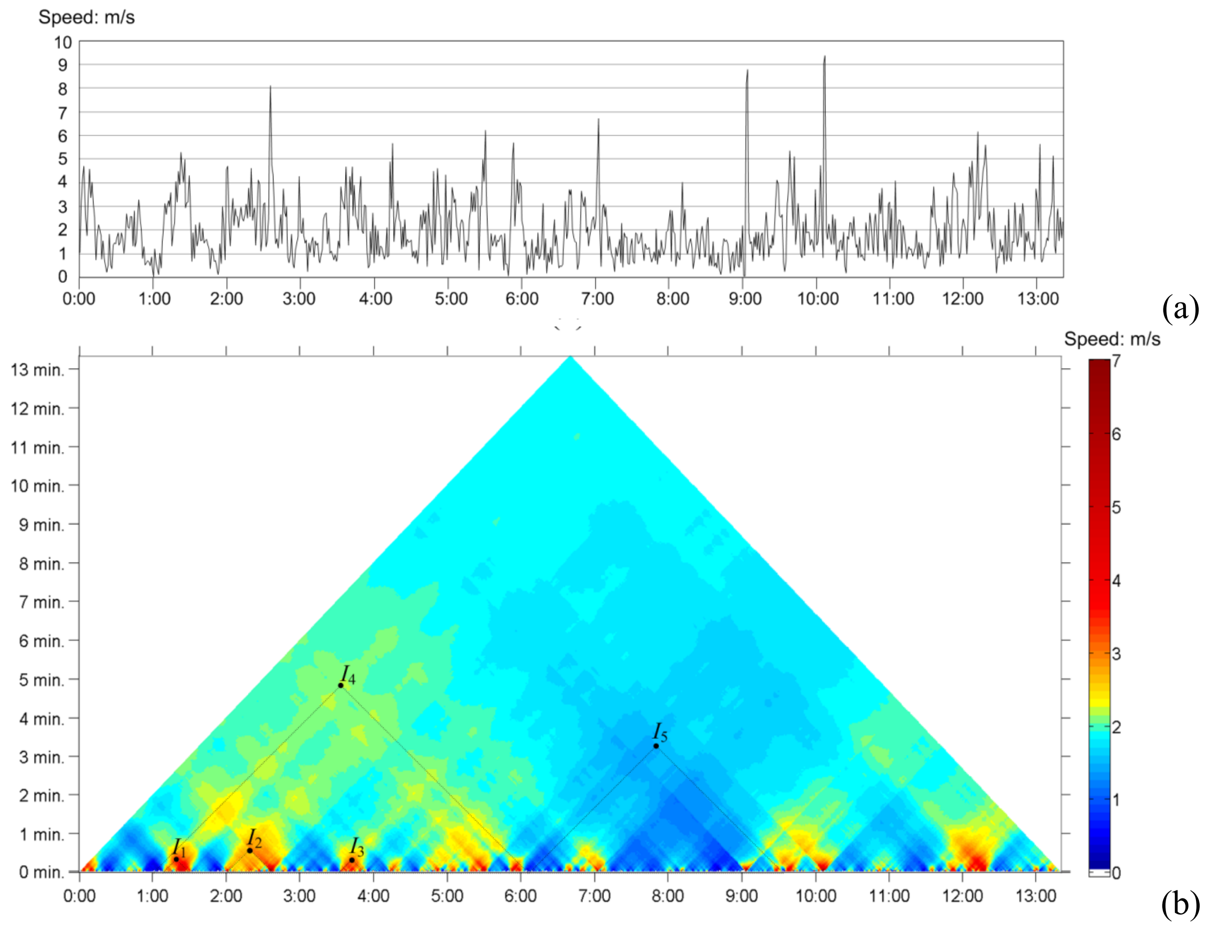
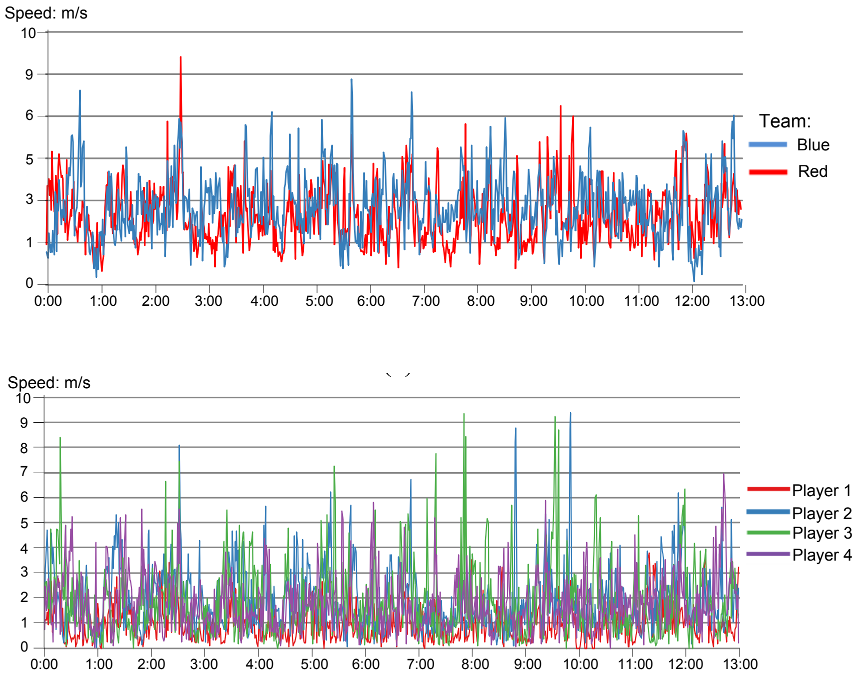


Figure 3-4: The time series of a soccer player's speed in a line diagram (a) and the CTM (b).

3.5 Analysing Multiple Time Series

As the CTM is based on a 2D coordinate interval spaces, many spatial analysis techniques in GIScience can be employed to analyse CTM diagrams. This section demonstrates how the method of map algebra is used to analyse multiple time series modelled by the CTM.

With the traditional line chart, the comparison of multiple time series can only be made in a fixed temporal scale and partition. For example, in Figure 3-5, one can only compare the speed between the indoor soccer teams or players at granularity of second. The speed in other scales is hard to compare. Alternatively, using map algebra in the CTM, these time series can be compared over all different time intervals. The time series of average running speed of the two competing teams can be compared by applying the ‘subtract’ algebra to their CTM diagrams. In the result CTM diagram, every point corresponds to a specific time interval, and the value at that point is the difference of the running speed between the two teams. Figure 3-6 illustrates the result of subtracting the CTM of the blue team from that of the red team. Blue represents a positive value, meaning that the average speed of the blue team is greater than that of the red team. Red represents a negative value, meaning that the average speed of the red team is greater. The result diagram can be interpreted as: in general, the blue team is more active (i.e. has greater running speed during long intervals), however, during some short time intervals, the red team has greater running speed, for instance, from the beginning to the 4th minute and from the 9th to the 13th minute. In Figure 3-6b, the darkness of the colours indicates the degree of difference, which gives a better sense of the actual difference. From Figure 3-6b, one can see that only during some very short intervals (less than 2 minutes), the red team is apparently more active than the blue team.

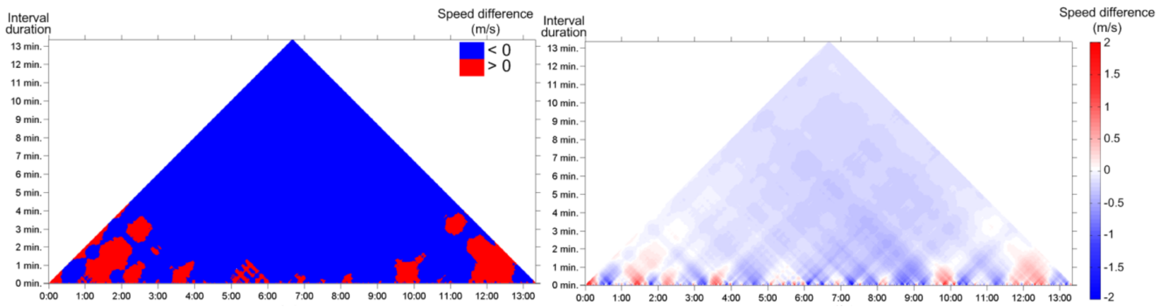


(a)

(b)

Figure 3-5: Line charts of multiple time series. (a): the average running speed of two competing soccer teams. (b): the running speed of individual players of the red team.

The CTM can also be used to compare more time series, for instance, the running speed of several soccer players. Here we compare the running speed of four players in the red team, who have played through the entire study period. This can be done by combining the CTM diagrams of these four players into one CTM using the following map algebra: at every point (i.e. during every interval), the player having the greatest running speed is selected.

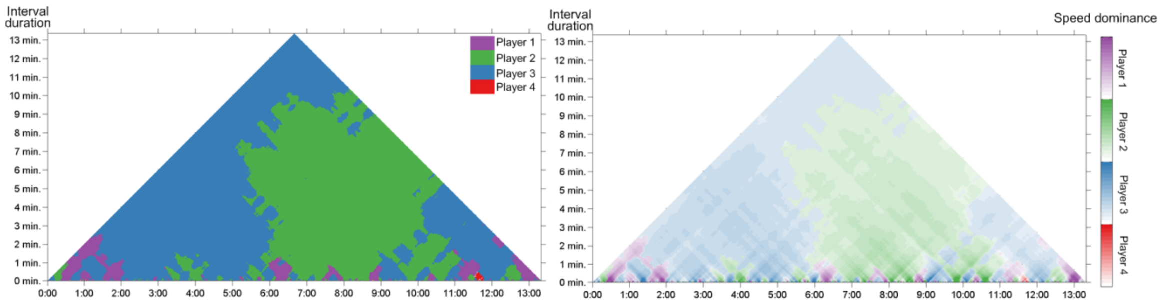


(a) (b)

Figure 3-6: The output of subtracting the CTM of the blue team from that of the red team. (a): Only colour hues are used to indicate the team with higher speed. (b): Colour darkness is applied to indicate the speed difference.

We define this player the dominant player of the interval. If the dominant player of every interval is displayed in a specific colour hue, the output becomes a nominal diagram with four zones (i.e. Figure 3-7a). Each zone represents a set of intervals during which a certain player is dominant. In Figure 3-7b, darkness of colours is used to indicate the degree of dominance. There are many ways to calculate the degree of dominance. In this case, we define the degree of dominance as the percentage that the dominant player's speed is greater than the average of the others. For example, if the speed of Player 1 is 5 and the average speed of Players 2-4 is 4, Player 1 is dominant over the others by 25%. Due to the many colour hues applied, it is a little hard to observe both the dominant players (represented by colour hues) and the degree of dominance (represented by darkness) in Figure 3-7b.

This problem can be overcome by representing players dominant at certain levels in multiple CTM diagrams. In these diagrams, each colour hue represents the player dominant by a certain percentage. For example, in the top-left diagram of Figure 3-8, every colour indicates a player that is dominant over the others by at least 10%, which means the speed of this player is more than the average speed of the others by at least 10%. Figure 3-9 uses two examples to illustrate this algebra in a raster space. From these diagrams, one can see that, with the increase of dominance threshold, the colour zones with low dominance gradually disappear, while the remaining zones represent the intervals during which a player is dominant above a certain level. With the dominance threshold of 30%, it becomes clear that Player 2 and Player 3 are much more active than the others during two successive 5-minute intervals, which possibly reveals a strategy change or position shift during the game. The variation of these CTM diagrams can be better observed through a controlled animation, where the CTM diagram dynamically responses to a slider setting the dominance level.



(a) (b)

Figure 3-7: Comparison of multiple players. (a): Dominant players is represented by discrete colour hues. (b): Dominance degree is represented by darkness of colours.

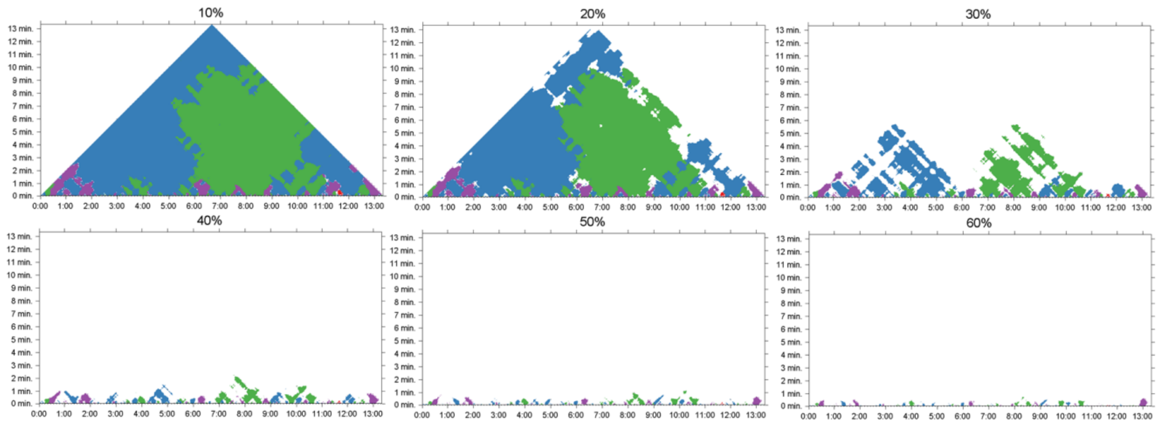


Figure 3-8: Players dominant at different degrees are represented in multiple diagrams. The meaning of colour hues is identical to that in Figure 3-7.

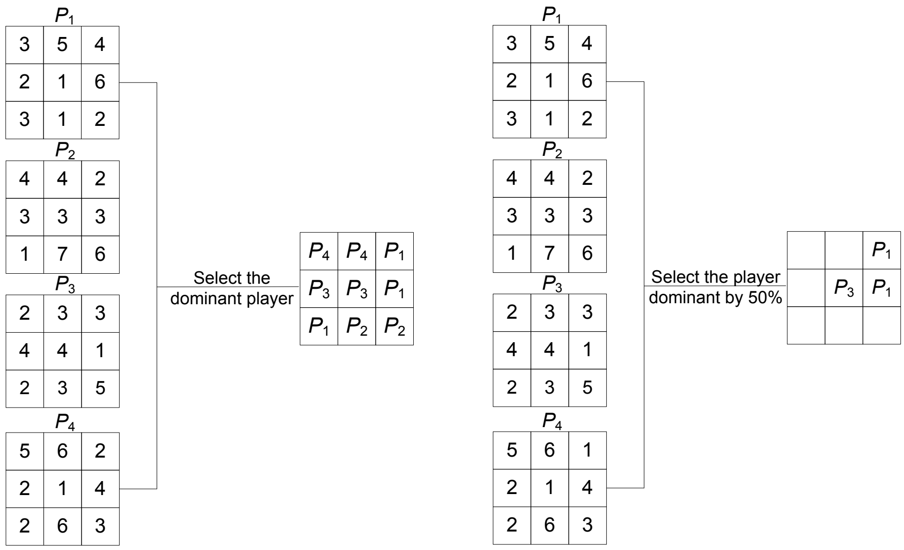


Figure 3-9: Select the dominant player at every point in a CTM diagram. P means player.

3.6 Conclusion

We introduced an innovative representation of linear data, namely the CTM. In the CTM, the linear data in different intervals are displayed in a two-dimensional space, constituting a basis for a multi-scale analysis of linear data. In general, the CTM has two major advantages. First, it provides an explicit and compact visualisation of linear data in different scales. In the CTM, moving statistics (e.g., the average and summation) during intervals of different lengths can be displayed in one diagram, which offers an explicit overview of patterns in different scales. Compared with traditional multi-scale visualisation approaches, which only display data in a few selected scales, the CTM can present the data in all scales. This feature is particularly useful for the exploration of unfamiliar datasets, in which interesting patterns may emerge in any scales. Also, the CTM offers an overview of the hierarchy of scales, which allows the observation of the interaction between large-scale patterns and small-scale patterns. Second, the CTM is based on a universal 2D coordinate space, which is very similar to prevalent geospatial datasets. This chapter demonstrated how existing techniques in GIScience can be used to manipulate and analyse CTM diagrams. By applying map algebra to the CTM, multiple time series can be compared at different scales. The CTM can use a single diagram to present the answer to the questions like ‘whether Brussels is warmer than New York’ or ‘whether Player 2 is more active than the other players’ according to all possible intervals within the considered time frame. We contend that the CTM representation of these answers is more informative and perceivable than traditional approaches such as line chart or colour lines. With the support of special visualisation techniques (e.g. setting a dominance threshold), it can have an extensive coverage of different analysis tasks. An example of soccer game was given to show the applicability of the CTM.

References

Aigner, W., Miksch, S., Muller, W., Schumann, H., & Tominski, C. (2007). Visualising time-oriented data - A systematic view. *Computers & Graphics, 31* (3), 401-409.

Aigner, W., Miksch, S., Schumann, H., & Tominski, C. (2011). *Visualization of Time-Oriented Data*. London: Springer-Verlag.

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M.-J., Schumann, H., & Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science, 24* (10), 1577-1600.

Andrienko, N., & Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Berlin Heidelberg: Springer-Verlag.

Enders, W. (2008). *Applied Econometric Time Series (2nd Ed.)*. New Jersey: Wiley.

Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science, 21* (3), 239-260.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Havre, S., Hetzler, B., & Nowell, L. (2000). ThemeRiver: Visualising theme changes over time. *Proceedings of the IEEE Symposium on Information Vizualization* (pp. 115-123).

Hochheiser, H., & Shneiderman, B. (2004). Dynamic query tools for time series data sets: Timebox widgets for interactive exploration. *Information Visualization, 3* (1), 1-18.

Keim, D., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006a). Challenges in visual data analysis. *Proceedings of the 10th International Conference on Information Visualization* (pp. 9 - 16).

Keim, D., Nietzschmann, T., Schelwies, N., Schneidewind, J., Schreck, T., & Ziegler, H. (2006b). A spectral visualization system for analysing financial time series data. In: T. Ertl, K. Joy & B. Santos (Eds.), *Proceedings of the Eurographics/IEEE-VGTC Symposium on Visualization (EuroVis 2006)* (pp. 195-200).

Kulpa, Z. (1997a). Diagrammatic representation for a space of intervals. *Machine Graphics & Vision, 6*, 5-24.

Kulpa, Z. (1997b). Diagrammatic representation of interval space in proving theorems about interval relations. *Reliable Computing, 3* (3), 209-217.

Kulpa, Z. (2006). A diagrammatic approach to investigate interval relations. *Journal of Visual Languages and Computing, 17* (5), 466-502.

Liao, T. W. (2005). Clustering of time series data: A survey. *Pattern Recognition, 38* (11), 1857-1874.

Lin, J., Keogh, E., Lonardi, S., Lankford, J. P., & Nystrom, D. M. (2004). Visually mining and monitoring massive time series. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 460-469).

Muller, W., & Schumann, H. (2003). Visualization methods for time-dependent data - An overview. *Proceedings of the 2003 Winter Simulation Conference* (Vol. 1, pp. 737-745).

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P., & Van de Weghe, N. (2010). Modelling imperfect time intervals in a two-dimensional space. *Control and Cybernetics, 39* (4), 983-1010.

Qiang, Y., Delafontaine, M., Neutens, T., Stichelbaut, B., De Tré, G., De Maeyer, P., & Van de Weghe, N. (2012a). Analysing imperfect temporal information in GIS using the triangular model. *The Cartographic Journal, 49* (3), 265-280.

Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P., & Van de Weghe, N. (2012b). Interactive analysis of time interval in a two-dimensional space. *Information Visualization, 11* (4), 255-272.

Smith, M. J. d., Goodchild, M. F., & Longley, P. A. (2007). *Geospatial analysis: A comprehensive guide to principles, techniques and software tools*. Leicester: Troubador Publishing Limited.

Tomlin, C. D. (1990). *Geographic Information Systems and Cartographic Modelling.* New Jersey: Prentice Hall.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B., & Ryckbosch, K. (2007). The triangular model as an instrument for visualising and analysing residuality. *Journal of Archaeological Science, 34* (4), 649-655.

Weber, M., Alexa, M., & Muller, W. (2001). Visualising time-series on spirals. *Proceedings of IEEE Symposium on Information Visualization* (pp. 7-13).

4

Repetitive Motion Patterns

A picture is worth a thousand words. Napoleon Bonaparte

Modified From: Seyed Hossein Chavoshi, Bernard De Baets, Yi Qiang, Guy De Tré, Tijs Nuetens, and Nico Van de Weghe (2015). A Qualitative Approach to the Identification, Visualisation and Interpretation of Repetitive Motion Patterns in Groups of Moving Point Objects. International Arab Journal of Information Technology, 12(5), First Online: http://ccis2k.org/iajit/PDF/Vol%2012,%20No.%205/7358.pdf

4 Repetitive Motion Patterns

Abstract: Discovering repetitive patterns is important in a wide range of research areas, such as bioinformatics and human movement analysis. This chapter puts forward a new methodology to identify, visualise and interpret repetitive motion patterns in groups of moving point objects (MPOs). The methodology consists of three steps. First, motion patterns are qualitatively described using the qualitative trajectory calculus (QTC). Second, a similarity analysis is conducted to compare motion patterns and identify repetitive patterns. Third, repetitive motion patterns are represented and interpreted in a continuous triangular model (CTM). As an illustration of the usefulness of combining these hitherto separated methods, two specific movement cases are examined: Samba dance, a rhythmical dance with many repetitive movements and mini-soccer, with fewer repetitive movements. The results show that in both cases the presented methodology is able to successfully identify, visualise and interpret the contained repetitive motion patterns.

4.1 Introduction

With recent advances in navigation and tracking systems, we are experiencing a dramatic growth in moving objects databases. These databases include the trajectories of human beings (Michael et al., 2006; Wang et al., 2003), animals (DeCesare et al., 2005; Laube et al., 2007) and vehicles (Brakatsoulas et al., 2004; Hvidberg, 2006). Discovering relevant information from these large and growing data sets is a challenging task. In recent years, significant research in a variety of disciplines has attempted to derive knowledge from motion data (see, among others, (Giannotti et al., 2009; Laube et al., 2005; Spaccapietra et al., 2008) for an overview). One way of discovering knowledge from large spatio-temporal datasets is by means of qualitative reasoning. To date, several qualitative spatial and temporal calculi have been introduced, e.g., interval algebra (Allen, 1983), cardinal direction calculus (Frank, 1991), Double-Cross calculus (Freksa, 1992) and region connection calculus (Randell et al., 1992). Of particular interest to the study of moving objects is the qualitative trajectory calculus (QTC) (Van de Weghe, 2004). QTC describes the interaction between moving point objects (MPOs) in a qualitative way.

In this chapter, we use QTC to identify repetitive motion patterns in the movement data of MPOs. The term ‘repetitive motion patterns’ refers to conceptual animations (sequences of QTC relations following the constraints imposed by qualitative reasoning) that occur more than once during the movement. Herein, conceptual animations are defined as movement sequences. Similarity analysis is used to calculate the degree of similarity between movement sequences. The movement sequences with high degrees of similarity are repetitive motion patterns. To display the degrees of similarity, a visualisation technique, the continuous triangular model (CTM), is applied. The methodology is illustrated with two real-world case studies; the first is Samba dance, in which the infrared-observed motions of different parts of the bodies of dancers are analysed, and the second is mini-soccer, in which the positions of the players are sampled from video frames.

With the introduction of this methodology, we seek to contribute to the procedure of knowledge discovery from movement of MPOs. The proposed methodology will help researchers and practitioners from various disciplines in analysing regularities and anomalies in moving object databases in their respective fields of expertise.

The remainder of this chapter is organised as follows. Section 4.2 introduces the preliminary concepts of QTC and CTM. Section 4.3 describes the methodology used to analyse the motion patterns in the context of QTC. In addition, the visualisation and interpretation of the repetitive motion patterns are presented. Section 4.4 gives a brief discussion, summarises the conclusions and presents possible future work.

4.2 Preliminaries

In this section, we briefly review some of the fundamental concepts related to qualitative trajectory calculus, similarity analysis between conceptual animations, and the continuous triangular model. These concepts will be used in the remainder of the chapter.

4.2.1 Qualitative Trajectory Calculus

The basic principle of qualitative trajectory calculus (QTC) is that the complex reality of moving objects can be simplified by describing the interaction between two disjoint point objects. Depending on the level of detail and the number of spatial dimensions, different types of QTC have been developed: QTC Basic (QTC_B) (Van de Weghe et al., 2006; Van de Weghe & De Maeyer, 2005), QTC Double-Cross (QTC_C) (Van de Weghe et al., 2005), and QTC Network (QTC_N) (Delafontaine et al., 2011). The first level of QTC Basic (i.e. QTC_{B1}) considers only the changing Euclidean distance between two objects, which is independent of the number of dimensions in which the movements take place. We restrict our calculations in this study to QTC_{B1} and from now on, we use the term QTC_B as a synonym for QTC_{B1}.

QTC_B relations are built from the distance constraints (A and B) introduced in Section 2.5 and created by a tuple of labels that have an identical three-valued qualitative domain $\{-, 0, +\}$. In total, there are nine QTC_B relations as shown in Figure 2-3a. At each time stamp, there is a QTC_B relation between two MPOs. Following the constraints imposed by continuity, a sequence of QTC_B relations (i.e. a conceptual animation) can be generated. For example, Figure 4-1 shows the interaction in a 2D space between two MPOs that are continuously moving. This interaction is represented by a sequence of three QTC_B relations during a given time interval $[t_1, t_2]$. In the beginning of the movement, the relation between the MPOs $(-\ -)$ is established during a time interval. The relation $(0\ 0)$ is an instantaneous QTC_B relation between the MPOs. The remaining relation $(+\ +)$ occurs during the last part of the movement (for a detailed explanation, see (Van de Weghe, 2004)).

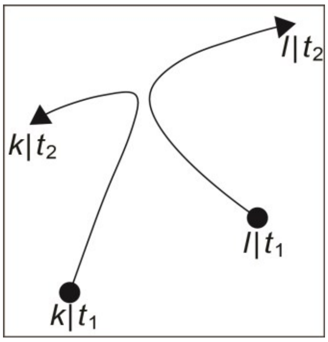


Figure 4-1: The conceptual animation of k and l during a time interval $[t_1, t_2]$

$(-\ -) \rightarrow (0\ 0) \rightarrow (+\ +)$

The relations between more than two MPOs can be presented in terms of a QTC_B matrix (see Section 2.7.1.). Consequently, for a time interval, a conceptual animation is proposed as a sequence of QTC_B matrices. For example, consider three MPOs, a, b, and c, at three consecutive time stamps (Figure 4-2). From time stamp t_1 to t_2, the QTC_B matrix X is formed by the QTC_B relations between all pairs of MPOs, and from time stamp t_2 to t_3, the QTC_B matrix Y is generated (Table 4-1).

In general, the goal of this approach is to identify, visualise and interpret the repetitive motion patterns in groups of MPOs by exploring their conceptual animations.

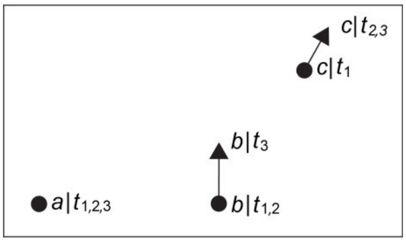


Figure 4-2: Three MPOs, a, b and c, during a time interval [t_1, t_3]

Table 4-1: A conceptual animation of two QTC_B matrices

$X[t_1 \rightarrow t_2]$	***a***	***b***	***c***
a		(0 0)	(0 +)
b	(0 0)		(0 +)
c	(+ 0)	(+ 0)	

⇨

$Y[t_2 \rightarrow t_3]$	***a***	***b***	***c***
a		(0 +)	(0 0)
b	(+ 0)		(− 0)
c	(0 0)	(0 −)	

4.2.2 Similarity Analysis between Conceptual Animations

Similarity analysis is used to express the degree of similarity between the conceptual animations. Prior to making a comparative analysis of two conceptual animations, we must decide how much detail needs to be considered in the comparison. For example, consider the following two conceptual animations, referring to the QTC_B relations among the three MPOs (a, b and c) during two time intervals in Table 4-2.

Table 4-2: A pair of conceptual animations among three MPOs during two time intervals [t_1-t_3] and [t_4-t_6]

conceptual animation [t_1-t_3]

$X[t_1 \rightarrow t_2]$	***a***	***b***	***c***
a		(+ −)	(− +)
b	(− +)		(0 +)
c	(+ −)	(+ 0)	

⇨

$Y[t_2 \rightarrow t_3]$	***a***	***b***	***c***
a		(+ 0)	(− −)
b	(0 +)		(0 −)
c	(− −)	(− 0)	

conceptual animation [t_4-t_6]

$X[t_4 \rightarrow t_5]$	***a***	***b***	***c***
a		(+ −)	(− −)
b	(− +)		(0 0)
c	(− −)	(0 0)	

⇨

$Y[t_5 \rightarrow t_6]$	***a***	***b***	***c***
a		(+ 0)	(− +)
b	(0 +)		(+ −)
c	(+ −)	(− +)	

For the sake of simplicity, each conceptual animation can be abstracted to a combined QTC_B matrix obtained by concatenating the ij^{th} cells of all QTC_B matrices in that conceptual animation (Table 4-3). Hence, each conceptual animation of any length (any time interval) can be represented by a combined QTC_B matrix.

Additionally, a movement during a time interval is divided into sub-intervals. In this study, to detect repetitive movement patterns, we start our comparison from the lowest level (level 1, which consists of only one QTC_B matrix) and extend it to the higher levels. For example, Table 4-3 shows the comparison of two sub-intervals of level 2. For the entire movement, all combined QTC_B matrices of level 2 should be compared to measure the degrees of similarity between them. This process is repeated for all levels, where the last level represents the entire movement.

Table 4-3: Combined QTC_B matrices during two time intervals

conceptual animation [t_1-t_3]

$t_1 \rightarrow t_2 \rightarrow t_3$	***a***	***b***	***c***
a		$(+\ -)(+\ 0)$	$(-\ +)(-\ -)$
b	$(-\ +)(0\ +)$		$(0\ +)(0\ -)$
c	$(+\ -)(-\ -)$	$(+\ 0)(-\ 0)$	

conceptual animation [t_4-t_6]

$t_4 \rightarrow t_5 \rightarrow t_6$	***a***	***b***	***c***
a		$(+\ -)(+\ 0)$	$(-\ -)(-\ +)$
b	$(-\ +)(0\ +)$		$(0\ 0)(+\ -)$
c	$(-\ -)(+\ -)$	$(0\ 0)(-\ +)$	

The combined QTC_B matrices can also be compared cell-by-cell. Two levels of detail are possible. In the highest level of detail, the fine comparison, the individual symbols of QTC_B notation in each cell are compared based on the topological distance presented by Egenhofer and Al-Taha (Egenhofer & Altaha, 1992) (for additional explanations, see (Van de Weghe, 2004)). In the coarse comparison, regardless of the details, a complete cell of a combined QTC_B matrix is compared to the corresponding cell in another combined QTC_B matrix at each level. In this study, we use the coarse comparison, which reflects the full equality of relations between pairs of MPOs. For this purpose, Eq. 4-1 is used to calculate the degree of similarity (expressed as a percentage) between a pair of combined QTC_B matrices as follows:

$$S = 100 * ((N - L)/N) \qquad \text{Eq. 4-1}$$

where N is the total number of cells in the combined QTC_B matrix after eliminating the elements below the diagonal of the matrix because they are interchangeable with the elements above the diagonal of the matrix and L is the number of non-identical cells above the diagonal. This expression is the simple matching similarity measure for categorical data. The degree of similarity for Table 4-3 is calculated as follows:

$$S = 100 * ((3-2)/3) = 33.33\%$$

As mentioned above, different levels of comparison are considered based on the length of the conceptual animations. In a subsequent section, the similarities between motion patterns are visualised using the continuous triangular model (CTM) to interpret the repetitive motion patterns.

4.2.3 The Continuous Triangular Model

CTM is derived from the idea of the triangular model (TM), which represents time intervals as points in a two-dimensional space (A detailed explanation can be found in Chapter 3). In the CTM, attribute data are associated with the points of the time intervals. Consequently, time series data can be mapped to a triangular plane in the 2D space, in which every point represents a specific interval of the time series and the grey scale at the point indicates a certain aggregation (e.g., summation and average) of time series of this interval. This representation of time series is the CTM. In Figure 3-4, we illustrated the two representations of a time series. Figure 3-4a showed a traditional line diagram of a time series. In the triangular plane in Figure 3-4b, every point corresponded to a time interval, following the coordinate space described in Figure 3-1, and the colour level at the point in Figure 3-4b indicated the average value of the time series within the interval. Using this approach, variations of short intervals can be observed in the lower levels of the triangular plane, and variations of long intervals can be observed in the higher levels. The CTM provides a direct overview of time series data at all temporal granularities. In addition to time series data, the CTM can be applied to other types of sequential data. In the following sections, we integrate CTM in the procedure of knowledge discovery from movement of MPOs.

4.3 Motion Pattern Analysis of MPOs

A comprehensive classification of movement patterns has been proposed by Dodge (Dodge et al., 2008). We focus on one of the primitive patterns in that classification: spatio-temporal periodicity (repetitive motion patterns). This chapter constitutes a novel contribution to the identification, visualisation and interpretation of the repetitive motion patterns. The workflow diagram presented in Figure 4-3 illustrates our approach. The procedure starts with raw data (trajectories of MPOs). Motion patterns of MPOs are obtained from the raw data. Then, similarity analysis is used to determine the degrees of similarity among the motion patterns. Finally, the degrees of similarity are visualised using the CTM to interpret them. In the following subsection, two different case studies are examined to analyse the repetitive motion patterns. First, the movements of three Samba dancers are analysed; second, the movements of mini-soccer players during a game are examined.

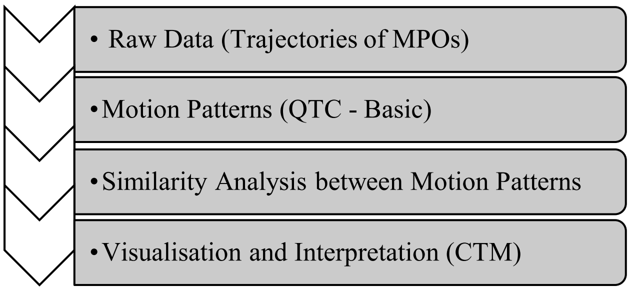


Figure 4-3: Procedure overview

4.3.1 Samba Dancers

In this subsection, the movement of the different parts of the bodies of Samba dancers is analysed. Relations between the different parts of the bodies of the dancers are described as QTC_B relations based on the positional information at each time stamp of the movement.

The positional information consists of locations of the MPOs in a three-dimensional space that includes the head, the root, the right finger (the right hand), the left finger (the left hand), the right toe (the right foot), and the left toe (the left foot) of every dancer's body, captured at every time stamp (temporal granularity of 0.04 s). For example, Table 4-4 shows a sequence of QTC_B matrices formed based on the positional information of all captured MPOs during a given time interval. The movement of the body is captured by an infrared motion capturing system, which yields the position of markers attached to the body. We use a normalised data set with respect to one reference point and the orientation of the dancer's body (the point is defined as the centroid of the body, root) (Naveda, 2011; Naveda & Leman, 2010). As mentioned earlier, similarity analysis is used to calculate the degrees of similarity between different movement sequences.

Based on the basic concept of CTM introduced in the previous section, we apply a modified version of CTM to map the similarities between different pairs of movement sequences into a triangular raster. Every cell in the raster represents a pair of movement sequences of equal length, and the grey scale of the cell indicates their degree of similarity.

4.3.1.1 The Horizontal and Vertical Dimensions

Here, the horizontal dimension of the raster represents the time line, and the vertical dimension represents the time distance between two sequences. The two sequences of the cell can be identified by drawing a 45°-45°-90° isosceles triangle on the horizontal axis (Figure 4-4). The 90° vertex is located in the cell. The two 45° vertices are located on the horizontal axis and identify the starting times of the two sequences. For illustration purpose, Figure 4-4 shows a highlighted triangle in which the cell at the 90° vertex represents a pair of movement sequences starting at 1.4 s and 3.2 s. The grey at the 90° vertex indicates the similarity between the pair of movement sequences.

The cell's position on the vertical axis indicates the distance between the starting points of the two represented sequences. For the highlighted triangle in Figure 4-4, the vertical position of the cell is 45 time stamps (1.80 s; the temporal granularity is 0.04 s), which is the temporal distance between the starting points of the two sequences.

Table 4-4: The movement sequence of the QTC_B matrices during a given time interval [0-0.24] (LF: left finger, RF: right finger, LT: left toe, RT: right toe, R: root, H: head).

0.00 – 0.04	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ 0)	(+ 0)	(+ −)	(+ 0)	(+ +)
RF			(+ −)	(+ −)	(0 0)	(− 0)
LT				(− 0)	(− 0)	(− 0)
RT					(− 0)	(− 0)
R						(0 0)
H						

0.04 – 0.08	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ −)	(+ −)	(+ −)	(+ 0)	(0 +)
RF			(+ 0)	(+ −)	(0 0)	(− +)
LT				(− −)	(− 0)	(− +)
RT					(− 0)	(− +)
R						(0 +)
H						

0.08 – 0.12	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ −)	(+ 0)	(+ −)	(+ 0)	(0 0)
RF			(0 +)	(0 0)	(− 0)	(− 0)
LT				(0 −)	(0 0)	(0 0)
RT					(− 0)	(0 0)
R						(0 0)
H						

0.12 – 0.16	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ −)	(+ 0)	(+ 0)	(+ 0)	(0 −)
RF			(− 0)	(− +)	(− 0)	(− 0)
LT				(+ −)	(0 0)	(0 −)
RT					(+ 0)	(+ −)
R						(0 −)
H						

0.16 – 0.20	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ −)	(+ −)	(+ 0)	(+ 0)	(− −)
RF			(− −)	(− +)	(− 0)	(− +)
LT				(+ 0)	(− 0)	(− 0)
RT					(0 0)	(0 0)
R						(0 0)
H						

0.20 – 0.24	*LF*	*RF*	*LT*	*RT*	*R*	*H*
LF		(+ −)	(+ 0)	(+ −)	(+ 0)	(− −)
RF			(− 0)	(− 0)	(− 0)	(− +)
LT				(+ −)	(0 0)	(0 +)
RT					(− 0)	(− +)
R						(0 +)
H						

4.3.1.2 The Level Number

The level of the CTM indicates the length of the movement sequences. For example, a level 1 CTM represents the similarities between any two movement sequences whose lengths are $1 * 0.04$ s (because the temporal granularity is 0.04 s), and a level 4 CTM represents the similarities between any two movement sequences whose lengths are

0.16 s ($4 * 0.04$ s). Hence, in Figure 4-4, the lengths of the movement sequences are 0.12 s ($3 * 0.04$ s). Therefore, the cell at the top of the highlighted triangle represents the similarity between the movement sequence during the temporal interval $[1.4, 1.4 + 0.12]$ and the movement sequence during the temporal interval $[3.2, 3.2 + 0.12]$.

4.3.1.3 The Grey Scale

In CTM, the grey scale of a cell indicates the similarity between two movement sequences, as calculated using Eq. 4-1. Black is 100% similarity, and white is 0% similarity. The grey bar on the right side of the CTM results displays the similarity scale.

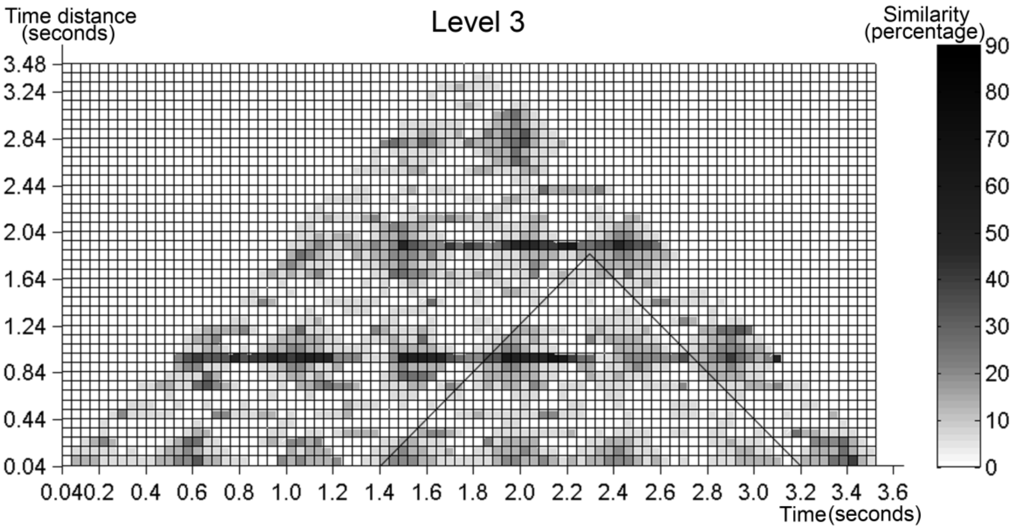


Figure 4-4: The CTM representation of similarities between movement sequences

4.3.1.4 Comparison of CTMs

The CTM visualises the similarity between the movement sequences of the person during two different time intervals. As explained above, a cell of a level 10 CTM displays the similarity between movements during the interval $[t_1, t_1+0.4]$ and movements of the same person during the interval $[t_2, t_2+0.4]$.

From the CTM of one person, temporal patterns of movements of the person can be observed. Now, the movements of three different Samba dancers (student 1, student 2 and their teacher) are analysed. The CTM representations show some regular patterns (Figure 4-5, Figure 4-6, and Figure 4-7). The first four levels of CTM for the three dancers are shown. High similarities (i.e. dark cells) are mostly distributed along lines that are parallel to the horizontal axis. These dark cells indicate high similarities in pairs of intervals with the same temporal distance between each other. For example, in

Figure 4-5, Figure 4-6, and Figure 4-7, the lower line of dark cells shows that movements in an interval are very similar to movements in another interval that is 0.92 s away from it. That is, the dancers regularly repeat similar movements every 0.92 s.

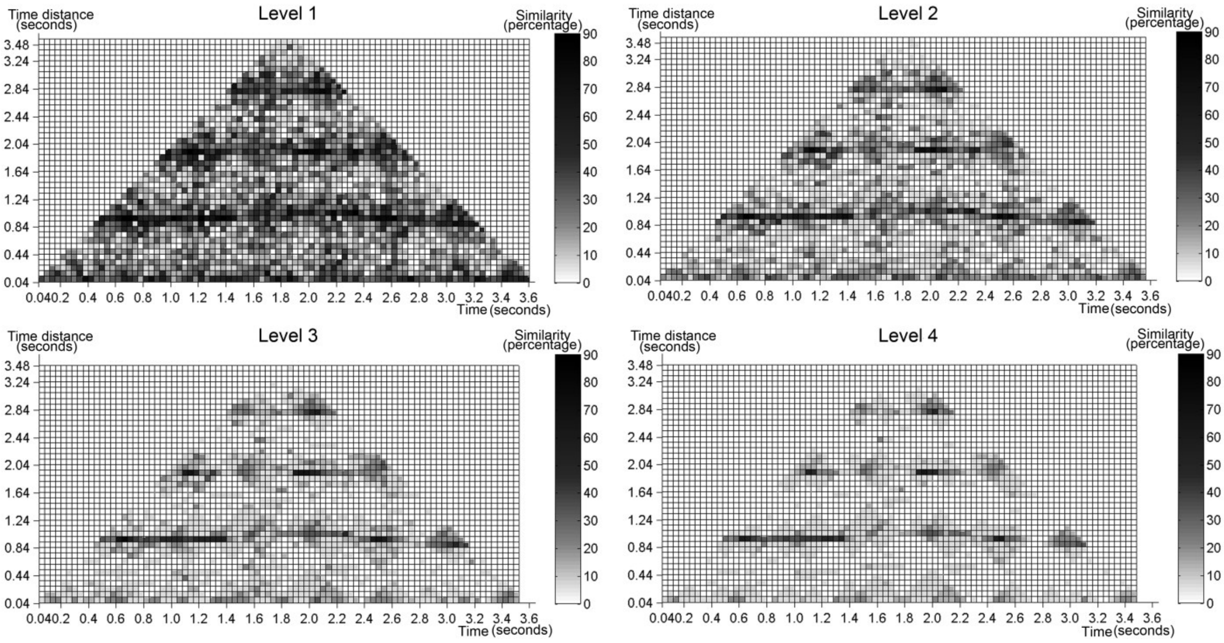


Figure 4-5: Levels 1 to 4 of the CTM of student 1 with 0.04 s time granularity

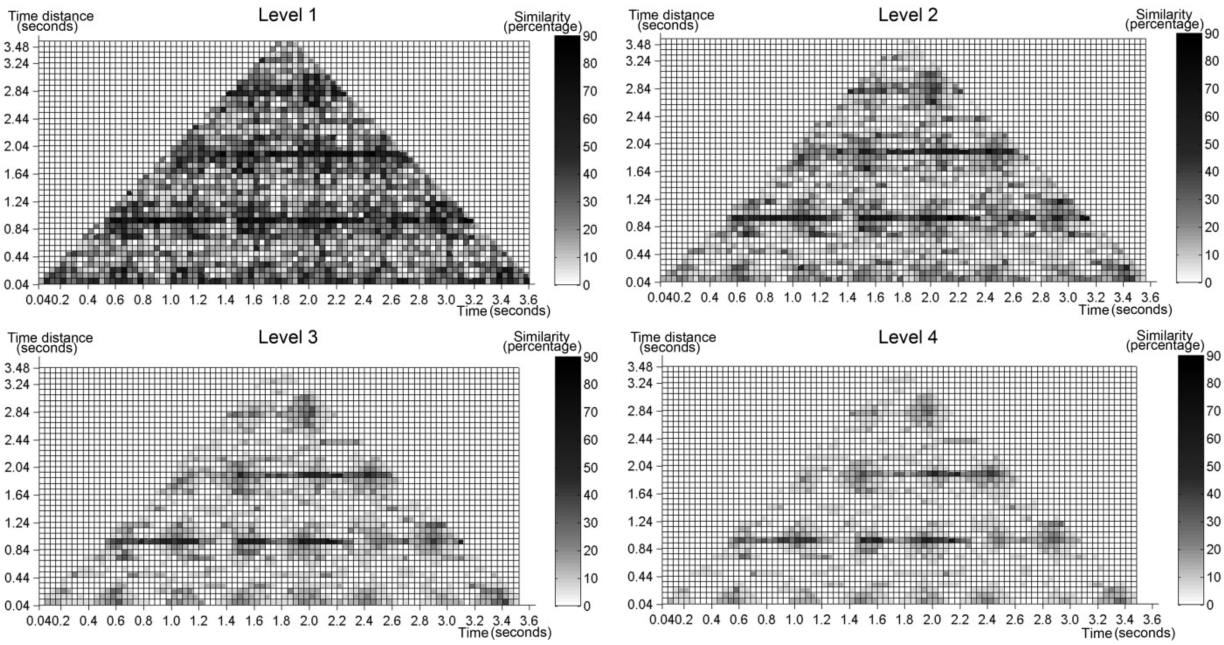


Figure 4-6: Levels 1 to 4 of the CTM of student 2 with 0.04 s time granularity

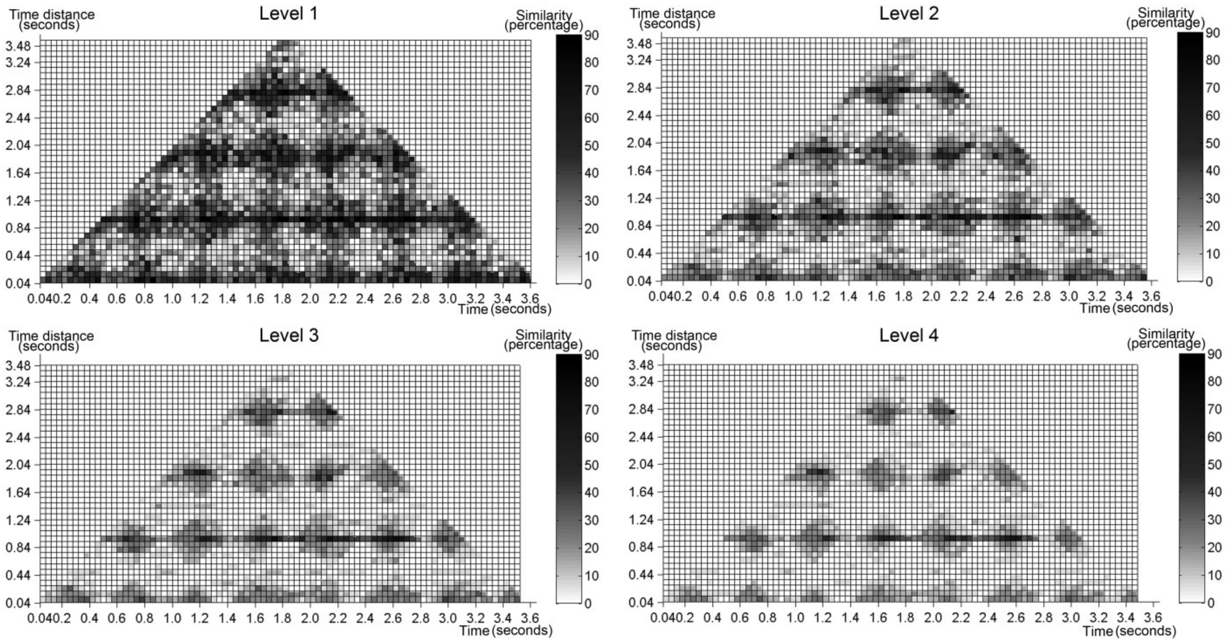


Figure 4-7: Levels 1 to 4 of the CTM of the teacher with 0.04 s time granularity

4.3.1.5 Interpretation of Motion Patterns

The results show some differences between the CTM of the teacher and the CTMs of the students. In the CTM of the teacher (Figure 4-7), dark similarities are strictly distributed along the line at 0.92 s. This indicates that the movements of the teacher are regularly repeated every 0.92 s. However, in the CTMs of students 1 and 2 (Figure 4-5 and Figure 4-6), the dark lines are not straight, compared with that of the teacher. Some parts of the dark line are located above or below the 0.92 s line. This is because there are some lag and lead times in the repetition of the same movements. From this observation, we can infer that the movements of students 1 and 2 are not as regular as the movements of the teacher. We also show some of the body configurations of student 1 and the teacher every 0.92 s in Figures 4-8 and 4-9. These visualisations are based on the MoCap toolbox (Toiviainen & Burger, 2010). The results show that student 1 and the teacher have an almost identical body configuration every 0.92 s. However, there are some time differences between the teacher and student 1 when performing the same movements.

4.3.2 Mini-Soccer

The relations between the movements of different mini-soccer players during a 13.33 m (i.e. minutes) portion of a game are analysed. The trajectories of the mini-soccer players are sampled from video frames taken at regular time stamps, with a temporal granularity of one second, by automatic computer vision-based tracking with field expert supervision. The procedure used for the Samba dancers is applied to identify the repetitive motion patterns of the mini-soccer players. All players' coordinates are extracted, and QTC_B relations between players are then formed at each time stamp. There are five players on each team; hence, different possible combinations of players form the QTC_B sequences, i.e. combinations of two, three, four or five players.

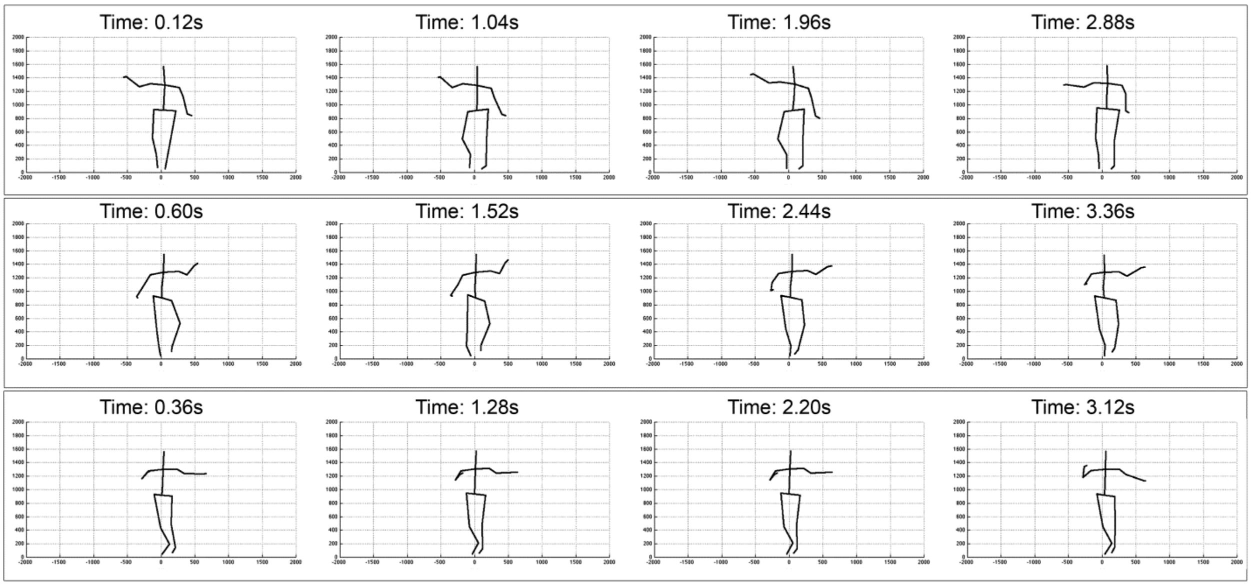


Figure 4-8: Some body configurations of student 1 every 0.92 s

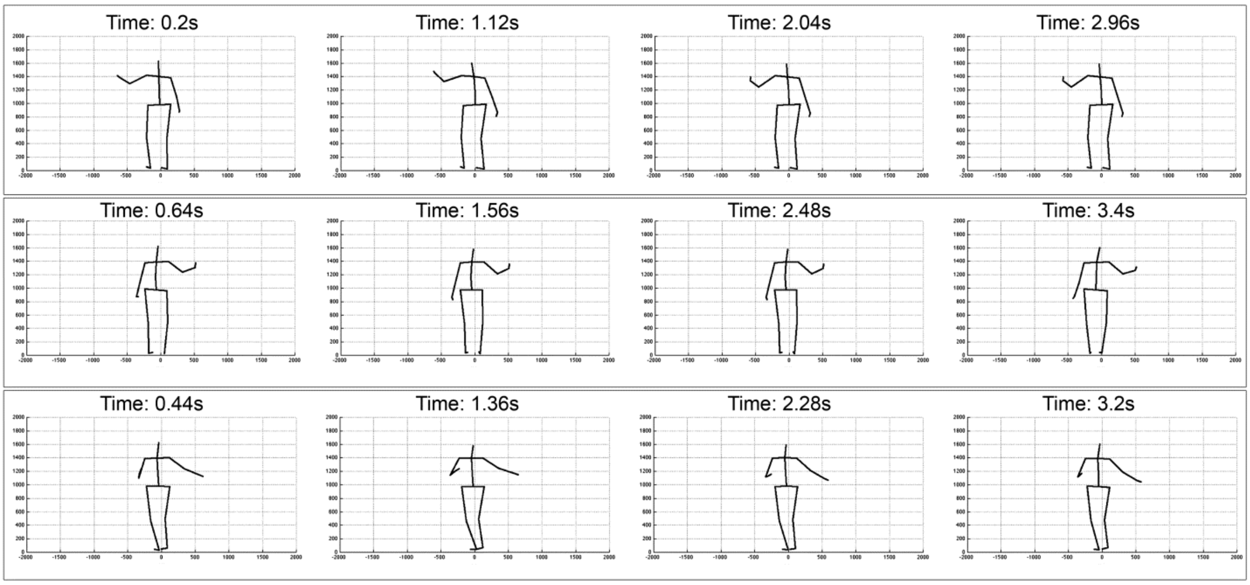


Figure 4-9: Some body configurations of the teacher every 0.92 s

A similarity analysis is performed between the movement sequences of the mini-soccer players for different combinations. Compared with the previous case study, the frequency of repetitive motion patterns in mini-soccer is much lower, as shown in the CTMs in Figures 4-10, 4-11, and 4-12. In these figures, the CTMs of three, four and five mini-soccer players of one team are presented for levels 1, 2, 3, and 4. Identifying the repetitive motion patterns from the CTMs of mini-soccer players is much less straightforward than in the case of the Samba dancers because fewer repetitive motion patterns occur in mini-soccer than in Samba dance. A higher CTM level indicates that fewer similar motion patterns are observed and that the observed motion patterns are less similar.

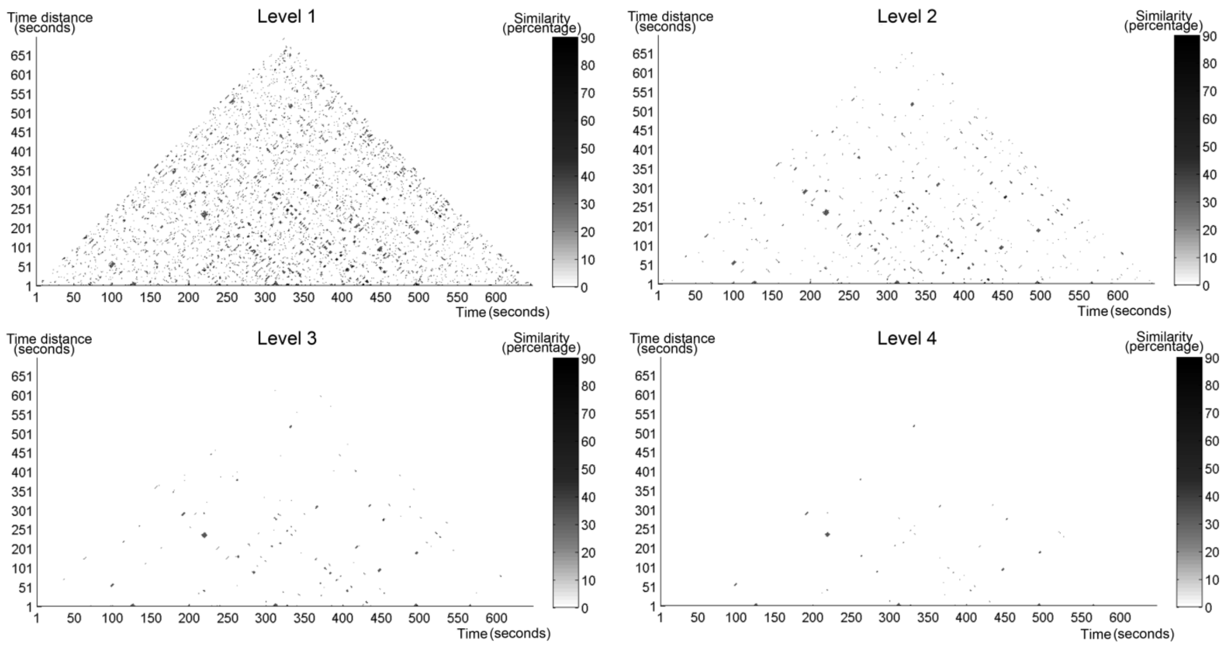


Figure 4-10: First four levels of the CTM representations of three mini-soccer players

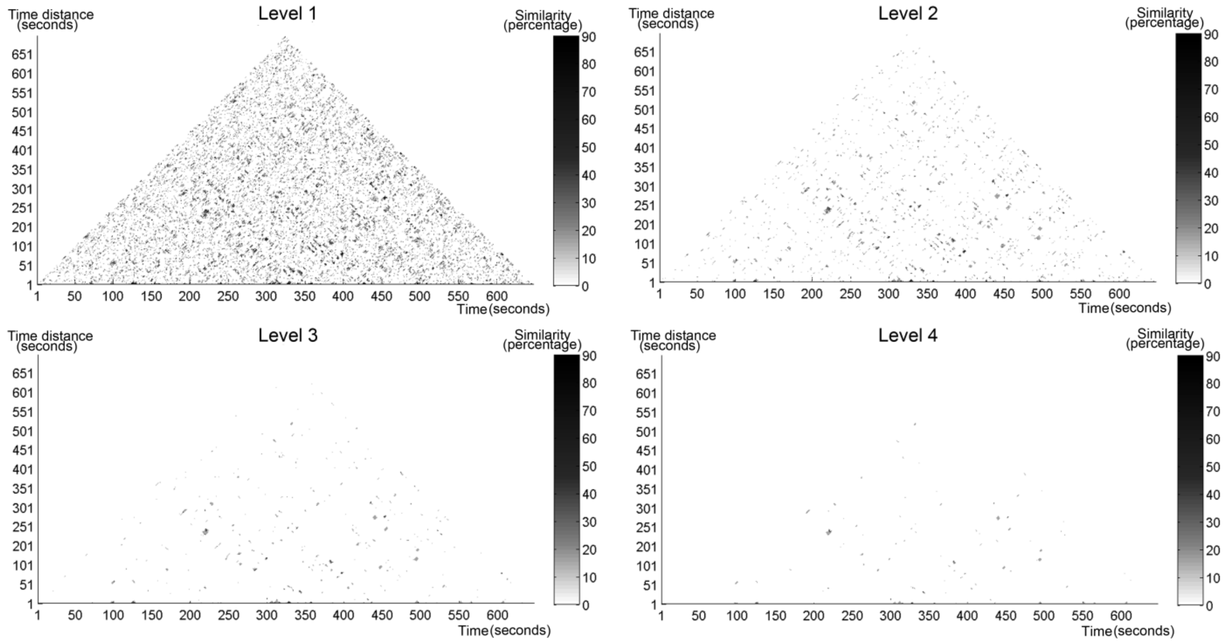


Figure 4-11: First four levels of the CTM representations of four mini-soccer players

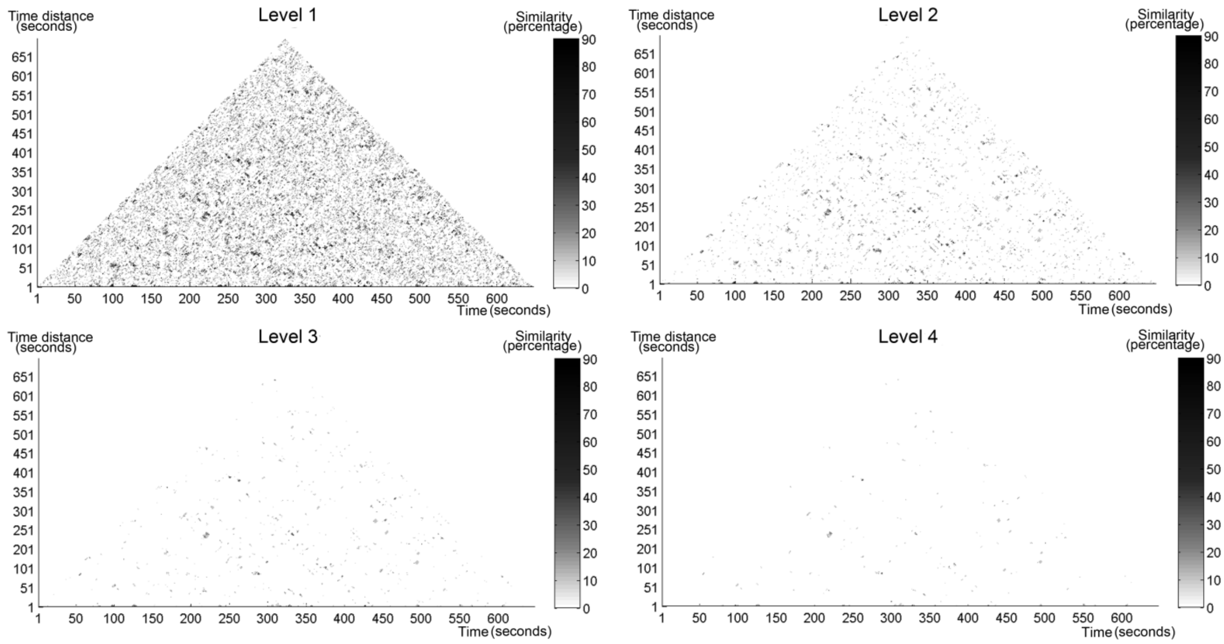


Figure 4-12: First four levels of the CTM representations of five mini-soccer players

4.4 Conclusions and Outlook

This chapter has proposed a three-tiered methodology to identify, visualise and interpret repetitive motion patterns in groups of MPOs. Movements of multiple MPOs are described in terms of sequences of QTC_B matrices, which in turn are used to identify the repetitive motion patterns. Next, similarity analysis is used to determine the degrees of similarity between pairs of movement sequences. Finally, CTM is applied to display the degrees of similarity between all pairs of movement sequences.

- The usefulness of the proposed methodology has been discussed in two real-world movement cases, i.e. Samba dance and mini-soccer. While the current chapter provides an intuitively appealing approach for studying repetitive movements of moving objects, the following aspects warrant further exploration in future work:
- Time granularity plays an important role in revealing the details of movement. The trajectories captured with the finest time granularity show more details of movement. It would be worthwhile to compare the results obtained from different time granularities.
- QTC_B relations are built based on changing Euclidean distances between two MPOs. In addition, directional information can also be considered to identify motion patterns using QTC Double-Cross (QTC_C). QTC_C provides more detail than QTC_B but increases the problem complexity.
- In the calculation of the similarity between QTC matrices, cell-by-cell comparison is made with the assumption that all cells are treated the same way. Some relations between the MPOs might be more important than others. These differences can be incorporated by assigning specific weights to each of those relations.
- *Map Algebra* (i.e. a set of algebraic operations applied on two or more raster layers with the same dimensions to produce a new raster layer) might be applied to infer additional results by comparing CTMs at different levels.

We hope to report on these and other aspects of movement pattern recognition and mining in the near future.

References

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Brakatsoulas, S., Pfoser, D., & Tryfona, N. (2004). Modelling, storing, and mining moving object databases. *Proceedings of the International Database Engineering and Applications Symposium* (pp. 68-77).

DeCesare, N. J., Squires, J. R., & Kolbe, J. A. (2005). Effect of forest canopy on GPS-based movement data. *Wildlife Society Bulletin, 33* (3), 935-941.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Dodge, S., Weibel, R., & Lautenschutz, A. K. (2008). Towards a taxonomy of movement patterns. *Information Visualization, 7* (3-4), 240-252.

Egenhofer, M. J., & Altaha, K. K. (1992). Reasoning about gradual changes of topological relationships. In: A. U. Frank, I. Campari & U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. (pp. 196-219).

Frank, A. U. (1991). Qualitative spatial reasoning about cardinal directions. *Technical Paper : ACSM-ASPRS Annual Convention, 6*, 148-167.

Freksa, C. (1992b). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Giannotti, F., Pedreschi, D., & Turini, F. (2009). Mobility, data mining and privacy the experience of the GeoPKDD project. In: B. Francesco, F. Elena, J. Wei & M. Bradley (Eds.), *Privacy, Security, and Trust in KDD. (pp. 25-32).* Springer-Verlag.

Hvidberg, M. (2006). Tracking human exposure to ultrafine particles in Copenhagen using GPS. *Epidemiology, 17* (6), S38.

Laube, P., Dennis, T., Forer, P., & Walker, M. (2007). Movement beyond the snapshot - Dynamic analysis of geospatial lifelines. *Computers, Environment and Urban Systems, 31* (5), 481-501.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Michael, K., McNamee, A., Michael, M. G., & Tootell, H. (2006). Location-based intelligence - Modelling behaviour in humans using GPS. *Proceedings of the IEEE International Symposium on Technology and Society* (pp. 97-104).

Naveda, L. (2011). *Gesture in Samba: A Cross-Modal Analysis of Dance and Music from the Afro-Brazilian Culture.* Ghent University, Ghent.

Naveda, L., & Leman, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (TGA). *Music Perception, 28* (1), 93-111.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic-based on regions and connection. In B. Nebel, C. Rich & W. R. Swartout (Eds.), *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR 92)* (pp. 165-176).

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Portoa, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data & Knowledge Engineering, 65* (1), 126-146.

Toiviainen, P., & Burger, B. (2010). MoCap Toolbox Manual. University of Jyväskylä, Jyväskylä.

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.*. Ghent University, Ghent.

Van de Weghe, N., Cohn, A. G., De Tre, G., & De Maeyer, P. (2006). A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems. *Control and Cybernetics, 35* (1), 97-119.

Van de Weghe, N., & De Maeyer, P. (2005). Conceptual neighbourhood diagrams for representing moving objects. *Perspectives in Conceptual Modelling, 3770*, 228-238.

Van de Weghe, N., Kuijpers, B., Bogaert, P., & De Maeyer, P. (2005b). A qualitative trajectory calculus and the composition of its relations. In: M. A. Rodriquez, I. F. Cruz, M. J. Egenhofer & S. Levashkin (Eds.), *Geospatial Semantics*. (pp. 60-76) Berlin Heidelberg: Springer-Verlag.

Wang, L. A., Hu, W. M., & Tan, T. N. (2003). Recent developments in human motion analysis. *Pattern Recognition, 36* (3), 585-601.

5

The Sequence Signature of Patterns

Make everything as simple as possible, but not simpler. Albert Einstein

Modified From: Seyed Hossein Chavoshi, Bernard De Baets, Tijs Neutens, Matthias Delafontaine, Guy De Tré, Nico Van de Weghe (2014). Movement Pattern Analysis Based on Sequence Signatures. Under Review.

5 The Sequence Signature of Patterns

Abstract: The increasing availability and affordability of fine-grained trajectory data sets have spurred GIS and computer scientists to push forward the frontier of movement pattern analysis. This chapter proposes a novel approach to identifying, representing and clustering patterns of relative motion between two disjoint moving point objects (MPOs). The approach consists of three major steps. First, relative motion is summarised in terms of qualitative relationships based on Euclidean distance between MPOs. Second, sequences of these qualitative relationships are represented in a sequence signature (SESI). A SESI is a fractal feature and provides a visual summary of the movement patterns of two MPOs. Third, SESIs are compared using a distance measure, making it possible to identify clusters of qualitatively distinct trajectory pairs. The proposed method is illustrated using two real-world examples of interacting MPOs: cars on a highway and squash players. These simple examples show the usefulness of our approach for uncovering movement patterns that are hidden in spatio-temporal trajectory databases.

5.1 Introduction

The increasing deployment of location-aware devices has given rise to an unprecedented wealth of trajectory information, documenting the movements of various types of moving objects, including vehicles (Haghani et al., 2009), animals (Cagnacci et al., 2010), bank notes (Brockmann et al., 2006), sportspersons (Wisbey et al., 2010), and tourists (Tiru et al., 2010). During the past two decades or more, the increased availability and affordability of these fine-grained data sets have aroused a burgeoning interest among (geographical) information scientists, who have steadily begun to develop and implement tools to discover, aggregate, and cluster meaningful patterns of individual or group behaviour in space-time (Ahlqvist et al., 2010; Delafontaine et al., 2012; Dodge et al., 2008; Gudmundsson et al., 2007; Laube et al., 2005; Shoval & Isaacson, 2007; Wang et al., 2012; Wilson, 1998). One specific area of interest concerns the line of inquiry that has developed qualitative formalisms to reason about moving objects. Adopting a qualitative approach implies that continuous information is being discretised by landmarks that classify neighbouring open intervals into discrete quantity spaces (Weld & Kleer, 1989). Key to this approach is that a distinction is introduced only if it is relevant to the research context at hand (Clementini et al., 1997; Cohn, 1996). The impetus for developing qualitative formalisms is that qualitative information aligns better with human intuition, communication and decision making than quantitative information (Egenhofer & Mark, 1995; Monferrer & Lobo, 2002; Renz et al., 2000), as illustrated in the following example taken from Clementini, et al. (1997, p. 318): "Saying that Alaska is 1, 518, 800 km^2 is sufficiently exact quantitative information about size and distances in Alaska, but very likely it is not meaningful in relation to the spatial knowledge of the average listener. On the other hand, saying that Alaska alone is bigger than all the states of the East coast from Maine to Florida is cognitively more immediate". As this example illustrates, spatial reasoning in our everyday interaction with the physical world is driven primarily by qualitative abstractions of the (often too precise) quantitative space (Cohn & Hazarika, 2001). By abstracting away from metrical details, qualitative representations are also much more appropriate for addressing incomplete information than quantitative methods (Cristani et al., 2000).

Much of the work in the area of qualitative reasoning concentrates on either the spatial or the temporal dimension of objects or phenomena. This observation is true for, among others, Interval Algebra (Allen, 1983), Point Algebra (Vilain et al., 1989), Double-Cross Calculus (Freksa, 1992b), Region Connection Calculus (Randell et al., 1992) and Oriented Point Reasoning Algebra (Moratz et al., 2005). Only a few studies, however, have systematically examined the qualitative properties of spatio-temporal information, such as the trajectory data of moving objects. One of these studies includes Qualitative Trajectory Calculus (QTC) proposed by Van de Weghe (2004). QTC is a powerful formalism for representing and reasoning about interactions between two moving point objects (MPOs). In QTC, the interactions are considered to be the changes in the Euclidean distance between two MPOs. While insightful, QTC has, until recently, remained largely conceptual, and new methods for inferring meaningful knowledge from qualitative spatio-temporal information remain sorely needed (Delafontaine et al., 2011). Hence, this chapter seeks to extend previous accomplishments in this area by developing a novel methodology to cluster trajectory pairs based on corresponding (repetitive) sequences of qualitative relationships between two MPOs.

Clustering enables to group data according to similarity into meaningful clusters. In particular, spatio-temporal clustering is a procedure for clustering objects based on their spatial and temporal similarities (Kisilevich et al., 2010). Figure 5-1, modified from (Kisilevich et al., 2010), visually depicts a possible classification of spatio-temporal data types. In this study, we focus on the clustering of trajectory data types and account for pointwise moving objects. Trajectories define the movement behaviour of moving objects, and consequently, the clustering of trajectories contributes to detecting groups of moving objects that have similar movement behaviour. Trajectory clustering is primarily application dependent. Until now, many attempts have been conducted in the field of trajectory clustering. Kisilevich et al (2010) has categorised trajectory clustering approaches into one of the following groups: Descriptive and generative model-based clustering (e.g., (Alon et al., 2003; Chudova et al., 2003; Gaffney & Smyth, 1999), distance-based clustering methods (e.g., (Nanni & Pedreschi, 2006; Pelekis et al., 2007)), density-based methods and the DBSCAN family (e.g., (Nanni & Pedreschi, 2006)), visual-aided approaches (e.g., (Andrienko & Andrienko, 2006)), micro clustering methods (e.g., (Hwang et al., 2005)), flocks and convoy (e.g., (Gudmundsson & van Kreveld, 2006; Kalnis et al., 2005)), important places (e.g., (Kang et al., 2004)), and pattern-based clustering (e.g., (Giannotti et al., 2007)). The presented approach could fall into pattern-based clustering. Patterns that are extracted from trajectories are referred to as trajectory patterns and express interesting behaviour of individual objects or multiple moving objects (Giannotti & Pedreschi, 2008). Different approaches exist in the mining of trajectory patterns. We use QTC to extract movement patterns of interactions of pairs of MPOs for clustering purposes (Figure 5-1). Additionally, a visually aided approach is applied to recognise the extracted patterns and to facilitate the understanding of the nature of the movement patterns that occur.

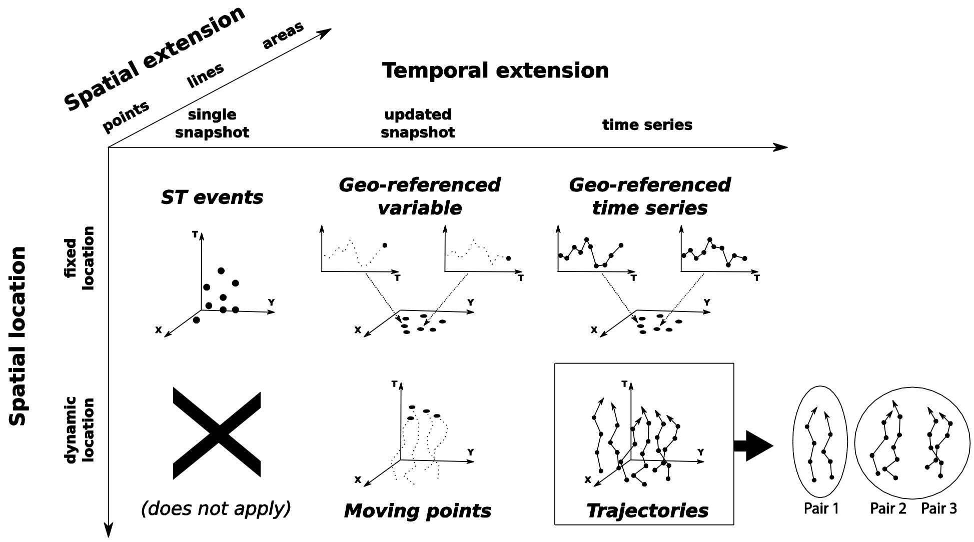


Figure 5-1: Spatio-temporal clustering modified from (Kisilevich et al., 2010)

The remainder of this chapter is organised as follows. The next section introduces the proposed approach for identifying, representing and analysing patterns of relative motion between two MPOs. This section is subdivided into three subsections, which correspond to the three major steps of the method: (i) description of movement patterns using QTC (Subsection 5.2.1), (ii) representation of movement patterns in a SESI (Subsection 5.2.2), and (iii) distance measurement and hierarchical clustering of multiple SESIs (Subsection 5.2.3). Section 5.3 then demonstrates the usefulness of our methodology in identifying and clustering the movement patterns of cars and squash players. Section 5.4 presents a detailed discussion. Lastly, Section 5.5 summarises and reflects on the major findings here and outlines avenues for future research.

5.2 Methodology

Moving objects constitute a principal unit of analysis in many major domains of both theoretical and applied scientific research, including geographical information science (GIScience), artificial intelligence, knowledge representation, sports science, and transportation. In particular, a number of qualitative calculi have been developed to reason about moving objects. This chapter examines whether the Qualitative Trajectory Calculus (QTC), a qualitative formalism for reasoning about motion, can be employed to cluster moving objects' trajectories based on their relative motions. In contrast to prior research that is related to the trajectory clustering of moving objects, which was primarily restricted to geometric summaries of trajectories, the present research is unique in that it specifically considers the interaction patterns of pairs of MPOs.

Studying the spatial interactions of MPOs over time is of interest in various research applications such as in sport sciences and choreography. The methodology comprises three major steps. First, raw trajectory data that stems from location-aware technologies is converted into qualitative relationships based on the Euclidean distance between two interacting MPOs. Second, drawing on the concept of an iterated function system, sequences of these qualitative relationships (i.e. movement patterns) are visually summarised in a sequence signature (SESI). Lastly, multiple SESIs are compared using a distance measure, which in turn is used to identify hierarchical clusters of qualitatively distinct trajectories. Each of these steps is discussed in depth below.

5.2.1 Step 1: Converting Raw Trajectory Data into Qualitative Relationships

To summarise raw trajectory data in terms of qualitative relationships, we use the Qualitative Trajectory Calculus (QTC) (Van de Weghe, 2004). In this study, we focus on QTC_B because it is the simplest type from which all of the other types are derived. The approach can *mutatis mutandis* also be applied for the other types, but this step is left for future work.

Discussed in Section 2.5, in QTC_B, a qualitative relationship between two objects *k* and *l* at a time stamp *t* is defined by a label that is composed of two characters *A* and *B*. The resulting syntax for QTC_B relation is the tuple (A B) that considers the distance constraints described in Section 2.5 (relationships A and B). In total, this approach yields 9 (3^2) base relations, which are represented in Figure 2-3. The nine represented relationships form a set of jointly exhaustive and pairwise disjoint (JEPD) base relations. Consequently, at each time instant, there is one and only one QTB_B relation for each pair of coexisting MPOs. To make this strategy clear, for example, QTC_B relations (0 +), (0 –), (+ 0), and (– 0), are explained as follows.

(0 +) : *k* is stable with respect to *l,* and *l* is moving away from *k*

(0 –) : *k* is stable with respect to *l,* and *l* is moving towards *k*

(+ 0) *:* *k* is moving away from *l,* and *l* is stable with respect to *k*

(– 0) *:* *k* is moving towards *l,* and *l* is stable with respect to *k*

A prototypical example of interaction between moving objects is the overtake event (see also (Van de Weghe et al., 2005)). Consider the interactions between two vehicles *k* and *l* during a time interval. The movements of these vehicles can be expressed by a conceptual animation, in other words, a chronological sequence of continuous transitions between the QTC_B relations. Suppose that car *k* is going to overtake car *l*. This interaction can be expressed in terms of the QTC_B relationships as follows:

State 1. car *k* and car *l* are both driving in the same traffic lane, and *k* is driving behind *l*: $(-\ +)$

State 2. car *k* is heading out to the second lane: $(-\ +)$

State 3. car *k* is driving in the second lane and is driving behind *l*, which is driving in the first lane: $(-\ +)$

State 4. car *k* is driving in the second lane and is driving in front of car *l*, which is driving in the first lane: $(+\ -)$

State 5. car *k* is heading back to the first lane: $(+\ -)$

State 6. car *k* and car *l* are both driving in the first lane, and *l* is driving behind *k*: $(+\ -)$

It is noted that, following Galton’s theory of dominance (2001) and Forbus' equality change law (1984), a direct change from – to + and vice versa is impossible, because such a change must pass the qualitative value 0. This landmark value 0 needs to hold only for an instant. Therefore, there must be another relationship between State 3 and 4, namely, (0 0) (see also Van de Weghe & De Maeyer, 2005). Lastly, the above overtake event is represented by the following movement pattern:

$$\{(-\ +) \rightarrow (-\ +) \rightarrow (-\ +) \rightarrow (0\ 0) \rightarrow (+\ -) \rightarrow (+\ -) \rightarrow (+\ -)\} \qquad \text{Eq. 5-1}$$

We use the earlier implementation prototype called *QTCAnalyst* (Delafontaine et al., 2011) to identify the qualitative movement patterns of MPOs. The next section presents a method to transform these QTC movement patterns into a fractal-based representation.

5.2.2 Step 2: Summarising Movement Patterns in a Sequence Signature

A sequence signature (SESI) is a fractal way of mapping patterns of the interactions between two MPOs in an indexed raster space. It is based on an iterated function system, as discussed in detail in Barnsley (1988). In a SESI, each sequence of *n* consecutive qualitative relationships is represented by a cell of length *n* (i.e. the number of subsequent QTC_B relationships that constitute a movement pattern). Thus, a SESI of length 1 represents movement patterns that are composed of one QTC_B relation, while, for example, a SESI of length 5 represents movement patterns that are composed of five consecutive QTC_B relations. The resolution of these cells depends on the length of that SESI.

Figure 5-2 shows SESIs of lengths 1 and 2. The SESI of length 1 shows the nine base QTC_B relationships. For higher lengths, each cell is further subdivided into nine cells, allowing each cell in a SESI of length n to correspond with a unique sequence of n qualitative QTC_B relationships. In this way, each sequence of QTC_B relationships has a specific location in the SESI. For example, the highlighted cells in the SESIs of length 1 and 2 in Figure 5-2 represent movement patterns $\{(-\,+)\}$, $\{(-\,+) \rightarrow (0\,+)\}$, respectively. In QTC_B, a SESI of length n contains 9^n cells.

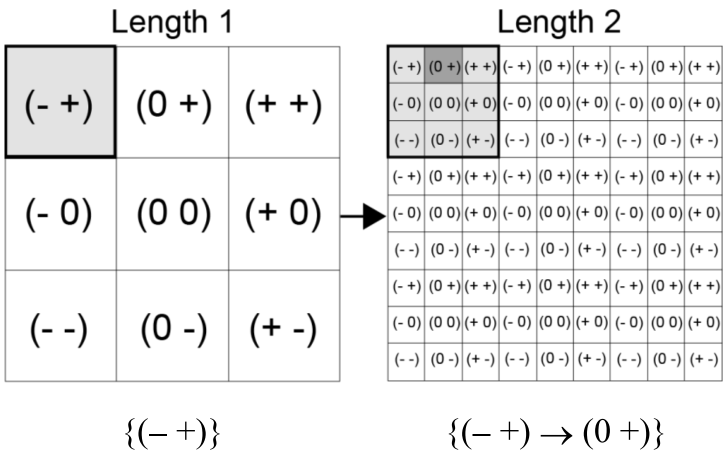
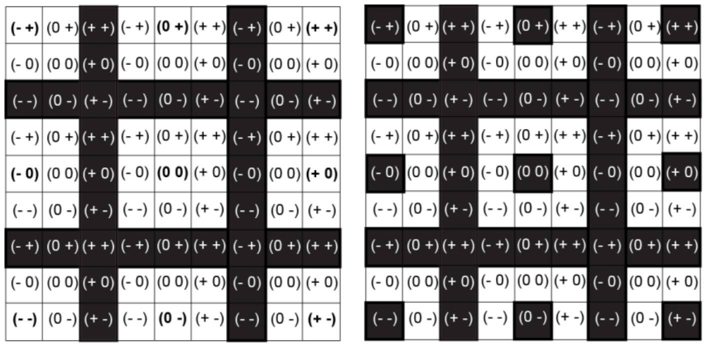


Figure 5-2: SESIs of length 1 and 2

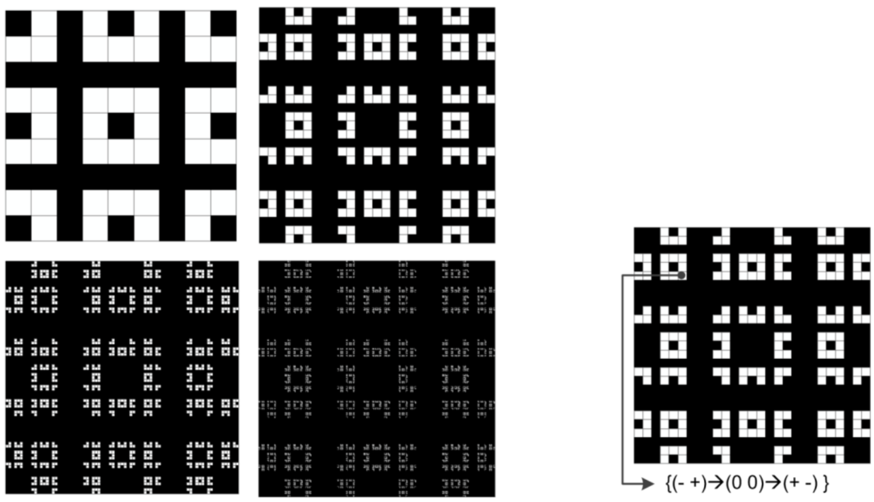
However, not all sequences of QTC_B relationships are possible or significant. First, based on the laws of continuity (see Subsection 2.3.2), we can exclude chronologically impossible combinations of QTC_B relationships in SESIs of length 2 or more. Figure 5-3a demonstrates SESIs of length 2 after imposing the continuity constraint.



(a) (b)

Figure 5-3: Continuity (a), and event-based constraints (b) imposed on a SESI of length 2

The black cells indicate discontinuous (and thus impossible) sequences of qualitative relationships between two MPOs. For example, it is not possible to have a direct transition from (– +) to relationships (+ +), (+ 0), (+ –), (0 –), and (– –). Second, because we are typically interested in the changes in the relative motion between the MPOs over time (i.e. events), some sequences of qualitative relationships are not significant. This arrangement is the case when a QTC_B relationship is invariant over time. For example, a transition from (– +) into itself is not very meaningful from a qualitative perspective. Figure 5-3b shows the result of imposing both of the restrictions above on a SESI of length 2. For illustrative purposes, we also provide the representations of SESIs of length 2 to 5 in Figure 5-4a. Figure 5-4b shows the SESI representation (length 3) of the overtake event given in Eq. 5-1.



(a) (b)

Figure 5-4: SESIs of lengths 2 to 5 (a), and the overtake event on a SESI of length 3 (b)

The fractal approach outlined above offers an insightful way to address repetitive movement patterns of any length.

5.2.3 Step 3: Clustering Trajectory Pairs Based on Their Relative Movements

Prior to clustering pairs of trajectories of MPOs based on their relative movements, we need some sort of measure to determine to what extent two SESIs are different. A comprehensive review of different clustering techniques has been performed by Rokach and Maimon (2010). Many clustering techniques employ distance measures to specify the distance between pairs of objects. We used a modified distance function that is based on the structure of the SESIs to calculate the distance between pairs of SESIs. We measure the distance $d_n(S_1, S_2)$ between two SESIs (S_1, S_2) of length n as follows:

$$d_n(S_1,S_2) = \sum_{i,j=1}^{3^n} \frac{(S_{1,N_{ij}} - S_{2,N_{ij}})^2}{3^{2*n} - \alpha_n} \qquad \text{Eq. 5-2}$$

where

$$S_{1,N_{ij}} = {S_{1,ij}}/{U_{ij}} \text{ and } S_{2,N_{ij}} = {S_{2,ij}}/{U_{ij}} \qquad \text{Eq. 5-3}$$

denote the normalised frequencies $S_{1,ij}$ and $S_{2,ij}$ of the mapped movement patterns in the ij^{th} cell of S_1 and S_2, respectively;

α_n denotes the number of impossible or insignificant cells in a SESI of length n; and

U_{ij} denotes the value of the highest frequency of the ij^{th} cell among all of the SESIs on which the distance measure is applied.

Normalisation is used to scale heterogeneous frequencies of movement patterns, to make it possible to compare them. As a result, the normalised frequency of each cell of a SESI is confined to the interval [0, 1]. To calculate the distance between two SESIs, we start from the top left cells in both SESIs to subtract their normalised frequencies from each other. The distance measure in Eq. 5-2 runs over all of the cells in the SESIs. The denominator in Eq. 5-2 represents the number of feasible sequences of QTC relationships in that SESI. By definition, infeasible sequences are assigned a frequency of 0. For example, there are 200 feasible cells in a SESI of length 3. The distance function in Eq. 5-2 ranges from 0 to 1, where 0 indicates that the SESIs are identical and 1 indicates that there is no correspondence at all between the SESIs. In fact, d_n is an indicator to show how much movement patterns of a pair of MPOs (with length n) are similar to movement patterns of another pair mapped on two distinct SESIs. As an illustration, Figure 5-5 depicts two arbitrary trajectory pairs of two objects during two different time intervals. The SESIs of length 3 for both trajectory pairs are shown in Figure 5-6. The black cells indicate impossible cells, whereas the green cells display the frequency of the movement patterns. In this simple example, the frequency of occurrence of a movement pattern is either 0 or 1. The distance between these two SESIs can be calculated based on Eq. 5-2:

$$d_3(S_1,S_2) = \sum_{i,j=1}^{3^3} \frac{(S_{1,N_{ij}} - S_{2,N_{ij}})^2}{3^{2*3} - 529} = 0.08$$

The low distance value indicates that the interaction patterns between the two MPOs are, according to QTC_B, qualitatively similar across both time intervals.

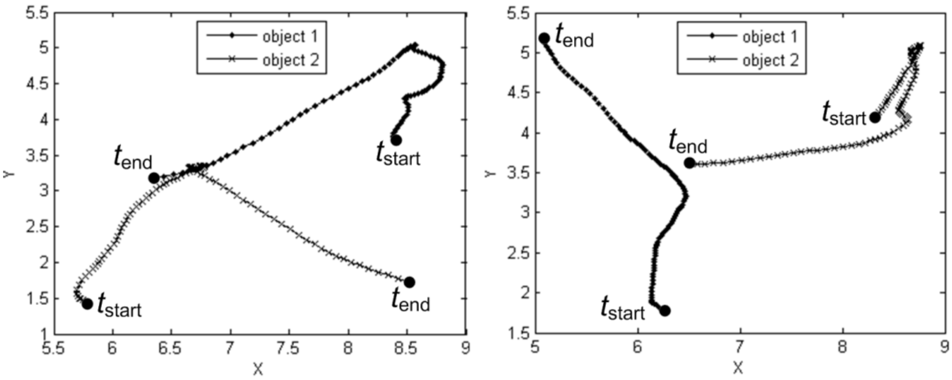


Figure 5-5: Two trajectory pairs during two different time intervals

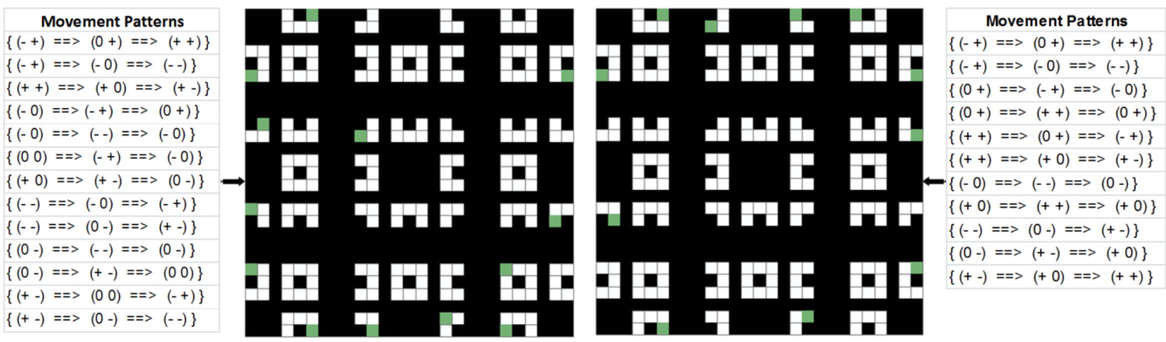


Figure 5-6: Movement patterns and SESIs for the trajectory pairs shown in Figure 5-5

In the case of multiple pairs of trajectories, we can apply a hierarchical clustering algorithm on the basis of the distance between multiple SESIs, to identify groups of trajectories that have similar interaction patterns. The idea behind hierarchical clustering is to build a binary tree of the data that successively merges similar groups of objects. This binary tree provides a visualisation of a useful summary of the data. In this study, we use an agglomerative hierarchical clustering algorithm, which yields a dendrogram that represents a nested grouping of objects (*in casu*, trajectory pairs). The horizontal axis in the dendrogram represents the clustered objects, while the vertical axis shows the distance between clusters. We use average-linkage to compute distances between the new cluster and each of the old clusters (A detailed explanation can be found in (Dawyndt et al., 2005)). The distance between two clusters is then applied as the average distance from any member of one cluster to any member of the other cluster. We will illustrate the three major steps outlined above in the next section.

5.3 Illustration

In this section, we will apply our methodology in two examples: a very simple overtake event and a more complicated interaction situation of squash players.

5.3.1 Example 1: Overtake Event on a Highway

Figure 5-7 illustrates five different traffic situations with two cars (*k* and *l*), where each situation lasts 7 s (time steps of 1 second). The first three situations are simple overtake events. Table 5-1 shows the movement patterns for all five trajectory pairs. Following our methodology, we must first transform the trajectories into sequences of qualitative relationships. These sequences are depicted in Table 5-1. We next convert the sequences into SESIs of length 2 and 3 (Figure 5-8). We then measure the distance between all of the pairs of SESIs of a specific length. These distance measurements are then placed in a (symmetric) distance matrix for each SESI length (Figure 5-9). We lastly group the trajectory pairs using a hierarchical clustering method (Figure 5-9). As presented in Table 5-1, the movement patterns of the first three situations (1, 2, and 3) are similar to each other, and the movement patterns of the last two situations (4 and 5) are the same. In fact, the distance matrix shows that the first three trajectory pairs are identical in terms of qualitative relationships (see Table 5-1).

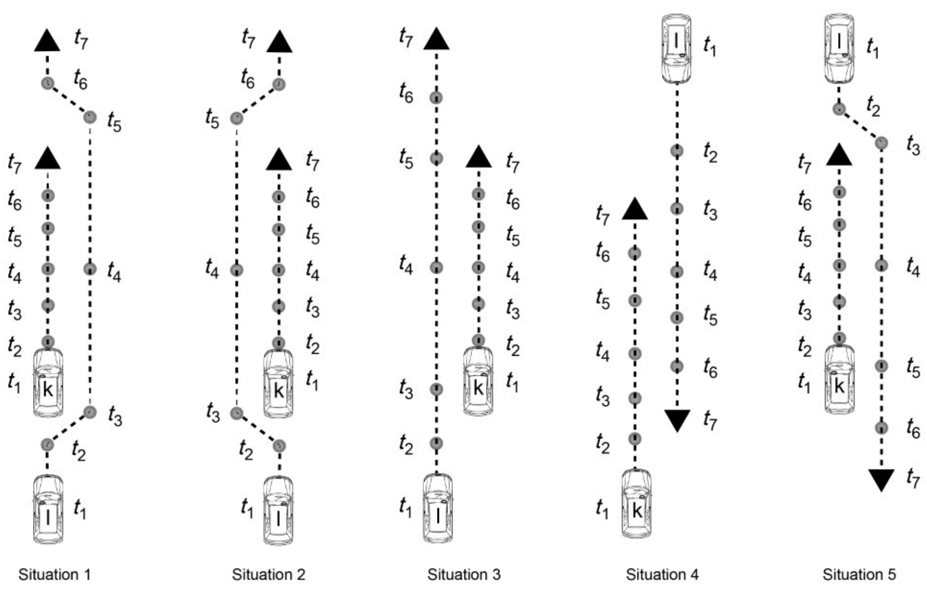


Figure 5-7: Five traffic situations

The trajectory pairs with similar movement patterns are then placed in the same clusters. Thus, the trajectory pairs 1, 2 and 3 are taken together in the same cluster, while the trajectory pairs 4 and 5 are grouped into another cluster for their given length. Here, cluster analysis allows for us to compare the movements of MPOs based on the QTC_B movement patterns extracted from the trajectory pairs. The distance value at each branching in the dendrogram represents the average distance between the SESIs in the branches.

Table 5-1: Movement patterns of the above five trajectory pairs between two cars, *k* and *l*

QTC Movement Patterns				
Situation 1	**Situation 2**	**Situation 3**	**Situation 4**	**Situation 5**
{ (+ -) ==> (0 0) }	{ (+ -) ==> (0 0) }	{ (+ -) ==> (0 0) }	{ (- -) ==> (0 0) }	{ (- -) ==> (0 0) }
{ (0 0) ==> (- +) }	{ (0 0) ==> (- +) }	{ (0 0) ==> (- +) }	{ (0 0) ==> (+ +) }	{ (0 0) ==> (+ +) }
{ (+ -) ==> (0 0) ==> (- +) }	{ (+ -) ==> (0 0) ==> (- +) }	{ (+ -) ==> (0 0) ==> (- +) }	{ (- -) ==> (0 0) ==> (+ +) }	{ (- -) ==> (0 0) ==> (+ +) }

Figure 5-9 represents only the hierarchical clustering of SESIs of length 2 and 3, respectively. The reason is that the maximum length of the considered sequences equals 3 in all of the situations presented in Table 5-1. In general, we can observe that the distance between the clusters with a shorter length is larger than the distance between the clusters with a longer length. The reason is that the frequencies of short movement patterns are higher than the frequencies of long movement patterns because the longer movement patterns are created from the shorter ones.

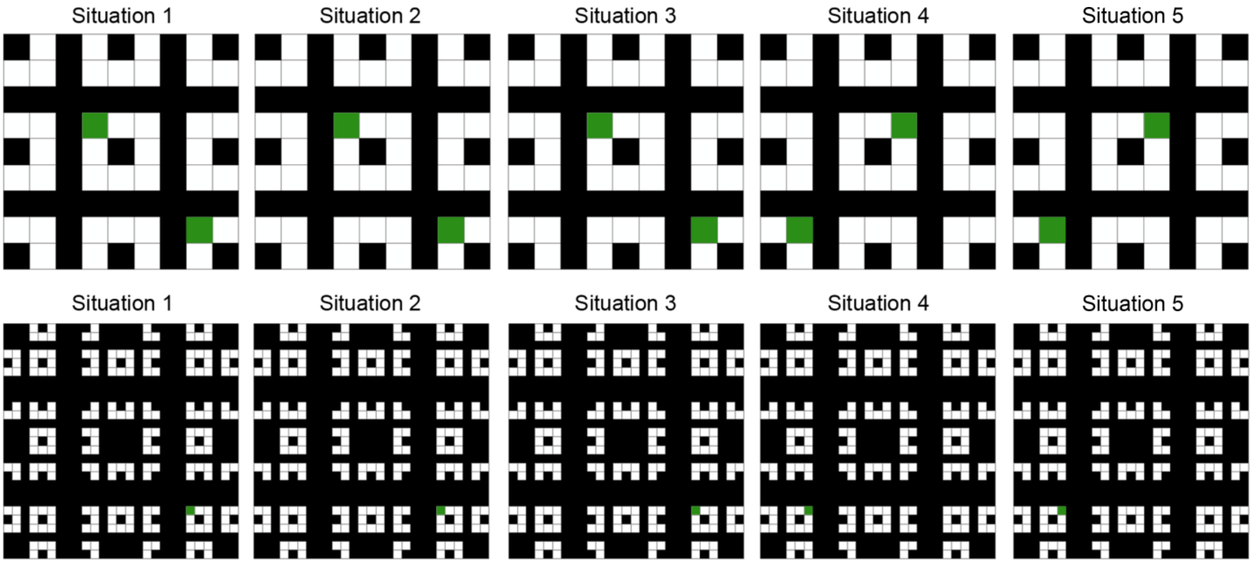
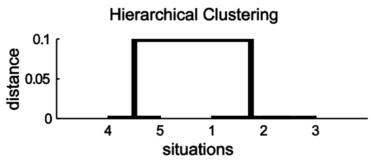


Figure 5-8: SESIs of lengths 2 and 3 for all of the situations

Length 2

Distance Matrix	Situation 1	Situation 2	Situation 3	Situation 4	Situation 5
Situation 1	0	0	0	0.1	0.1
Situation 2	0	0	0	0.1	0.1
Situation 3	0	0	0	0.1	0.1
Situation 4	0.1	0.1	0.1	0	0
Situation 5	0.1	0.1	0.1	0	0



Length 3

Distance Matrix	Situation 1	Situation 2	Situation 3	Situation 4	Situation 5
Situation 1	0	0	0	0.01	0.01
Situation 2	0	0	0	0.01	0.01
Situation 3	0	0	0	0.01	0.01
Situation 4	0.01	0.01	0.01	0	0
Situation 5	0.01	0.01	0.01	0	0

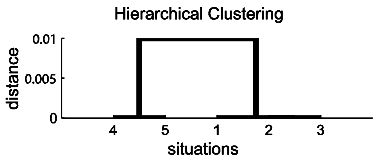


Figure 5-9: Distance matrices and hierarchical clustering of five traffic situations (five trajectory pairs) for SESIs of length 2 and 3 based on the distance function in Eq. 5-2

5.3.2 Example 2: A Squash Rally

In the second example, the qualitative relationships between two squash opponents are analysed. The data used is the public standard CVBase'06 dataset (Pers et al., 2006), which contains coordinates of two squash players derived from video frames taken at regular time steps (temporal resolution of 0.04 s). We consider the QTC_B relationships between both players during four different rallies of the game (Figure 5-10). Long rallies have movement patterns that have higher frequencies compared to shorter rallies. However, if two players interact with a high frequency during a short rally, that rally might give rise to movement patterns that have high frequencies as well. For the analysis, we considered rallies that have lengths equal to 8 s. The rallies were selected randomly from the game.

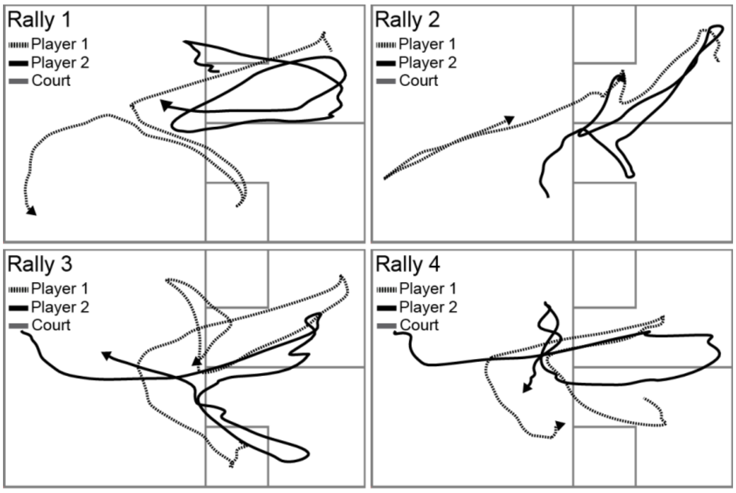


Figure 5-10: Interaction between two squash players in four rallies of a set of a game

The quantitative information of the trajectories was first transformed to QTC_B. As an example, the frequencies of the movement patterns of length 4 realised during the fourth squash rally are shown in Figure 5-11. It is observed that $\{(-\,+) \rightarrow (0\,+) \rightarrow (+\,+)\} \rightarrow (0\,+)\}$ is a frequently occurring movement pattern. To interpret it, one might express that at the beginning of the interaction, the first player was moving towards the second player, while the second player was moving away from the first. Afterward, the first player changed his direction of movement and moved away from the second player. Later in the interaction, the first player stood still, and the second player was still moving away from him.

Figure 5-12 illustrates the SESIs of length 2 to 4 of the four rallies. In addition, the distance matrices and dendrograms that result from the comparison of the SESIs are shown. From the SESIs, the complexity of the interactions of the squash players can be observed. Movement patterns mapped into the SESIs can reveal specific strategies taken by players. For example, the SESI of the third rally illustrates more frequent movement patterns than others. $\{(-\,0) \rightarrow (-\,+)\}$ and $\{(-\,-) \rightarrow (-\,0)\}$ are the most frequent movement patterns of length 2 in this rally. This arrangement could arise because one of the players has taken a strategy to hit the ball to certain parts of the front wall, which has caused the other player to move towards him repeatedly. This approach can be used to obtain knowledge from a large moving object database. In the squash game, for example, the strategies taken by a player can be examined by comparing the SESIs in different rallies and games.

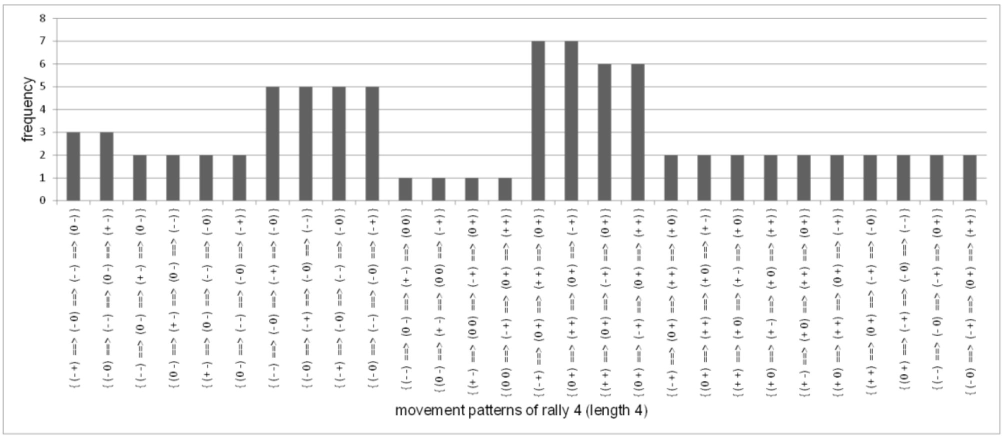
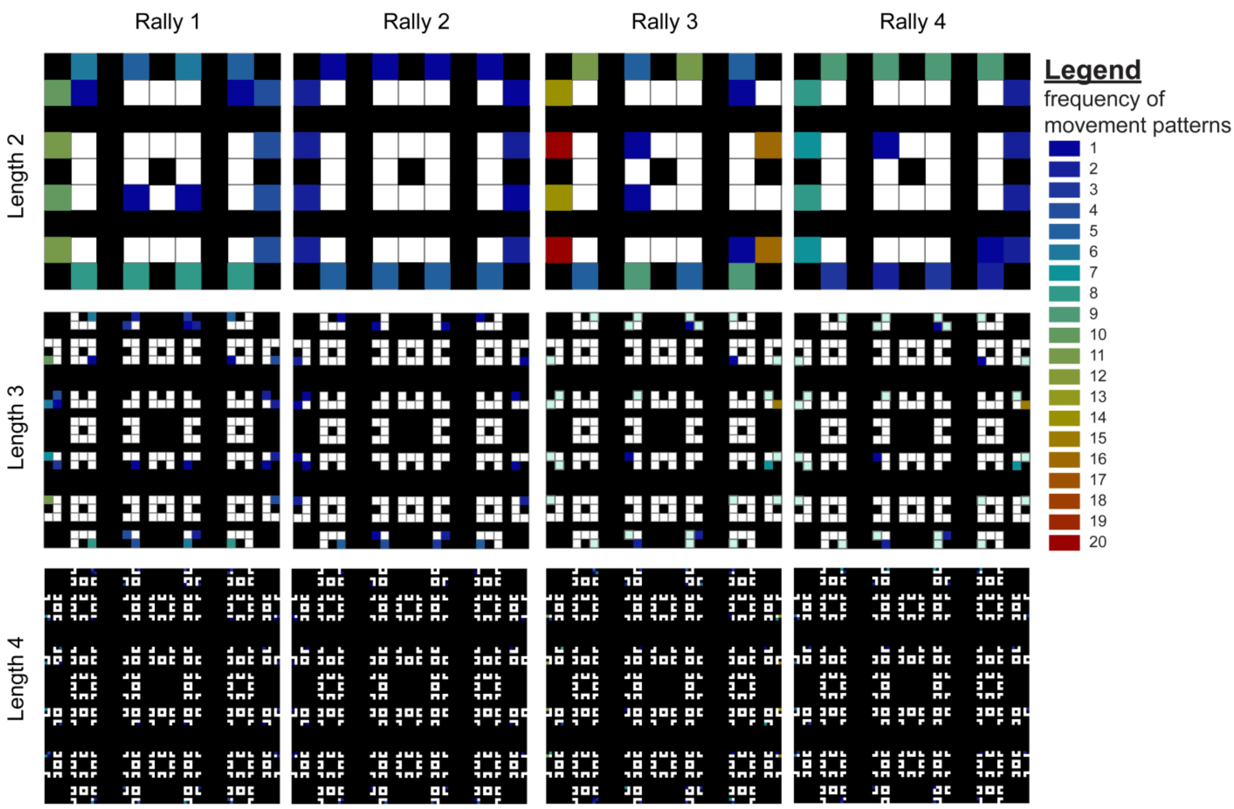


Figure 5-11: Frequency for the 28 movement patterns of length 4 of rally 4

Similar to the previously shown example of traffic events, in the squash game, the distance matrices are formed based on Eq. 5-2, and finally, the results are used for clustering purposes. The dendrograms show clusters of similar rallies, which share more common movement patterns. From the dendrograms presented in Figure 5-12, we see that the interactions of the squash players in rally 2 are more similar to those in rally 4 than in other rallies. This finding could support our understanding of the movement behaviour of MPOs during complex interactions based on qualitative relationships.



Length 2

Distance Matrix	Rally 1	Rally 2	Rally 3	Rally 4
Rally 1	0	0.16	0.19	0.2
Rally 2	0.16	0	0.32	0.13
Rally 3	0.19	0.32	0	0.2
Rally 4	0.2	0.13	0.2	0

Length 3

Distance Matrix	Rally 1	Rally 2	Rally 3	Rally 4
Rally 1	0	0.04	0.07	0.06
Rally 2	0.04	0	0.09	0.04
Rally 3	0.07	0.09	0	0.06
Rally 4	0.06	0.04	0.06	0

Length 4

Distance Matrix	Rally 1	Rally 2	Rally 3	Rally 4
Rally 1	0	0.01	0.02	0.02
Rally 2	0.01	0	0.02	0.01
Rally 3	0.02	0.02	0	0.02
Rally 4	0.02	0.01	0.02	0

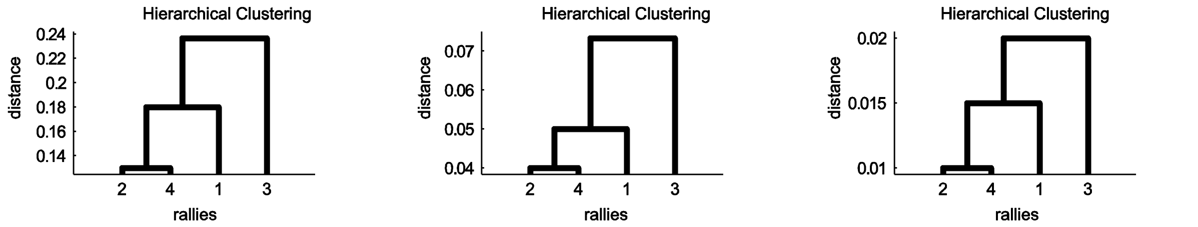


Figure 5-12: SESIs for rallies 1, 2, 3, and 4 of length 2, 3, and 4 along with the corresponding distance matrices and hierarchical clustering

5.4 Discussion

In this chapter, we proposed a method to cluster trajectory pairs of MPOs, and we demonstrated its applicability in two real-world examples. In what follows, we will elaborate on the strengths and weaknesses of our approach.

(a) According to (Kisilevich et al., 2010), "*analysis of movement behaviour is a complex process that requires understanding of the nature of the movement and phenomena it incurs. Automatic methods may discover interesting behavioural movement patterns with respect to the optimisation function but it may happen that these patterns are trivial or wrong from the point of view of the phenomena that is under investigation. Visual analytics field tries to overcome the issues of automatic algorithms introducing frameworks implementing various visualisation approaches of spatio-temporal data and proposing different methods of analysis including trajectory aggregation, generalisation and clustering*". In this study, SESIs provide a visual summary of the movement patterns of MPOs. The scattering of movement patterns that have occurred is visible in SESIs. The length of SESI is very influential to the scale of SESI. Therefore, the availability of fine-grained movement data at small temporal sampling intervals increases the difficulty of detecting QTC_B movement patterns in SESIs. That is, SESIs are used as a visual-aided approach to perceive the dispersion of movement patterns. However, we see the SESI mathematically as a rectangular array of numbers, symbols, or expressions, which are arranged in rows and columns. The individual items in this matrix (i.e. SESI) are referred to as its elements or entries (i.e. the frequency of QTC_B movement patterns). From a computational perspective, the longer the length of a SESI is, the more time is needed to compute the distance between two SESIs (time complexity $O(m^2 + 9^n)$; with m = length of chain of elements, n = length of SESI). In short, it is possible to calculate distances for any length of SESI, but visually detecting long-movement patterns in SESIs will become challenging.

(b) Throughout this study, two different examples were analysed to show that the presented approach is persuasive and reliable. In the first example, the movement patterns of two cars at five different traffic scenarios were evaluated. The movement patterns could be identified through the trajectories themselves. The first example was deliberately considered to be simple, to examine the result of the clustering. As predicted, the results of clustering were justifiable according to the trajectories of the cars. In the second example, a more complex movement was given. Unlike a traffic scenario, squash players were moving freely in the squash court. Consequently, the clustering of the trajectory pairs was not possible by looking only at the trajectories. We applied the approach that was evaluated in the first example to cluster the trajectory pairs of the squash players. Comparing the experimental results of the presented approach with other existing methods was not our first priority in this study. However, this issue will enable more persuasive and reliable results.

(c) Another important issue is whether incomplete knowledge (e.g., noise) can be handled with QTC. Of course, not always everything is known when making inferences about an issue at hand (Van Belleghem et al., 1994). Obviously, in these situations, the available information could be lacking with respect to offering complete answers to queries. However, "*a partial answer may be better than no answer at all.*" as Freksa (1992a) argues. The development of QTC has been inspired by some major qualitative reasoning calculi such as the temporal Semi-Interval Calculus (Freksa, 1992a) and the spatial Double-Cross Calculus (Freksa, 1992b; Zimmermann & Freksa, 1996). Principal in these theories is that specific attention has been paid to reasoning about incomplete knowledge. Hence, QTC can handle incomplete knowledge as well. (For more explanation, see (Van de Weghe et al., 2007).)

5.5 Conclusions and Future Work

Knowledge discovery from moving objects' trajectories is an important and challenging issue in many research domains. This chapter presented a new technique for identifying, representing, and analysing patterns of relative motion between disjoint MPOs, which is based on three major steps. In the first step, we described movement patterns of MPOs using qualitative trajectory calculus (QTC). QTC enables us to express the interactions between moving objects qualitatively. In the next step, movement patterns were represented in a sequence signature (SESI), which is a fractal way of mapping patterns of interactions between MPOs in an indexed raster space. Then, in the third step, a distance function was used to cluster SESIs and aims to improve the understanding of the movement patterns. We demonstrated the usefulness of our methodology to identify and cluster the movement patterns of cars and squash players.

Our experimental results showed that the proposed mining technique achieved good clustering quality. In addition, unlike many trajectory similarity measures that are restricted to geometric abstractions of the trajectories, we were able to measure the distance between the trajectory pairs based on the movement patterns. The proposed methodology could be used in a wide range of research applications. Movement patterns such as walking, running, jumping, lifting, striking, and swimming can be investigated for different purposes. For example, the proposed approach can be used in sports sciences to analyse the movement of athletes with the purpose of rehabilitation, physical education and practice. The emphasis of the therapy can be diverse, ranging from upper limb rehabilitation and balance rehabilitation to the rehabilitation of specific body parts (Schönauer et al., 2011). In many cases, full body interactions are captured using various types of capturing systems, such as Motion Capture (MoCap), to determine how well the body parts, such as hands and feet, move through space and time to regain normal function (Allard et al., 1998; Fernandez-Baena et al., 2012; Higginson, 2009; Schönauer et al., 2011). The proposed clustering approach can cluster pairs of interactions and assess whether body interactions are sufficiently improved relative to a ‘normal’ body interaction. In the case of swimming, for example, the investigation of movement patterns of lower/upper limbs of swimmers identifies common features of novice swimmers and how these features change with increasing skill during an instructional period. The proposed approach can be applied in dance analysis, where an examination of the movement patterns of the body parts of dancers is important to assist instructors for educational purposes.

Another possibility is to apply the presented approach to investigate the movement behaviour of animals. Because QTC considers the relationships between trajectory pairs, we can analyse the movement of pairs of animals to observe their movement patterns. For example, the interaction between two males can be different from the interaction between a male and a female, or the interaction of a baby and its mother can be different from the interaction of a baby and its father. These differences can be observed from the QTC movement patterns. A library of various identified QTC movement patterns (interactions) can enhance the understanding of unknown movement patterns of animals.

We could also exploit the proposed approach as a complementary analysis technique to support coaching in certain sports, such as football. For example, movement behaviour of certain players (e.g., strikers) in different temporal periods of the game could be examined to discover techniques or strategies that lead to frequent movement patterns on SESI. It would be an interesting step towards the visual analysis of the movement behaviour of players using SESIs when the team is attacking or defending.

Prior to applying the proposed approach in such case studies, we will need to extend our approach towards other types of QTC, such as QTC Double-Cross (QTC_C). These approaches provide more details about the movement of MPOs by including the direction of the movement of disjoint MPOs.

Another interesting issue is the ontological aspect of knowledge discovery from qualitative data. The ontological commitments for incomplete, uncertain, and erroneous data must be linked to the decision process to see how they affect the quality of the decisions (Frank, 2007). In this study, we did not address the ontological aspects of the presented approach. However, we intend to extend the on-going research in future work and assess the ontological commitments for incomplete data on the resulting clusters.

References

Ahlqvist, O., Ban, H., Cressie, N., & Shaw, N. Z. (2010). Statistical counterpoint: Knowledge discovery of choreographic information using spatio-temporal analysis and visualization. *Applied Geography, 30* (4), 548-560.

Allard, P., Cappozzo, A., & Vaughan, C. (1998). *Three-Dimensional Analysis of Human Locomotion*. Wiley.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Alon, J., Sclaroff, S., Kollios, G., & Pavlovic, V. (2003). Discovering clusters in motion time-series data. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Vol. 1, pp. 375-381).

Andrienko, N., & Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Berlin Heidelberg: Springer-Verlag.

Barnsley, M. F. (1988). *Fractals Everywhere*. New York: Academic Press.

Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature, 439* (7075), 462-465.

Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365* (1550), 2157-2162.

Chudova, D., Gaffney, S., Mjolsness, E., & Smyth, P. (2003). Translation-invariant mixture models for curve clustering. *Proceedings of the 9th International Conference on Knowledge Discovery and Data mining (ACM SIGKDD)* (pp. 79-88).

Clementini, E., Felice, P. D., & Hernández, D. (1997). Qualitative representation of positional information. *Artificial Intelligence, 95* (2), 317-356.

Cohn, A. G. (1996). Calculi for qualitative spatial reasoning. *Artificial Intelligence and Symbolic Mathematical Computation*, 124-143.

Cohn, A. G., & Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae, 46* (1-2), 1-29.

Cristani, M., Cohn, A. G., & Bennett, B. (2000). Spatial locations via morpho-mereology. *Proceedings of the Principles of Knowledge Representation and Reasoning (KR'2000)* (pp. 15-25). Colorado, USA.

Dawyndt, P., De Meyer, H., & De Baets, B. (2005). The complete linkage clustering algorithm revisited. *Soft Computing, 9* (5), 385-392.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Delafontaine, M., Van de Weghe, N., Bogaert, P., & De Maeyer, P. (2008). Qualitative relations between moving objects in a network changing its topological relations. *Information Sciences, 178* (8), 1997-2006.

Delafontaine, M., Versichele, M., Neutens, T., & Van de Weghe, N. (2012). Analysing spatiotemporal sequences in Bluetooth tracking data. *Applied Geography, 34*, 659-668.

Dodge, S., Weibel, R., & Lautenschutz, A. K. (2008). Towards a taxonomy of movement patterns. *Information Visualization, 7* (3-4), 240-252.

Egenhofer, M. J., & Mark, D. M. (1995). Naive geography. *Spatial Information Theory, 988*, 1-15.

Fernandez-Baena, A., Susin, A., & Lligadas, X. (2012). Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments. *Proceedings of the 4th International Conference on Intelligent Networking and Collaborative Systems (INCoS)* (pp. 656-661).

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence, 24* (1-3), 85-168.

Frank, A. U. (2007). Data quality ontology: An ontology for imperfect knowledge. In: S. Winter, M. Duckham, L. Kulik & B. Kuipers (Eds.), *Spatial Information Theory* (pp. 406-420). Berlin Heidelberg: Springer.

Freksa, C. (1992a). Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54* (1), 199-227.

Freksa, C. (1992b). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Gaffney, S., & Smyth, P. (1999). Trajectory clustering with mixtures of regression models. *Proceedings of the 5th International Conference on Knowledge Discovery and Data mining (ACM SIGKDD)* (pp. 63-72).

Galton, A. (2001). Dominance diagrams: A tool for qualitative reasoning about continuous systems. *Fundamenta Informaticae, 46* (1-2), 55-70.

Giannotti, F., Nanni, M., Pinelli, F., & Pedreschi, D. (2007). Trajectory pattern mining. *Proceedings of the 13th International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD)* (pp. 330-339).

Giannotti, G., & Pedreschi, D. (2008). *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*. Berlin Heidelberg: Springer-Verlag.

Gudmundsson, J., & van Kreveld, M. (2006). Computing longest duration flocks in trajectory data. *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems* (Vol. 10, pp. 35-42).

Gudmundsson, J., van Kreveld, M., & Speckmann, B. (2007). Efficient detection of patterns in 2D trajectories of moving points. *Geoinformatica, 11* (2), 195-215.

Haghani, A., Hamedi, M., Sadabadi, K. F., Young, S., & Tarnoff, P. (2009). Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board, 2160*, 60-68.

Higginson, B. K. (2009). Methods of running gait analysis. *Current Sports Medicine Reports, 8* (3), 136-141.

Hwang, S. Y., Liu, Y. H., Chiu, J. K., & Lim, E. P. (2005). Mining mobile group patterns: A trajectory-based approach. *Advances in Knowledge Discovery and Data Mining*, 145-146.

Kalnis, P., Mamoulis, N., & Bakiras, S. (2005). On discovering moving clusters in spatio-temporal data. *Advances in spatial and temporal databases*, 923-923.

Kang, J. H., Welbourne, W., Stewart, B., & Borriello, G. (2004). Extracting places from traces of locations. *Proceedings of the 2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots* (pp. 110-118).

Kisilevich, S., Mansmann, F., Nanni, M., & Rinzivillo, S. (2010). Spatio-temporal clustering. In: O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 855-874). Springer.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Monferrer, M. T. E., & Lobo, F. T. (2002). Qualitative velocity. In: M. T. E. Monferrer, F. T. Lobo & E. Golobardes (Eds.), *Topics in Artificial Intelligence* (pp. 29-39). Berlin Heidelberg: Springer.

Moratz, R., Dylla, F., & Frommberger, L. (2005). A relative orientation algebra with adjustable granularity. *Proceedings of the Workshop on Agent in Real-Time and Dynamic Environments (IJCAI 05).* Edinburgh, Scotland.

Nanni, M., & Pedreschi, D. (2006). Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems, 27* (3), 267-289.

Pelekis, N., Kopanakis, I., Marketos, G., Ntoutsi, I., Andrienko, G., & Theodoridis, Y. (2007). Similarity search in trajectory databases. *Proceedings of the 14th International Symposium on Temporal Representation and Reasoning (TIME 2007)* (pp. 129-140).

Pers, J., Bon, M., & Vuckovic, G. (2006). CVBASE '06 Dataset. *Available Online:* http://vision.fe.uni-lj.si/cvbase06/downloads.html.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In B. Nebel, W. Swartout & C. Rich (Eds.), *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR)* (Vol. 92, pp. 165-176).

Renz, J., Rauh, R., & Knauff, M. (2000). Towards cognitive adequacy of topological spatial relations. *Spatial Cognition II, 1849*, 184-197.

Rokach, L., & Maimon, O. (2010). A survey of clustering algorithms. In: L. Rokach & O. Maimon (Eds.), *Data Mining and Knowledge Discovery Handbook*, (pp. 269-298).

Schönauer, C., Pintaric, T., & Kaufmann, H. (2011). Full body interaction for serious games in motor rehabilitation. *Proceedings of the 2nd Augmented Human International Conference* (pp. 4-11).

Shoval, N., & Isaacson, M. (2007). Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers, 97* (2), 282-297.

Tiru, M., Saluveer, E., Ahas, R., & Aasa, A. (2010). Web-based monitoring tool for assessing space-time mobility of tourists using mobile positioning data: Positium barometer. *Journal of Urban Technology, 17* (1), 71-89.

Van Belleghem, K., Denecker, M., & De Schreye, D. (1994). Representing continuous change in the abductive event calculus. In: P. Van Hentenryck (Ed.), *Proceeding of the International Conference on Logic Programming (ICLP)* (pp. 225-240).

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Van de Weghe, N., Bogaert, P., Cohn, A., Delafontaine, M., De Temmerman, L., Neutens, T., De Maeyer, P., & Witlox, F. (2007). How to handle incomplete knowledge concerning moving objects. In: B. Gottfried (Ed.), *Proceedings of the Workshop on Behaviour Monitoring and Interpretation (BMI 2007)* (pp. 91-101).

Van de Weghe, N., Cohn, A. G., Maeyer, P. D., & Witlox, F. (2005a). Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems with Applications, 29* (4), 977-983.

Van de Weghe, N., Cohn, A. G., De Tre, G., & De Maeyer, P. (2006). A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems. *Control and Cybernetics, 35* (1), 97-119.

Vilain, M., Kautz, H., & Van Beek, P. (1989). Constraint propagation algorithms for temporal reasoning: A revised report. In: D. S. Weld & J. D. Kleer (Eds.), *Readings in Qualitative Reasoning about Physical Systems* (pp. 373-381). San Francisco: Morgan Kaufmann.

Wang, C., De, D., & Song, W. Z. (2013). Trajectory mining from anonymous binary motion sensors in Smart Environment. *Knowledge-Based Systems. 37 (2013): 346-356.*

Weld, D. S., & Kleer, J. D. (1989). *Readings in Qualitative Reasoning about Physical Systems*. San Francisco: Morgan Kaufmann.

Wilson, W. C. (1998). Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A, 30* (6), 1017-1038.

Wisbey, B., Montgomery, P. G., Pyne, D. B., & Rattray, B. (2010). Quantifying movement demands of AFL football using GPS tracking. *Journal of Science and Medicine in Sport, 13* (5), 531-536.

Zimmermann, K., & Freksa, C. (1996). Qualitative spatial reasoning using orientation, distance, and path knowledge. *Applied Intelligence, 6* (1), 49-58.

6

Multi-Dimensional Patterns

Knowledge is love and light and vision. Helen Keller

Modified From: Seyed Hossein Chavoshi, Bernard De Baets, Tijs Neutens, Matthias Delafontaine, Guy De Tré, Nico Van de Weghe (2014). Movement Pattern Analysis Based on Sequence Signatures. Under Review.

6 Multi-Dimensional Patterns

Abstract: Increased affordability and deployment of advanced tracking technologies have led researchers from various domains to address the resulting spatio-temporal sets of movement data for the purpose of knowledge discovery that can help better understand the behaviour of moving objects. Two different approaches can be considered in the analysis of moving objects: quantitative analysis and qualitative analysis. This research focuses on the latter and uses the qualitative trajectory calculus (QTC), a type of calculus that represents qualitative data on moving point objects (MPOs) and establishes a framework to analyse the movement behaviour of multiple MPOs. This chapter uses a visualisation technique called sequence signature (SESI) which is a way of mapping QTC movement patterns in a 2D indexed rasterised space, to measure the similarity of movement patterns of multiple MPOs. The concept of SESI is fully introduced in Chapter 5. In this chapter, we extend this concept to extract knowledge from movement of multiple moving objects. The applicability of the proposed methodology is illustrated by means of a practical example of comparing Samba dance movements during different time intervals. The results show that the proposed method can be effectively used to analyse interactions of multiple MPOs in different domains.

6.1 Introduction

Due to technological advances in positioning and tracking systems such as GPS, mobile positioning with Bluetooth and Wi-Fi, and video tracking, enormous amounts of movement data have been acquired in various domains. As a result, much attention has recently been paid to the analysis of movement data in many research areas. A large number of studies have been conducted with respect to the analysis of trajectory data, the mining of movement patterns, and exploratory visual analytics (see, for example (Andrienko & Andrienko, 2007; Andrienko & Andrienko, 2012; Bak et al., 2012; Dodge, 2011; Giannotti & Pedreschi, 2008; Imfeld, 2000; Laube et al., 2005; Mountain, 2005)).

Despite these broad efforts, only scant progress has been made in the field of qualitative reasoning about moving objects. Qualitative formalisms that are suited to express qualitative spatial or temporal relationships between entities have gained wide acceptance as a useful way of abstracting the real world and, consequently, reducing the complexity of reasoning about moving objects. In this respect, qualitative trajectory calculus (QTC), which was introduced by Van de Weghe (2004), represents a powerful calculus for assessing the interaction between disjoint moving point objects (MPOs) qualitatively.

Based on concepts from geographic knowledge discovery (GKD) (Giannotti & Pedreschi, 2008) for extracting meaningful information, discovering interesting patterns, and interpreting them in a plausible way, we propose a visualisation technique for analysing the movements of multiple disjoint MPOs using QTC. In this chapter, we will focus on the usefulness of QTC in identifying movement patterns of MPO pairs. An innovative visualisation technique called sequence signature (SESI) is used to transform the QTC movement patterns of MPOs into a structure that is suitable for being analysed through traditional data mining techniques such as clustering algorithms. Using a similarity measure between two SESIs, we can quantify the similarity of the movements of the MPOs. The presented methodology reduces the complexity that is associated with the analysis of MPO movements.

The remainder of this chapter is organised as follows. In the next section, we briefly discuss related work, while in Section 6.3, we introduce the key concepts of QTC and SESI. Section 6.4 discusses the proposed methodology for analysing the movement behaviour of multiple MPOs. In this section, data from an experiment (Samba dance) is presented to exemplify and validate the proposed method. In Section 6.5, we discuss our findings. Finally, Section 6.6 draws some conclusions and directions for future research.

6.2 Related Work

Movement data sets are growing rapidly due to the wholesale collection of spatio-temporal data on various phenomena using time-efficient and accurate positioning technologies such as GPS. As the most fundamental perception of movement, much effort has been devoted to explore the trajectories of moving objects, which enables finding behavioural patterns that can be used in different applications (Dodge, 2011). Despite the abundance of work that is related to the analysis of moving object traces, there are still many questions to answer, such as what type of movement patterns is one looking for? And which approaches and algorithms should be accounted for to extract the movement patterns? Because the focus of this chapter is developing a method to support our understanding of movement patterns of moving objects, three main issues are covered in this section: (i) relationships among MPOs (ii) visualisation of MPOs, and (iii) clustering movements.

6.2.1 Relationships among MPOs

Generally, movement patterns can be derived either from the movement of an individual object over time or from an interaction between two or more moving objects during a time interval of movement. A comprehensive classification of patterns in movement data, which is applicable for all of the common types of moving objects in different domains, such as humans, animals, cars, and eye movements, has been studied by Dodge (Dodge, 2011; Dodge et al., 2008). It is obvious that movement patterns of an isolated individual can differ from movement patterns of an individual who is part of a group. These differences between individual and group movement behaviour have been discussed extensively (Andrienko & Andrienko, 2007). Our research seeks to contribute to the exploration of interactions of multiple MPOs. The key idea is to examine movement patterns of multiple objects over space and time. One relevant concept in this respect is the *relative motion matrix* (REMO) (Laube et al., 2005), which was introduced by Laube, Imfeld and Weibel, in which users can search in a large movement data set for instances of pre-defined movement patterns that were constructed based on the existing knowledge about the movement of the objects under study. In REMO, the movement patterns are described by changes in the motion attributes of objects (i.e. the change in the speed or motion azimuth over space and time), and therefore, relate one object's movement to that of others. The REMO representation transforms the trajectories of MPOs into an analysis matrix that allows for the matching of movement patterns. The main difference between the current work and REMO is that, at the very basic level, we are investigating the interaction between pairs of MPOs instead of solely looking at the movement of individuals over time, and then, we infer the collective behaviour of multiple objects. Furthermore, this chapter does not investigate the changes in the motion attributes of MPOs. Instead, we examine how the relative changes in the Euclidean distances between MPOs can reveal movement patterns. For this purpose, we apply a qualitative approach called QTC (Van de Weghe, 2004) to express the relative change in the Euclidean distance among MPOs in a qualitative manner. Basically, three notions have been introduced in QTC between two MPOs to express the relationships between objects, namely *moving towards*, *moving away from*, and *stable,* which will be discussed in detail in Subsection 6.3.1.

6.2.2 Visualisation of MPOs

Visual representation is used as an effective technique to represent and support the analysis of movement patterns of objects. A substantial amount of research has been conducted to express the significance of visualisation in understanding movements in different domains (Andrienko & Andrienko, 2010; Andrienko & Andrienko, 2007; Andrienko et al., 2008; kraak & Van De Vlag, 2007; Ooms et al., 2012; Pelekis et al., 2012; Rinzivillo et al., 2008; Tominski et al., 2012; Xia & Kraak, 2010). For example, an effort has been made by Andrienko, Andrienko, Wachowicz, and Orellanal (Andrienko et al., 2008) to uncover interactions between moving objects by combining visual and filtering techniques. To achieve that goal, they have focused on two main issues: (i) the development of a theoretical foundation that includes a formal definition of an interaction and its indications, such as the spatial proximity between two or more objects at some moment in time or during a time interval, and (ii) a methodology for visually exploring and analysing possible interactions in large sets of movement data. Detecting, understanding, and visualising movement patterns are not limited to certain applications. For example, in Chapter 3, a technique for identifying, visualising and interpreting repetitive movement patterns within groups of moving point objects based on QTC information has been studied. As a case study, movement data of Samba dancers has been examined to evaluate the performance of Samba dancers in a visualisation technique, i.e. the continuous triangular model (CTM). From the results, repetitive movement patterns have been visually detected and interpreted. In fact, when movement patterns of MPOs are visualised, more tangible information might be extracted. This work is an extension of our previous effort (Chapter 5) of analysing the movement patterns of MPOs. Formerly, the relative motions and associated movement patterns between two disjoint MPOs were analysed using a visualisation technique called SESI. The proposed approach was only able to analyse the movement patterns of a single pair of MPOs, while in this chapter, this weakness is resolved by extending the approach to multiple MPOs, as explained in detail in Section 6.4.

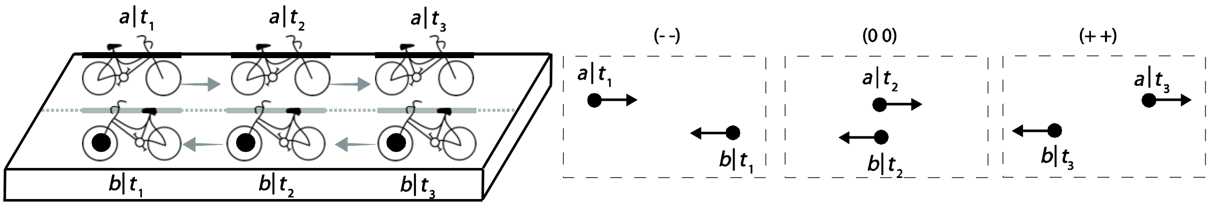
6.2.3 Clustering of Movements

Over the past decade, many studies have addressed knowledge discovery and data mining issues that are related to moving object data. Among them, some contributed to the clustering of moving objects (Buzan et al., 2004; Jensen et al., 2007; Li et al., 2004; Nanni & Pedreschi, 2006; Rinzivillo et al., 2008; Zhang & Lin, 2004), the mining of movement patterns (Demsar & Virrantaus, 2010; Dodge et al., 2008; Gudmundsson et al., 2007; Laube et al., 2011; Laube et al., 2008; Laube et al., 2005; Wilson, 2008) and exploring the similarity of moving objects (Buchin et al., 2009; Ding et al., 2008; Dodge et al., 2009; Lin & Su, 2008; Pelekis et al., 2007; Rinzivillo et al., 2008). The idea behind the approach proposed in this chapter is to contribute to all three classes of knowledge discovery, i.e. mining patterns, similarity assessment, and clustering. In addition to detecting and visualising movement patterns, we measure the similarity in the movement behaviour of moving objects. Knowledge about similarities in movement data can be valuable in the prediction, modelling and simulation of the collective behaviour of various dynamic phenomena (Dodge, 2011). In most of the cases, researchers have found similarities in movement patterns solely from the trajectories of moving objects. For example, in a prior study (Dodge, 2011), variations in the movement parameters, such as the speed, acceleration, or direction of the objects over time (which are obtained from the trajectories of moving objects) have been used to assess the similarity. Our work intends to explore similarities of movement patterns using a visualisation technique called SESI. From multiple SESIs, in addition to obtaining a visual synopsis of the movement patterns of MPOs, we can measure the degree of similarity in the movement patterns. The measurement leads towards a clustering of trajectory pairs, being the identification of groups of similar trajectory pairs.

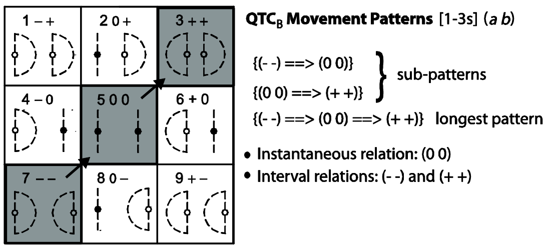
6.3 Background

6.3.1 Qualitative Trajectory Calculus

The movement of objects can sometimes be described more satisfactorily by using a qualitative description rather than a quantitative description (Delafontaine et al., 2011a). This possibility arises because qualitative measures better support the intuition of human beings compared with quantitative measures (Freksa, 1992). For example, stating that people are moving faster when they ride a bicycle than when they walk is easier to understand than stating that the average speed of riding a bicycle is 17-19 km/h and that of walking is 4-6 km/h. Until now, various qualitative calculi have been developed to reason about space (Cohn et al., 1997; Egenhofer & Franzosa, 1991) and time (Allen, 1983), but few of them have systematically examined the qualitative properties of spatio-temporal information, such as trajectory data of moving objects. This subsection briefly presents a general overview of the principal theoretical aspects of QTC, which is an integrated spatio-temporal calculus that represents and reasons about moving objects (A detailed explanation can be found in Chapter 5). We show how QTC_B, the basic level of QTC, is employed to represent raw moving object data in order to provide the basis for our aim. We will briefly introduce some basic concepts of QTC_B with an example. An interaction between two moving objects, say two bicycles, *a* and *b* (Figure 6-1a), during a time interval of movement can be described by QTC_B relations. QTC_B relations are constructed based on changes in the Euclidean distance between two MPOs over time.



(a)



(b) (c)

Figure 6-1: Interaction between two moving objects (*a*, *b*) during a time interval of movement (a), nine possible QTC_B relation icons (b), and all QTC_B movement patterns between two bicycles *a* and *b* (c).

The QTC_B relations mainly imply the *towards* and *away from* relations. They are built up from the distance constraints (A and B) introduced in Section 2.5. As noted earlier in Chapter 5, each QTC_B relation is represented by a two-tuple $(A\ B)_B$[1], where each tuple implies the distance constraints between both MPOs. There are 9 (3^2) base relations in QTC_B, which are shown in Figure 6-1b; these relations are all possible in a one or higher-dimensional space, i.e. $(-\ +)$, $(0\ +)$, $(+\ +)$, $(-\ 0)$, $(0\ 0)$, $(+\ 0)$, $(-\ -)$, $(0\ -)$, and $(+\ -)$. For example, the QTC_B relation $(+\ +)$ expresses that both objects are moving away from each other (for more explanation about QTC_B, see (Van de Weghe, 2004)). Each time stamp/interval of movement can be represented by a QTC_B relation.

[1] In $(A\ B)_B$ relation syntax, A refers to movement of the first object with respect to the second object at *t* and B refers to movement of the second object with respect to the first object at *t*.

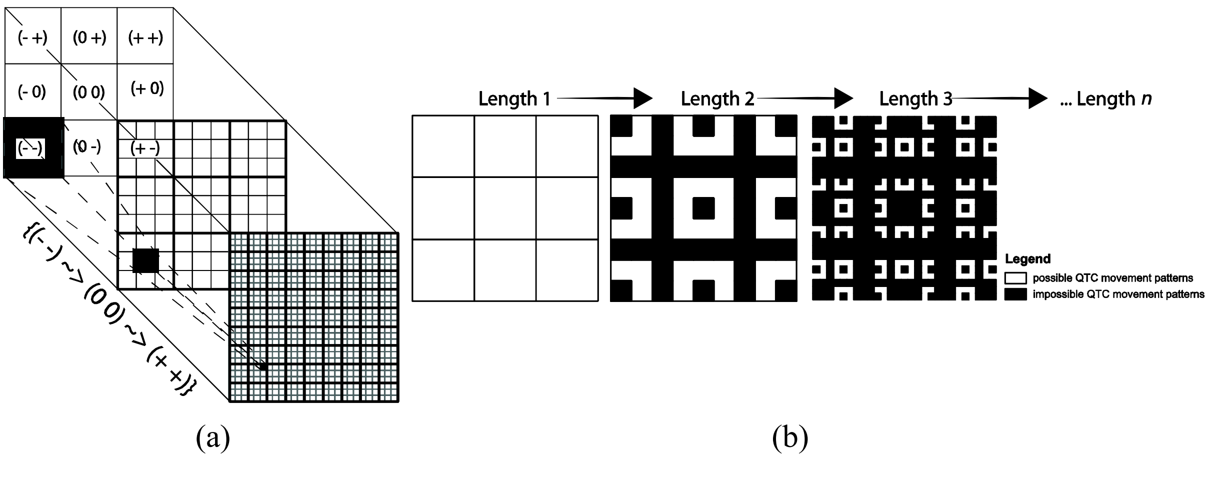
In fact, depending on the temporal granularity by which movement data has been captured, details of the information and, consequently, the accuracy of the results will vary. A QTC_B relation represents the interaction between two MPOs at a time stamp of movement and, therefore, during a time interval of interaction between two MPOs, a sequence of successive QTC_B relations can be obtained in which some of the QTC_B relations are instantaneous relations (e.g., (0 0) is an instantaneous relation in Figure 6-1). Hereafter, the term QTC_B movement patterns will be replaced by sequences of successive QTC_B relations. The longest QTC_B movement pattern shows the whole interaction between two MPOs in terms of the QTC_B relations during a time interval of movement, while the sub-QTC_B movement patterns constructed from the longest movement pattern represent the sectional patterns (Figure 6-1c). In this instance, given that the temporal granularity of capturing the movement is one second, three QTC_B relations, $(-\,-)$, $(0\ 0)$, and $(+\,+)$, were observed. Accordingly, the QTC_B movement pattern $\{(-\,-) \rightarrow (0\ 0) \rightarrow (+\,+)\}$ represents the whole interaction between two bicycles *a* and *b* during a time interval of 2 s. Sub-QTC_B movement patterns disclose the details of the interactions during a time interval. In a previous example, $\{(-\,-) \rightarrow (0\ 0)\}$ and $\{(0\ 0) \rightarrow (+\,+)\}$ are two sub-QTC_B movement patterns, each of which are made of two QTC_B relations. It is noted that, following Galton's theory of dominance (2001) and Forbus' equality change law (1984), a direct change from $-$ to $+$ and vice versa is impossible because such a change must pass the qualitative value 0. Therefore, there are some impossible direct transitions between the QTC_B relations. For example, there must be another QTC_B relation between the relations $(-\ +)$ and $(+\ +)$, which is $(0\ +)$.

The example shown in Figure 6-1 was simply designed to explain the QTC_B relations and QTC_B movement patterns. However, the interaction between the MPOs could be much more complex in reality. In this chapter, we will specifically examine the movement of the pairs of MPOs by using the properties of the QTC_B movement patterns, i.e. the frequency and duration. Depending on the complexity of the movement between interacting MPOs, the frequency and duration of the sub-QTC_B movement patterns will vary. To investigate the movement of the pairs of MPOs during different time intervals of movement, it will be sufficient to compare their sub-QTC_B movement patterns and, respectively, their frequency and duration. A novel visual technique called SESI is used to map QTC_B movement patterns in a 2D indexed rasterised space. SESIs provide a framework for comparing the interaction of pairs of MPOs. The next subsection introduces the principal concept of SESI.

6.3.2 Sequence Signature

A Sequence Signature (in short SESI) is a fractal approach to mapping QTC_B movement patterns that result from interactions between MPOs in an indexed raster space (see also Chapter 5). It is based on the concept of an iterated function system (IFS), which is a method of constructing fractals, as discussed in detail previously (Barnsley, 2000). In brief, a fractal is a geometric object that is similar to itself at all scales, and the overall pattern is repeated at any scale. Iterated function systems are methods for constructing fractals. An IFS fractal exhibits self-similarity, which means that its structure is constructed from the repetition of copies of itself, each copy being transformed by a function. The functions are typically contractive, which results in smaller shapes while bringing points closer together.

Each cell in a SESI represents a specific QTC_B movement pattern. The resolution of the cells in a SESI depends on the length of the QTC_B movement patterns. A SESI with length 1 represents all of the basic QTC_B relations, while, for example, a SESI with length 2 illustrates all of the QTC_B movement patterns with two consecutive QTC_B relations. Figure 6-2a illustrates the longest QTC_B movement pattern from the previous example in a SESI with length 3. A compact SESI is a superimposition of all of the SESIs of any length during a given time interval. This concept will be explained later in this chapter.



(a) (b)

Figure 6-2: SESI of length 1, 2, and 3 (the longest movement pattern in Figure 6-1 is indicated in the SESI of length 3) (a), continuity and event-based constraints imposed on SESIs with length 1, 2, and 3 (b).

As mentioned before, not all QTC_B transitions and, consequently, not all QTC_B movement patterns are possible or significant. First, we exclude chronologically impossible combinations of QTC_B relationships in SESIs of length 2 or more. Second, because we are typically interested in the changes in the relative motion between MPOs over time (i.e. events), some sequences of qualitative relationships are insignificant. This outcome occurs when a QTC relationship is invariant over time. For example, a transition from $(-\ +)$ into itself is not very meaningful from a qualitative perspective. Figure 6-2b demonstrates SESIs of length 1, 2, and 3 after imposing both constraints. The black cells indicate discontinuous (and thus impossible) and non-significant sequences of qualitative relationships between two MPOs (for more explanation, see (Chapter 5)). Here, each cell in a SESI has two values, which are the frequency and duration of a unique QTC_B movement pattern located in that cell. The frequency and duration of the cells are considered in the similarity measure, which will be discussed in detail in the methodology section.

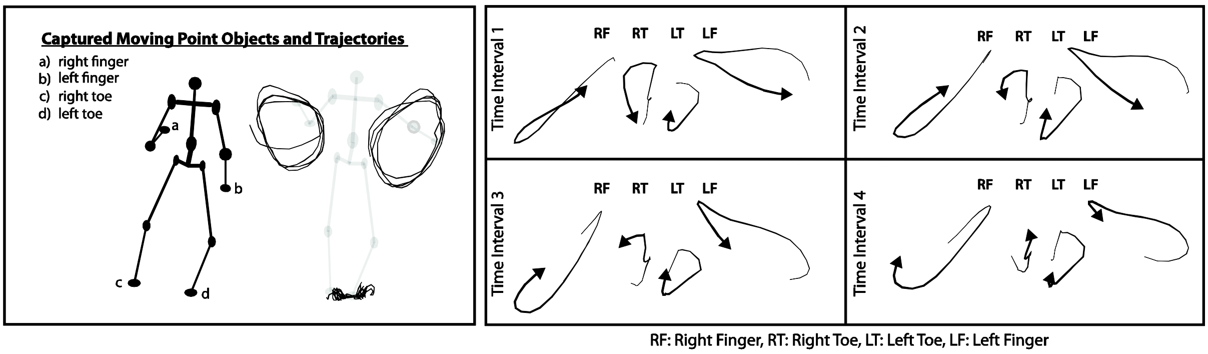
6.4 Methodology

Our methodology for analysing the movement of multiple MPOs is composed of four steps: (i) data preparation; (ii) QTC_B movement pattern extraction; (iii) building SESIs and hyper-SESIs for individual and multiple pairs of MPOs, respectively; and (iv) building a similarity measure for hyper-SESIs. Each of the steps is explained in detail in the following subsections.

6.4.1 Step 1: Data Preparation

A major step in analysing movement is the data preparation. The first issue that must be considered is the spatio-temporal granularity of the movement data, which in turn depends on the resolution of the capturing. Some applications require very high resolution data, such as dance analysis, while others, such as traffic flow analysis, can be derived from more roughly captured data. Selecting and assigning optimal spatio-temporal resolution for capturing data are out of the scope of this chapter and will not be considered. In the remainder of this subsection, an experiment is designed to examine and validate the applicability of the proposed methodology. However, the proposed methodology can be applied to a wide range of applications in which exploring and understanding movement patterns is significant, including traffic management, in which building an ontology based on the movement patterns of vehicles can inform flow modelling; the analysis of human body movements, where identifying and examining certain patterns in the interactions of different body parts of disabled people can help them to recover; and sports analysis, such as squash and football, where understanding the movement strategies of the players can be used for coaching.

For illustrative purposes, this chapter will use a data set of movements of a Samba dancer during four different time intervals, each lasting 0.8 s. The time intervals have been deliberately considered short to simply recognise the similarity between trajectories of movements. The movements of four parts of the body of the dancer, including the right finger (the right hand), the left finger (the left hand), the right toe (the right foot), and the left toe (the left foot), are captured at each time stamp of the movement with a temporal granularity of 0.04 s (Figure 6-3a). The main reason for considering only four body parts to investigate the dance movement is their higher number of interactions compared with other parts of the body of the dancer.



(a) (b)

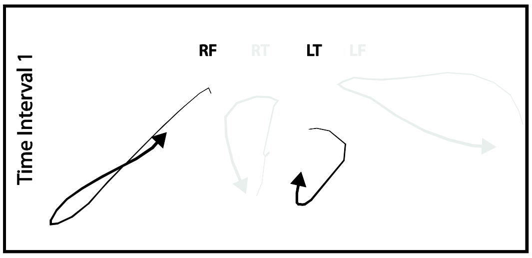
Figure 6-3: An abstracted movement of a Samba dancer based on four parts of the body (right finger, left finger, right toe, and left toe) (a), and top view of movements of the Samba dancer during four different time intervals, each lasting 0.8 s (b).

The positional information has been captured by an infrared motion capture system that yields the position of the markers attached to the body in three-dimensional space (Figure 6-3b). The data have been normalised with respect to one reference point and the orientation of the dancer's body (that point is defined as the centroid of the body, called the root).

In this chapter, we investigate only the movements of the listed parts of the body because their movement is more noticeable compared to that of other parts. We intend to identify, visualise, and interpret the existing movement patterns obtained from the interactions between the different body parts of the dancer. Relationships between different parts of the body are described by QTC_B relations based on the positional information described in the next subsection.

6.4.2 Step 2: QTC_B Movement Pattern Extraction

In the second step, *QTCAnalyst* is used to extract QTC_B movement patterns of MPO pairs by giving the trajectories. *QTCAnalyst* was developed as a prototype QTC-based information system in Visual Basic 6.5 using AutoCAD to automatically generate and export QTC representations that model relations among moving objects (Delafontaine et al., 2011b). Trajectories of MPOs can be loaded in *QTCAnalyst* through a GUI. There are possibilities for visualising the trajectories in a conventional two-dimensional space to see how the MPOs are interacting with each other. Here, we use *QTCAnalyst* to calculate and export the QTC information, such as the QTC_B movement patterns and the sub-QTC_B movement patterns, i.e. the chains of subsequent QTC_B relations. The complexity of interactions between two MPOs during a given time interval gives rise to different QTC_B movement patterns. Figure 6-4 illustrates trajectories of the right finger and the left toe of the Samba dancer during Time Interval 1 followed by all of the QTC_B movement patterns at different lengths obtained from *QTCAnalyst*.



RF: Right Finger, RT: Right Toe, LT: Left Toe, LF: Left Finger

QTC_B Movement Patterns	Frequency	Duration
[+- ==> 0-]	1	10
[0- ==> --]	1	2
[-- ==> -0]	1	2
[-0 ==> -+]	1	7
[+- ==> 0- ==> --]	1	12
[0- ==> -- ==> -0]	1	2
[-- ==> -0 ==> -+]	1	9
[+- ==> 0- ==> -- ==> -0]	1	12
[0- ==> -- ==> -0 ==> -+]	1	9
[+- ==> 0- ==> -- ==> -0 ==> -+]	1	19

Figure 6-4: Summary of the QTC_B movement patterns, with their frequencies and durations, for the interaction between the right finger and left toe of the Samba dancer.

From the QTC_B movement patterns obtained from the interactions between the right finger and left toe of the dancer, we can observe that there is no regularity in the movement because the occurrences of all of the patterns are equal to one. In other words, there is no repetitive motion between these two parts of the body during the performance. Unlike the frequency, the durations of the QTC_B movement patterns are different, and some of the movement patterns have lasted longer while others are shorter. As mentioned earlier, the temporal granularity of capturing dance movement is 0.04 s. For the sake of simplicity, we assume every 0.04 s to be a single time unit. For example, pattern $\{(+-) \rightarrow (0-)\}$ has lasted 10 time units, which is equivalent to $10 * 0.04 = 0.4\ s$. The variety in the movement patterns that was achieved from the interactions between MPOs implies that there is complexity in the relative motion, which is our concern in this study.

6.4.3 Step 3: Building SESIs and Hyper-SESIs for Individual and Multiple Pairs of MPOs, Respectively

In the third step, a SESI is used as an iterative mapping technique to represent QTC_B movement patterns obtained from interactions of a pair of MPOs. As the main contribution of the chapter, we propose SESI as a structure to compare the QTC_B movement patterns of pairs of MPOs and hence attain the degree of similarity between them. Basically, acquiring similarities as part of knowledge discovery on movement data is beneficial to assessing the degree of closeness of the movement behaviour of different pairs of MPOs. In our previous effort, the basic concept of SESI has been introduced. We will now extend the use of the SESI concept for documenting and representing the interaction of multiple pairs of MPOs, accounting for an additional property of movement patterns (i.e. duration) in the calculation of similarity. As explained earlier, SESIs are representing the QTC_B movement patterns of the interaction of pairs of MPOs in indexed raster spaces. In addition, it is an effective way to provide a representation of the distribution of QTC_B movement patterns. A SESI can have different lengths. A SESI of length 1 shows the nine base QTC_B relationships, which are exclusively mapped in nine cells (Figure 6-2a). We do not represent length 1 of SESIs because they do not reveal interesting information about the movement patterns of objects. For higher lengths, each cell is further subdivided into nine cells, to make each cell in a SESI of length n correspond to a unique sequence of n qualitative QTC_B relationships. Figure 6-5 illustrates a SESI that is obtained from the interaction between the right toe and left finger. All QTC_B movement patterns of the interaction have been represented in a compact SESI, which is a superimposition of all SESIs of any length (i.e. five lengths). To enhance the visibility of the transformed QTC_B movement patterns, impossible and insignificant QTC_B movement patterns (which correspond to the black cells in Figure 6-2b) are not represented.

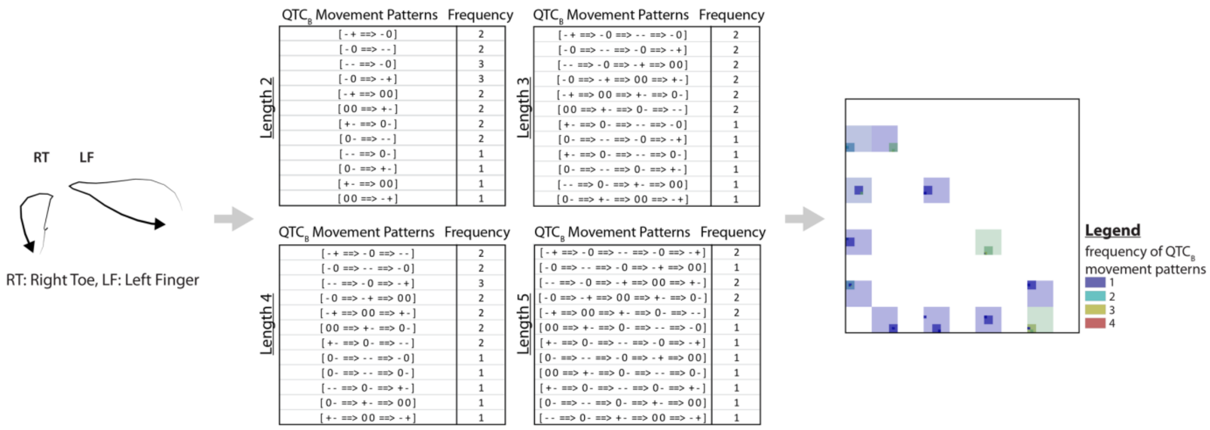


Figure 6-5: A SESI that represents all of the QTC_B movement patterns of interactions between the right toe and left finger of the Samba dancer.

The values of the cells in the SESIs are either the frequency of QTC_B movement patterns or the duration of them. Therefore, two types of SESI are defined, namely frequency-based SESI and duration-based SESI. Because any QTC_B movement pattern with any length is located in an exclusive cell, it is obvious which QTC_B movement patterns are more frequent than others or which ones last longer. More changes in the interactions between two MPOs results in a SESI with a greater variety of QTC_B movement patterns spread all over the cells. Figure 6-6 demonstrates that the right toe and left toe of the Samba dancer have more interactions with each other during the first time interval of the movement than the second time interval. In this figure, from the frequency-based SESIs of both time intervals, we observe that the QTC_B movement patterns $\{(+\,+) \rightarrow (+\,0)\}$, $\{(+\,0) \rightarrow (+\,+)\}$, $\{(+\,0) \rightarrow (+\,-)\}$, and $\{(+\,-) \rightarrow (+\,0)\}$ are more frequent than others of length 2. The QTC_B movement patterns $\{(+\,+) \rightarrow (+\,0) \rightarrow (+\,-)\}$ and $\{(+\,-) \rightarrow (+\,0) \rightarrow (+\,+)\}$ are the most frequent movement patterns of length 3, and so forth. Considering the duration-based SESI of the second time interval, the QTC_B movement pattern $\{(-\,+) \rightarrow (0\,0) \rightarrow (+\,-)\}$ is one of the long duration patterns of length 3. This arrangement means that the right toe was moving towards the left toe, while at a time in between, the movement behaviour changed, with the left toe moving towards the right toe, while the right toe was moving away from the left toe.

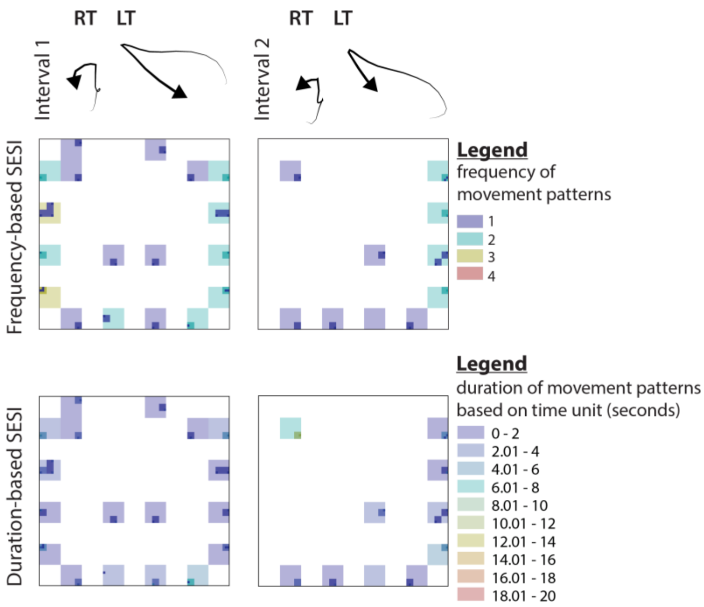


Figure 6-6: Frequency-based SESI and duration-based SESI of the right toe and left toe of the Samba dancer during two equal time intervals of movement.

Because each SESI represents the interaction between two MPOs, the interaction of multiple MPOs can be visualised simultaneously on a newly formed SESI called hyper-SESI, which is tessellated into all sub-SESIs. Based on the successive QTC_B relations of all possible pairs of MPOs during each time interval, we create hyper-SESIs for four equal time intervals based on the frequency and duration of the QTC_B movement patterns (Figures 6-7 and 6-8). Because in this chapter only the movement of four parts of the body of a dancer has been studied, there are six possible pairs of interactions. The pairs of movement interactions between different body parts are shown in terms of hyper-SESIs for each time interval of movement in Figures 6-7 and 6-8. In each hyper-SESI, the leftmost cell expresses the interaction between the right finger and right toe of the dancer, and the lowest cell represents the interaction, or in fact QTC_B movement pattern, between the left toe and left finger. Note that only the upper parts of the main diagonal of the hyper-SESIs must be considered because the interactions of the moving parts are symmetrical. We attempt to discover whether there is a degree of closeness/similarity between the movement patterns of different pairs of body parts during different time intervals of movement. Based on the context of the movement data, we can interpret the hyper-SESIs. For example, we know that Samba dance is a rhythmical dance that has regularity in the movements. The regularities in the movements can be distinguished on the hyper-SESIs cells. In Figure 6-7, the upper left cell of hyper-SESI shows the interactions between the right finger and right toe at four different time intervals. One can observe that there are more repetitive QTC_B movement patterns during the time interval 4 because more frequent patterns are visualised on its hyper-SESI. This finding probably occurs because, during that time interval, dancers should perform some more repetitive movements. As another example, the SESI of right toe-left toe at the first time interval of movement are slightly different from the SESIs at other time intervals.

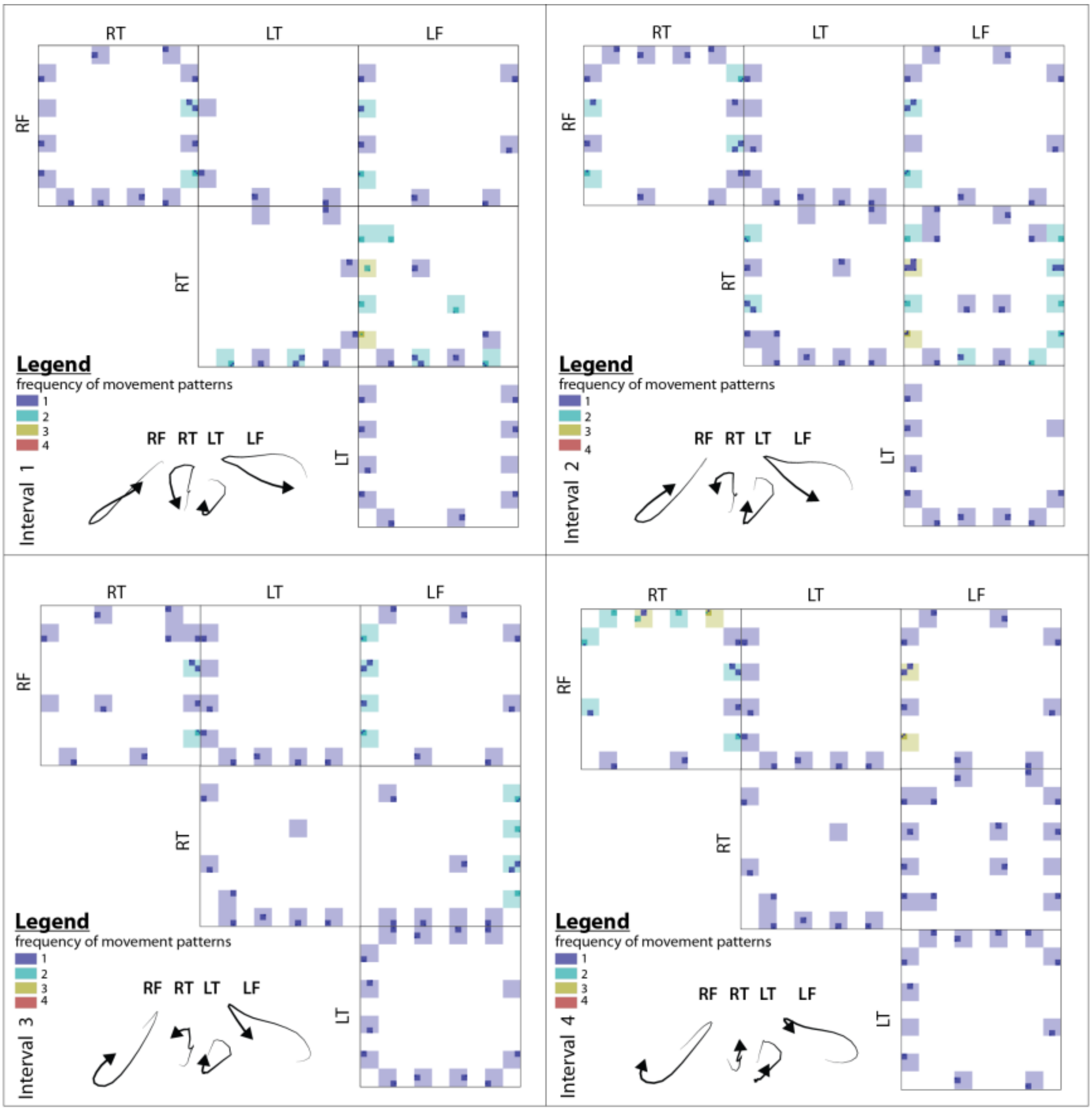


Figure 6-7: Hyper-SESIs of four equal time intervals of movements based on the frequency of QTC_B movement patterns.

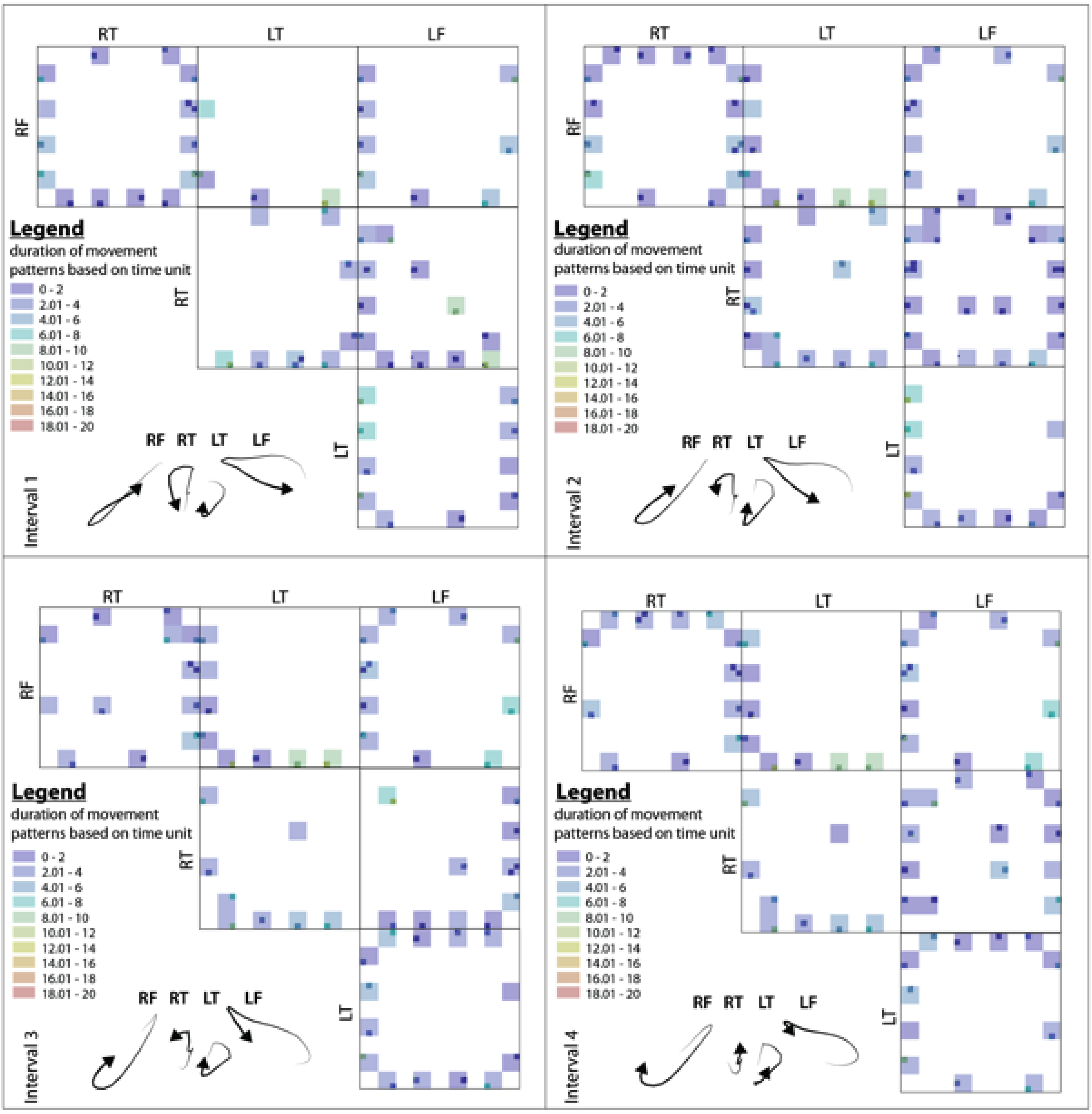


Figure 6-8: Hyper-SESIs of four equal time intervals of movements based on the duration of QTC_B movement patterns.

QTC_B movement patterns $\{(+\,0) \rightarrow (+\,+)\}$ and $\{(+\,-) \rightarrow (+\,0)\}$ of length 2 and $\{(+\,0) \rightarrow (+\,+)) \rightarrow (0\,+)\}$ and $\{(+\,-) \rightarrow (+\,0) \rightarrow (+\,+)\}$ of length 3 are observed in the first time interval of movement but not in the other intervals. From the perspective of visual assessment, many differences are found when we look more closely at the hyper-SESIs, both for frequency and duration.

It makes more sense when we expect certain patterns with a specific frequency and duration. For example, in our case study, a dance tutor can ask the dance amateur to perform some movements. The requested movement patterns should be observed in the corresponding SESIs unless the dance amateur has failed to perform successfully.

In addition to a visual judgement of SESIs or hyper-SESIs, the quality of the dance performances can be measured. One might measure the similarity of the performances of a dancer (taken as a benchmark) with the performances of other dancers and investigate the quality of the performances. For this purpose, we measure the degree of closeness/similarity between corresponding hyper-SESIs. In fact, we can investigate not only the movements of each individual pair, such as the left finger-right finger at different time intervals of movement, but also the overall measurement of whole body part interactions. Corresponding SESIs are compared to drawing an analogy between individual pairs; however, the average of all similarity measures of individual SESIs creates the overall measurement. Extracting such similarities can meaningfully contribute to the analysis of the movement behaviour of MPOs. The similarity measure is discussed in detail in the next subsection.

6.4.4 Step 4: Building a Similarity Measure for Hyper-SESIs

Measuring the similarity between hyper-SESIs assists in discovering the degree of closeness in the collective movement behaviour of pairs of MPOs. Usually, selecting an appropriate similarity function depends on the data type and the context of the problem. We use a cosine-based similarity function to express the similarity between two hyper-SESIs. The proposed similarity function is calculated based on the values of the cells in the SESIs (i.e. the frequency and duration of superimposed QTC_B movement patterns). In fact, the cosine-based similarity function is a measure of similarity between two vectors, here two SESIs, by measuring the cosine of the angle between them (Eq. 6-1). The cosine-based similarity is non-negative and bounded between [0, 1] because the values of frequency and duration of QTC_B movement patterns are non-negative. The distance between two SESIs can be calculated by using a distance measure (Eq. 6-2), which in turn can be used in traditional clustering algorithms.

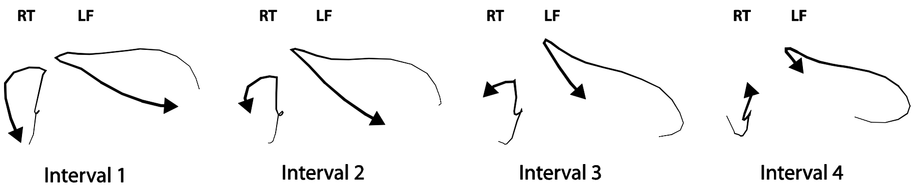
$$Sim(A,B) = \frac{A.B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \cdot B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} * \sqrt{\sum_{i=1}^{n}(B_i)^2}} \quad \text{Eq. 6-1}$$

$$Dis\,(A,B) = 1 - Sim(A,B) \quad \text{Eq. 6-2}$$

The individual scalar components A_i and B_i are called features or attributes, which denotes the frequencies /durations of the i^{th} cell in SESI A and SESI B, respectively.

The result of the distance measure is used in building an agglomerative hierarchical clustering, which yields a dendrogram that represents a nested grouping of hyper-SESIs based on their distance from each other. Figure 6-9 demonstrates the distances between SESIs of right toe-left finger at four time intervals of movement, both for frequency- and duration-based SESIs. In this figure, the outcomes of the distance measure are presented in the form of distance matrices in which each cell in the matrix represents the distance between the movements of right toe-left finger at two time intervals. Accordingly, the dendrograms represent the clusters of movements, where they confirm that more similar movements are grouped into a cluster. Figure 6-9 illustrates that the movements of right toe-left finger at time interval 1 and 4 share more similarity in QTC_B movement patterns than at other time intervals. This finding is true for both measurements, those either based on the frequency or the duration values of the mapped QTC_B movement patterns.

To calculate the overall similarity between two hyper-SESIs, an average similarity is taken into account. This approach means that an average is taken from all of the similarity measures that were obtained from the peer to peer comparison of all of the SESIs (i.e. right finger-right toe, right finger-left toe, right finger-left finger, right toe-left toe, right toe-left finger, and left toe-left finger) at each time interval of movement. The results are shown in Figure 6-10, where the overall performance of the Samba dancer at time interval 1 is more similar to that at time interval 3 considering the frequency values, while the results of comparing duration-based SESIs confirms the higher similarity between time intervals 3 and 4. The result reveals that there are not as many differences between hyper-SESIs at different time intervals as expected. Figure 6-3b shows that this result is reasonable because the trajectories of the MPOs are almost the same. Note that the dance application is only one example of a wide range of applications that the proposed methodology can be applied to. In summary, the proposed methodology is appropriate for comparing interactions of MPOs based on the generated QTCB movement patterns between them.



frequency - based similarity

Distance Matrix (RT, LF)	Interval 1	Interval 2	Interval 3	Interval 4
Interval 1	0	0.042	0.053	0.022
Interval 2		0	0.089	0.068
Interval 3			0	0.047
Interval 4				0

duration- based similarity

Distance Matrix (RT, LF)	Interval 1	Interval 2	Interval 3	Interval 4
Interval 1	0	0.245	0.372	0.048
Interval 2		0	0.349	0.246
Interval 3			0	0.062
Interval 4				0

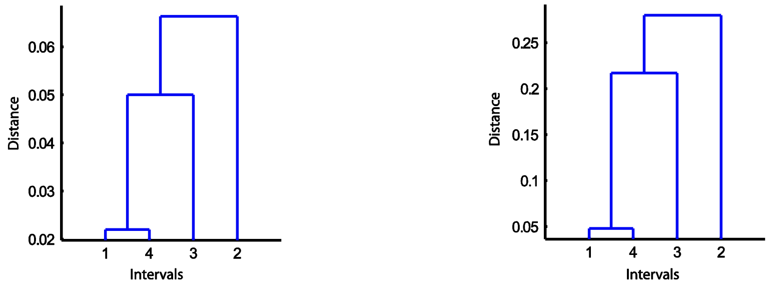


Figure 6-9: Distance between the movements of the right toe-left finger at four time intervals of movement alongside the dendrograms that represent the agglomerative hierarchical clustering of movements.

frequency - based similarity

Distance Matrix (RF, RT, LT, LF)	Interval 1	Interval 2	Interval 3	Interval 4
Interval 1	0	0.030	0.023	0.031
Interval 2		0	0.039	0.043
Interval 3			0	0.030
Interval 4				0

duration- based similarity

Distance Matrix (RF, RT, LT, LF)	Interval 1	Interval 2	Interval 3	Interval 4
Interval 1	0	0.054	0.090	0.044
Interval 2		0	0.086	0.081
Interval 3			0	0.023
Interval 4				0

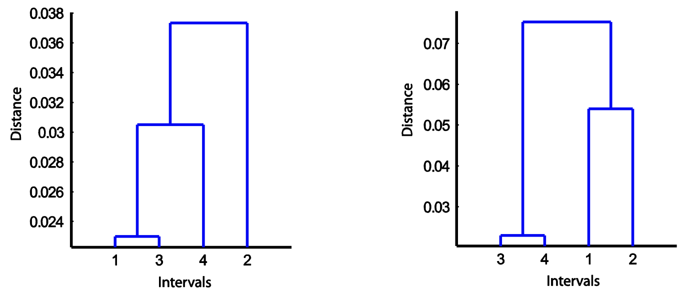


Figure 6-10: Distance between hyper-SESIs at four time intervals of movement alongside the dendrograms that represent the agglomerative hierarchical clustering of movements.

Throughout the chapter, we have shown the possibilities of the proposed method for extracting knowledge from the hyper-SESIs obtained from the interactions of MPOs. For example, a complex movement of a pair of MPOs is visually distinguishable from a simple movement by comparing the SESIs that cover the same time interval. Apart from a visual analysis, however, we could also measure the similarity between hyper-SESIs. For example, a complex movement of a pair of MPOs is visually distinguishable from a simple movement of a pair of MPOs by comparing the SESIs that cover the same time interval.

6.5 Discussion

This chapter has addressed the usefulness of QTC information when employed in the following knowledge discovery tasks: (i) representation of movement patterns obtained from qualitatively interpreted interactions of multiple MPOs, (ii) similarity assessment of movement patterns using hyper-SESIs, and (iii) clustering of collective movement. The applicability of the method has been presented using a dance movement data set. Next, we will note some strengths and weaknesses of the presented approach.

(a) In contrast to many research areas in knowledge discovery of moving objects that rely on comparisons of trajectories, this research has focused more attention to the interaction patterns of MPOs.

(b) The presented approach, which is based on qualitative relations, enables us to investigate the movement behaviour of multiple MPOs. However, it is generally accepted that qualitative and quantitative formalisms can complement each other. This statement means that quantitative approaches can be applied as well to investigate the outputs of our approach.

(c) From a representational point of view, distinguishing long movement patterns in the SESI is challenging because of resolution issues. Given that SESIs are fractal concepts, mapping long movement patterns requires high resolution cells obtained by iterating the SESI based on the length of the considered time interval. However, from a computational perspective, this concern is not necessarily problematic.

(d) Usually, the required accuracy of the results of the analysis depends on the type of application at hand. Representing the movement interactions with a qualitative representation such as QTC_B (our approach) significantly reduces the accuracy of the result. However, other types of QTC, such as QTC Double-Cross (QTC_C), can incorporate more relevant information about the movement. In addition to the Euclidean distance considered in QTC_B, in QTC_C, directional information between two moving objects is included. Consequently, to visualise the QTC_C movement patterns, a new high resolution hyper-SESI is needed because each QTC_C relation is represented by a four-tuple [for more explanation about QTC_C, see (Van de Weghe et al., 2005)].

(e) The proposed approach can be applied in different domains where the analysis of the interactions between objects is important. For example, the interactions between football players during a game can be examined by using the proposed methodology. However, it is a complicated task for investigating the interactions between moving objects in a large data set of movements, such as migratory birds.

(f) Delineating the visualised QTCB movement patterns on SESIs might be a difficult task, especially for infrequent and non-durable movement patterns. However, having prior knowledge of which movement patterns could occur during the interactions of MPOs based on the context of the application can greatly help in interpreting the results.

6.6 Conclusions and Future Work

In this chapter, we introduced a new methodology for similarity detection of movement behaviour of multiple MPOs. This method is based on QTC. In this study, the most fundamental type of QTC, namely, QTC Basic (QTC_B), has been used. The usefulness and applicability of QTC to identify, visualise and analyse the movement behaviour of multiple MPOs has been demonstrated, starting from raw trajectory data to hyper-SESIs. As a future extension of our approach, we intend to develop a SESI for other types of QTC. Furthermore, we intend to enrich the developed approach by incorporating descriptive statistical analyses. These will provide summaries about QTC movement patterns in different time intervals of movement, to have more insight into the movement data. In an applied setting, it is not possible to relate interactions of MPOs (i.e. QTC movement patterns) to trajectory pairs. However, in a more advanced implementation of the work, this weakness can be addressed.

References

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Andrienko, G., & Andrienko, N. (2010). A general framework for using aggregation in visual exploration of movement data. *The Cartographic Journal, 47* (1), 22-40.

Andrienko, N., & Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. *Cartographica, 42* (2), 117-138.

Andrienko, N., & Andrienko, G. (2013). Visual analytics of movement: An overview of methods, tools and procedures. *Information visualization, 12* (1), 3-24

Andrienko, N., Andrienko, G., Wachowicz, M., & Orellana, D. (2008). Uncovering interactions between moving objects. In: T. J. Cova, H. J. Miller, K. Beard, A. U. Frank & M. F. Goodchild (Eds.), Proceedings of the *5th International Conference on GIScience* (Vol. 5266, pp. 16-26). Park City, Utah.

Bak, P., Marder, M., Harary, S., Yaeli, A., & Ship, H. J. (2012). Scalable detection of spatiotemporal encounters in historical movement data. *Computer Graphics Forum, 31* (3pt1), 915-924.

Barnsley, M. F. (2000). *Fractals Everywhere* (2nd ed.). San Francisco: Morgan Kaufmann.

Buchin, K., Buchin, M., Van Kreveld, M., & Luo, J. (2009). Finding long and similar parts of trajectories. *Proceedings of the 17th International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS)* (pp. 296-305).

Buzan, D., Sclaroff, S., & Kollios, G. (2004). Extraction and clustering of motion trajectories in video. *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)* (Vol. 2, pp. 521-524).

Cohn, A. G., Bennett, B., Gooday, J., & Gotts, N. M. (1997). Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica, 1*, 275-316.

Delafontaine, M., Bogaert, P., Cohn, A. G., Witlox, F., De Maeyer, P., & Van de Weghe, N. (2011a). Inferring additional knowledge from QTC_N relations. *Information Sciences, 181* (9), 1573-1590.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011b). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Demsar, U., & Virrantaus, K. (2010). Space-time density of trajectories: Exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science, 24* (10), 1527-1542.

Ding, H., Trajcevski, G., Scheuermann, P., & Soc, I. C. (2008). Efficient similarity join of large sets of moving object trajectories. *Proceedings of the 15th International Symposium on Temporal Representation and Reasoning (Time 08)* (pp. 79-87).

Dodge, S. (2011). *Exploring Movement Using Similarity Analysis*. University of Zurich, Zurich.

Dodge, S., Weibel, R., & Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers Environment and Urban Systems, 33* (6), 419-434.

Dodge, S., Weibel, R., & Lautenschutz, A. K. (2008). Towards a taxonomy of movement patterns. *Information Visualization, 7* (3-4), 240-252.

Egenhofer, M. J., & Franzosa, R. D. (1991). Point-set Topological spatial relations. *International Journal of Geographical Information Systems, 5* (2), 161-174.

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence, 24* (1-3), 85-168.

Freksa, C. (1992). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Galton, A. (2001). Dominance diagrams: A tool for qualitative reasoning about continuous systems. *Fundamenta Informaticae, 46* (1-2), 55-70.

Giannotti, F., & Pedreschi, D. (2008). Mobility, data mining and privacy: A vision of convergence. In: F. Giannotti & D. Pedreschi (Eds.), *Mobility, Data Mining and Privacy-Geographic Knowledge Discovery (pp. 1-11)*. Berlin Heidelberg: Springer.

Gudmundsson, J., van Kreveld, M., & Speckmann, B. (2007). Efficient detection of patterns in 2D trajectories of moving points. *Geoinformatica, 11* (2), 195-215.

Imfeld, S. (2000). *Time, Point and Space - Towards a Better Analysis of Wildlife Data in GIS.* University of Zurich, Zurich.

Jensen, C. S., Lin, D., & Ooi, B. C. (2007). Continuous clustering of moving objects. *IEEE Transactions on Knowledge and Data Engineering, 19* (9), 1161-1174.

kraak, M. J., & Van De Vlag, D. (2007). Understanding spatiotemporal patterns: Visual ordering of space and time. *Cartographica, 42* (2), 151-161.

Laube, P., Duckham, M., & Palaniswami, M. (2011). Deferred decentralized movement pattern mining for geosensor networks. *International Journal of Geographical Information Science, 25* (2), 273-292.

Laube, P., Duckham, M., & Wolle, T. (2008). Decentralized movement pattern detection amongst mobile geosensor nodes. In: T. Cova, H. Miller, K. Beard, A. Frank & M. Goodchild (Eds.), *Geographic Information Science (pp. 199-216)*. Berlin Heidelberg, Springer.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Li, Y., Han, J., & Yang, J. (2004). Clustering moving objects. In: W. Kim, R. Kohavi, J. Gehrke & W. DuMouchel (Eds.), *Proceedings of the 10th International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD)* (pp. 617-622). Washington: ACM.

Lin, B., & Su, J. W. (2008). One way distance: For shape based similarity search of moving object trajectories. *Geoinformatica, 12* (2), 117-142.

Mountain, D. M. (2005). *Exploring mobile trajectories: An investigation of individual spatial behaviour and geographic filters for information retrieval.* City University, London.

Nanni, M., & Pedreschi, D. (2006). Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems, 27* (3), 267-289.

Ooms, K., Andrienko, G., Andrienko, N., De Maeyer, P., & Fack, V. (2012). Analysing the spatial dimension of eye movement data using a visual analytic approach. *Expert Systems with Applications, 39* (1), 1324-1332.

Pelekis, N., Andrienko, G., Andrienko, N., Kopanakis, I., Marketos, G., & Theodoridis, Y. (2012). Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems, 38* (2), 343-391.

Pelekis, N., Kopanakis, I., Marketos, G., Ntoutsi, I., Andrienko, G., & Theodoridis, Y. (2007). Similarity search in trajectory databases. *Proceedings of the 14th International Symposium on Temporal Representation and Reasoning (TIME 2007)* (pp. 129-140).

Rinzivillo, S., Pedreschi, D., Nanni, M., Giannotti, F., Andrienko, N., & Andrienko, G. (2008). Visually driven analysis of movement data by progressive clustering. *Information Visualization, 7* (3-4), 225-239.

Tominski, C., Schumann, H., Andrienko, G., & Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. *IEEE Transactions on Visualization and Computer Graphics, 18* (12), 2565-2574.

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Van de Weghe, N., Cohn, A. G., Maeyer, P. D., & Witlox, F. (2005a). Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems with Applications, 29* (4), 977-983.

Wilson, C. (2008). Activity patterns in space and time: calculating representative Hagerstrand trajectories. *Transportation, 35* (4), 485-499.

Xia, L., & Kraak, M. J. (2010). A temporal visualization concept: A new theoretical analytical approach for the visualization of multivariable spatio-temporal data. *Proceedings of the IEEE 18th International Conference on Geoinformatics* (pp. 1-6), Beijing, China.

Zhang, Q., & Lin, X. (2004). Clustering moving objects for spatio-temporal selectivity estimation. In: K. D. Schewe & W. Hugh (Eds.), *Proceedings of the 15th Australasian database conference (ADC '04)* (Vol. 27, pp. 123-130).

7

Alignment of Patterns

Science never solves a problem without creating ten more. George Bernard Shaw

Modified From: Seyed Hossein Chavoshi, Bernard De Baets, Tijs Neutens, Matthias Delafontaine, Guy De Tré, Nico Van de Weghe (2014). Exploring Movement Data Based on Sequence Alignment Methods. Under Review.

7 Alignment of Patterns

Abstract: Despite the abundance of methodological research concerning knowledge discovery from moving object databases, only a limited number of studies have examined the interaction between moving point objects in space over time. This chapter describes a novel approach for measuring the similarity among interactions between moving objects. The approach is based on a qualitative formalism that describes the relative motion of two disjoint moving point objects in terms of distance. The proposed approach consists of three steps. First, we transform movement data (i.e. geographical coordinates of moving objects) into sequences of successive qualitative relations based on Qualitative Trajectory Calculus (QTC). A sequence of qualitative relations represents the interaction between a pair of moving objects during a particular time interval. Second, sequence alignment methods are applied to measure the similarity between the movement sequences. Finally, the movement sequences are grouped based on similarity by means of an agglomerative hierarchical clustering method. We seek to investigate the extent to which the interaction of one pair of moving objects resembles the interaction of another pair. The applicability of this approach is tested using movement data from Samba dancers.

7.1 Introduction

Technological advances in tracking and navigation systems make it possible to capture, efficiently and cost-effectively, the trajectories of a wide range of moving objects, including human beings (Michael et al., 2006; Wang et al., 2003), animals (DeCesare et al., 2005; Gagliardo et al., 2007; Gau et al., 2004; Laube et al., 2005), and vehicles (Brakatsoulas et al., 2005; Hvidberg, 2006). With access to an unprecedented wealth of accurate motion data, researchers today can apply pattern discovery techniques to moving object databases and generate knowledge to inform many disciplines, including urban planning (van Shaick & van der Spek, 2008), event management (Versichele et al., 2012a; Versichele et al., 2012b), crisis management (Pan et al., 2007), traffic (Ong et al., 2011), and tourism (Orellana et al., 2012). In addition to their usefulness for processing large-scale movement data sets, data mining and knowledge discovery techniques can also be applied to small-scale movement data sources. For example, movement patterns, such as walking, running, jumping, lifting, striking and swimming, can be investigated for various purposes. Investigating the movement of swimmers, for instance, helps coaches to analyse the performance of their swimmers (Guerra-Salcedo et al., 2005). Nonetheless, the specific techniques and methods chosen for extracting movement patterns from a data set depend on the context of the movement under examination. Among the wide range of research methodologies, similarity analysis has attracted considerable attention from many researchers. The similarity between two entities is measured as the cost of transforming one entity into another via a similarity measure (Faloutsos et al., 1997). In the context of movement, trajectories (i.e. representative paths that moving objects follow through space as a function of time) are typically considered to be the entities in similarity analyses of the dynamic behaviour of moving objects. Among the existing research that has applied similarity analysis to the study of moving object trajectories, most studies have focused on the spatial dimension (Chen et al., 2005; Lin & Su, 2005; Vlachos et al., 2002; Yanagisawa et al., 2003), whereas several studies have considered both spatial and temporal aspects (Buchin et al., 2009; Frentzos et al., 2007; Pelekis et al., 2007; Sinha & Mark, 2005; Van Kreveld & Luo, 2007). However, despite extensive research in this field (Giannotti & Pedreschi, 2008; Miller & Han, 2009), certain aspects of moving object trajectories have received only scant attention to date.

In this chapter, instead of presenting a spatial or spatio-temporal similarity analysis of trajectories, we propose a framework in which the similarity measure is used to quantify similarity when pairs of moving objects interact with one another. We believe that a focus on the similarity in the interactions among moving object pairs may reveal more information on object movement than a sole focus on object trajectories.

To form the basis of the similarity analysis, a qualitative formalism appropriate for the representation of spatio-temporal human cognition is used to express the interactions between objects. To date, researchers have proposed several formalisms for the qualitative analysis of spatial and temporal phenomena. However, the existing work in this area has been limited to either spatial or temporal qualitative calculi (Allen, 1983; Frank, 1996; Freksa, 1992; Randell et al., 1992), with only a few studies presenting an integrated, spatio-temporal treatment of object movements. One notable example of an integrative approach is Qualitative Trajectory Calculus (QTC) (Van de Weghe, 2004). QTC reduces the complexity of interacting real-world continuously disjoint moving objects by representing the interaction in terms of qualitative relationships (Van de

Weghe et al., 2007). By converting relative motion attributes (i.e. distance) into symbolic representations, QTC transforms quantitative data on movement (positional information) into qualitative data (QTC relations), resulting in the simplified representation of trajectory pairs. The practicality and appropriateness of QTC for analysing the interaction of moving objects have been successfully demonstrated in various applications (Delafontaine et al., 2011a; Delafontaine et al., 2012a; Delafontaine et al., 2011b; Van de Weghe et al., 2005b).

In this chapter, we cross-pollinate QTC with sequence alignment methods (SAMs) to identify similarities in the movement behaviour between pairs of interacting moving objects over time. Although SAMs have long been used in bioinformatics for the analysis of DNA strings (Morrison, 2010), it has only recently been applied to the field of movement analysis (2007a). In the current study, SAM is used to assess the similarity between movement sequences of QTC relations for three reasons. First, SAMs allow us to visually distinguish movement patterns from sequences and extract insightful information from them. Second, the comparison of movement patterns using SAMs results in a quantitative measure of similarity between movement patterns. Finally, the results of a similarity analysis are used to cluster movement data into groups that share similar properties. The usefulness of our approach will be demonstrated in an empirical case study in which SAM is used to examine the movement patterns of different parts of the body of Samba dance performers.

The remainder of this chapter is organised as follows: Section 7.2 provides a brief review of the background and basis of QTC and SAM. Section 7.3 presents a description of the data set used in this chapter is given. Section 7.4 presents the methodology that is applied in this research. Section 7.5 discusses the concept, compares it with related approaches and identifies strengths and open problems. Finally, Section 7.5 presents our concluding remarks and outlines the directions for future work.

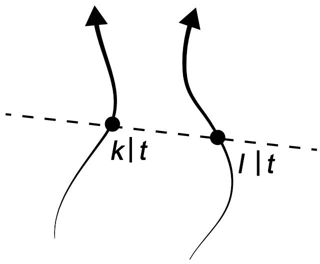
7.2 Background

7.2.1 The Qualitative Trajectory Calculus

The qualitative Trajectory Calculus (in short QTC) was introduced by Van de Weghe (2004) as a qualitative calculus technique to represent and reason about moving objects. The technique expresses the spatio-temporal relationship between two disjoint moving point objects (MPO_S). Different types of QTC have been developed, namely QTC_B (QTC-Basic) (Van de Weghe et al., 2006), QTC_C (QTC Double-Cross) (Van de Weghe et al., 2005a), QTC_N (QTC-Network) (Bogaert et al., 2007), and QTC_S (QTC-Shape) (Van de Weghe et al., 2004). For simplicity, this chapter will focus only on the basic level of QTC_B (see Section 2.5).

QTC_B provides a qualitative representation of the two-dimensional movement of a pair of MPOs (Figure 7-1). Binary relations between two MPOs are evaluated based on Euclidean distance (Van de Weghe, 2004). QTC_B relations are constructed from the relationships A and B introduced in Section 2.5. Accordingly, the $(A\ B)_B$ relationship syntax is used to represent the relation between two MPOs. In total, there are 9 (3^2) base relations for QTC_B (Figure 7-1b). For example, the QTC_B relation (+ +) indicates that the two objects are moving away from each other.

(a)



(b)

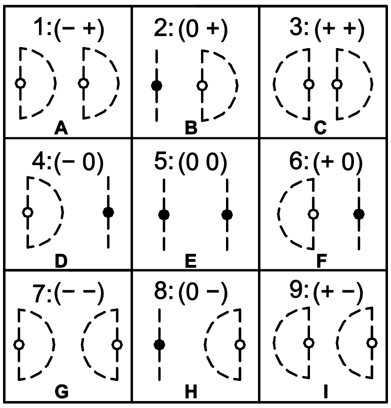
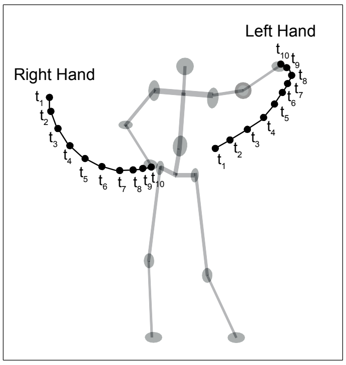


Figure 7-1: Two MPOs *k* and *l* and their trajectories used to form QTC_B base relations. The frame of spatial reference is illustrated by the dashed line. (a), nine QTC_B base relations (b)

The trajectory of an MPO comprises a set of observations through space and time (Hornsby & Egenhofer, 2002). MPOs have been used to represent different objects in diverse research fields ranging from animal studies to fleet management. The interactions between two MPOs during a time interval of movement can be expressed in the form of a QTC movement sequence - a chronological sequence of consecutive transitions between QTC relations. Figure 7-2 illustrates the movement of a pair of MPOs (i.e. hands of a dancer) during a 10-second interval with its QTC_B relations at each time stamp.



QTC_B Relations between
Left and Right Hands of a dancer

QTC_B	time
Unknown	1
[+ -]	2
[+ -]	3
[+ -]	4
[+ -]	5
[+ -]	6
[+ -]	7
[+ -]	8
[0 -]	9
Unknown	10

Figure 7-2: QTC_B relations of the movement of two hands of a dancer

7.2.2 Sequence Alignment Methods

Sequence alignment methods (SAMs) have played an important role in many research fields. In the early 1980s, biochemists began to use sequence analysis to analyse DNA sequences (Shoval & Isaacson, 2007a). Later, social scientists, such as the sociologist Abbott (1995), have applied sequence alignment to the analyses of musicians' careers. More recently, sequence alignment methods have been used in fields including transportation (Joh et al., 2002; Wilson, 2001; Wilson, 2008), cartography (Fabrikant et al., 2008), tourism (Shoval & Isaacson, 2007b), and crowd behaviour analysis (Delafontaine et al., 2012c), among others.

Sequence alignment is the process of aligning two or more character sequences based on a set of conventional operations. Specifically, dynamic programming algorithms are used to equate sequences with the goal of maximising a similarity measure or minimising a distance measure between them (Wilson, 2008). Two of the most widely used SAMs are pairwise alignment and multiple alignment. Pairwise alignment is the comparison of two sequences, whereas multiple alignment is the comparison of more than two sequences. Pairwise alignment and multiple alignment both operate on the basis of two primary types of algorithms: (i) global alignment and (ii) local alignment. Global alignment forces the alignment to span the entire length of all sequences, whereas local alignments identifies regions of similarity in long sequences (for a detailed explanation, see e.g., (Rosenberg, 2009)).

Pairwise alignment equates two sequences using four conventional operations: *identity, substitution, insertion*, and *deletion.* Based on the scope of the research, each operation is associated with a numbrt of cost/penalty that is defined *a priori* using a scoring matrix. To clarify the process, an example of a pairwise alignment of two sequences is presented in Figure 7-3. Specifically, the character strings '*DANCE'* and '*TANZ'*

(English and its German equivalent) are subjected to pairwise alignment, which reveals two *identities*, two *substitutions* and one *insertion/deletion*. The more operations required for the alignment of two sequences, the greater the distance (or the lower the similarity) between the sequences. A scoring matrix contains the entire set of pairwise substitution scores. In this scoring matrix, additive scoring is defined in which *identify* denotes the highest similarity and, as a result, the highest score in the scoring matrix. Typically, *substitution* has a lower score and *insertion/deletion* is associated with negative scores (i.e. penalties). Similar to pairwise alignment, a type of multiple pairwise alignment is known as progressive alignment (Wilson, 2006).

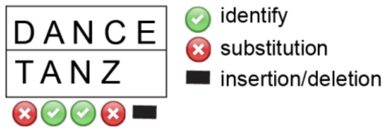


Figure 7-3: Pairwise alignment of two words dance and tanz in English and German, respectively

In Section 7.4, we describe how SAMs can be used to align QTC movement sequences derived from the way in which dancers move different parts of their bodies and cluster these sequences based on similarity. Samba dance is intentionally chosen in light of the rhythmicality of movement patterns in Samba and similarity assessment problems using SAMs.

In addition to the visual analysis of aligned QTC movement sequences, we also present an objective assessment regarding how well dancers (i.e. students in this case study) follow the instructions given by an instructor. In other words, our goal was to identify the aspects of students' performances in which movement patterns of the dancers matched or deviated from the instructor's movements. Samba dance is a rhythmical dance based on many set movements. Characterising the conformity of Samba dance movements is highly meaningful given that Samba is a dance that involves a group of dancers rather than a single one.

Synchronicity in performance is the factor that most effectively draws people's attention. Not only is synchronicity important to dancing, it may be used as a qualification measure for other types of movements such as synchronised swimming, i.e. a hybrid form of swimming, dance and gymnastics, which consists of swimmers performing a synchronised routine of complicated moves in the water, accompanied by music.

7.3 Data

The raw data used in this study come from the movements of three Samba dancers (a teacher, student 1, and student 2) during a given time interval. The movements of the dancers' heads, torsos, right and left hands, and right and left feet at each time stamp of a considered time interval of 3.64 s (temporal granularity of 0.04 s) were recorded from real dancers using the MoCap (MotionCapture) system owned by the Department of Musicology at Ghent University. MoCap is a movement retrieval technique that records the position of objects over time by means of reflective markers attached to these objects in combination with infrared cameras. It is used in a wide range of research fields. For example, MoCap has been used in sport sciences to capture the movement of athletes as part of rehabilitation, physical education and practice (Brodie et al., 2008; Mirabella et al., 2011). In the medical sciences, physiotherapists, orthopaedists and neurologists may examine MoCap measurements of human gaits in conjunction with biomechanical modelling to evaluate a patient's status and develop plans for treatment and rehabilitation (Colombo et al., 2013). Our data set is in the following format: t (i.e. the time stamp of movement), x, y, and z (i.e. the local positional information in a three-dimensional space) of each captured body part. The recorded positions of the markers were transformed into coordinates using the torso of a dancer's body as origin. Figure 7-4 depicts the configuration of the MoCap system, consisting of 14 infrared cameras, used to capture the Samba dance movements of the teacher and students performing basic Samba movements. Across 92 time units, many repetitive movements were observed from the performances of the teacher and the two students.

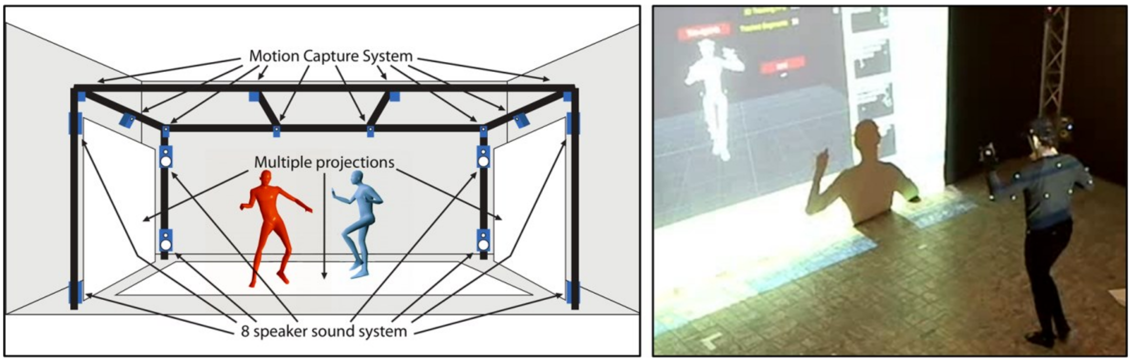


Figure 7-4: The MoCap system at the Institute for Psychoacoustics and Electronic Music (IPEM), Department of Musicology at Ghent University, where the data were captured

7.4 Methodology

The existing measures of trajectory similarity can be classified into two major categories (Dodge, 2011): (i) spatial similarity measures that concentrate on the geometric shape of the trajectories; and (ii) spatio-temporal similarity measures that consider both the spatial and temporal dimensions of the trajectories. In contrast to existing methods of classification, in this study, we measure the similarity of interactions between pairs of MPO. In other words, instead of comparing individual trajectories, we compare pairs of trajectories for similarity. We follow three major steps (Figure 7-5). First, raw trajectories of interacting MPOs from location-aware technologies are converted into qualitative relations (QTC_B). Second, sequences of the qualitative relationships are aligned for the interpretation of the movement patterns of MPOs. Finally, the results of the alignment are used to identify hierarchical clusters of qualitatively distinct sequences. Each step is discussed in depth below. To show the applicability of the proposed methodology, we explain and test the approach on real movement data from three Samba dancers.

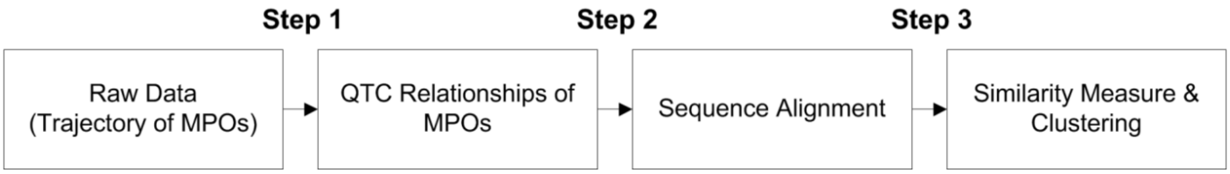


Figure 7-5: Process overview

7.4.1 Step 1: Converting Raw Trajectories into Qualitative Relations

In the first step, the relationships between different parts of the body of the three Samba dancers are described in terms of QTC_B relations. Figure 7-6 presents the movement of the dancers' heads, torsos, right and left hands, and right and left feet in a given time interval from both the front view and the side view (45º). The trajectories of the teacher's body parts to those of the students reveal several minor differences. For example, from the front views displayed in Figure 7-6, we can observe that the space used by students to move their hands was quite different compared to that of their teacher. Next, for simplicity, QTC_B relations were transcoded into single-character sequences. The corresponding character code for each base relation in QTC_B is presented in Figure 7-1b (below each representation).

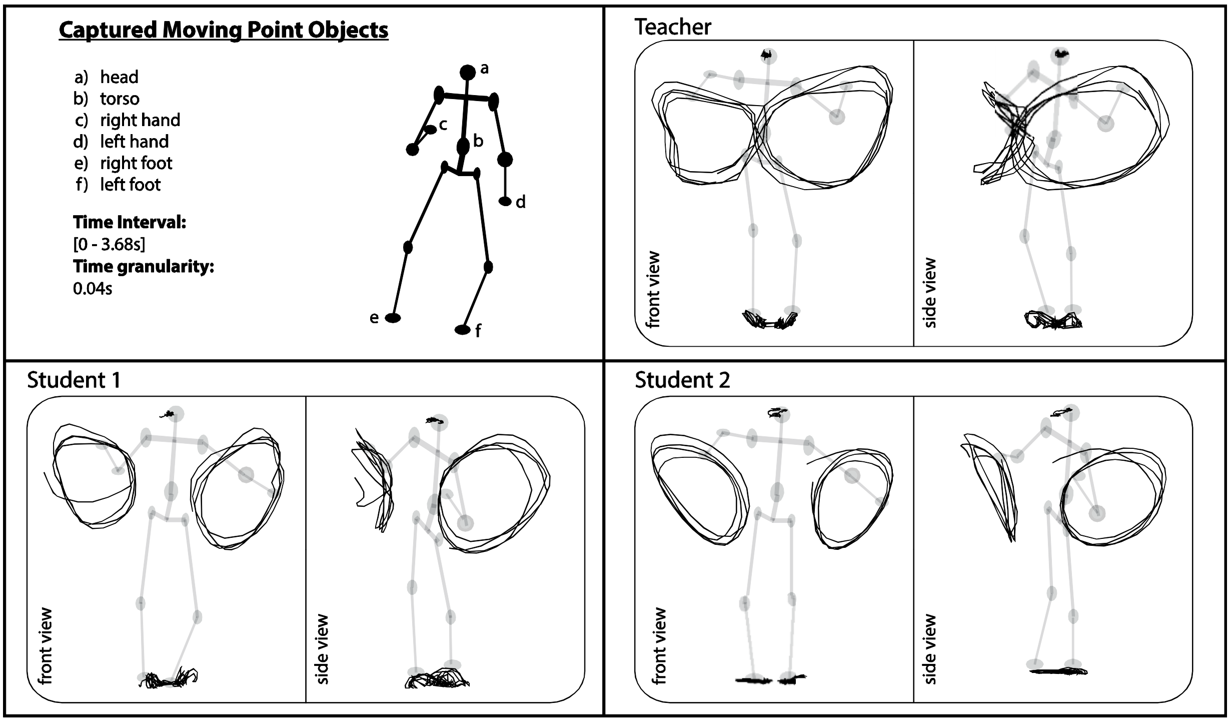


Figure 7-6: Derived trajectories of the movements of different parts of the body of Samba dancers using the MoCap system

Figure 7-7 presents the entire set of transcoded sequences of QTC_B relations between the different limbs of the teacher (i.e. $n(n-1)/2$ with n the number of body parts) in a movement lasting 3.64 s (temporal granularity of 0.04 s).

Index	Relation		
1	Head - Root	1:	F F F F EDD E F E E F F F EDD E F F F F F F F F F F F F E E F F F E E F F F EDDDDDD E F F EDDDD E EDD E F F EDDDDDDDD E F EDDDDD EDD E F F F EDDDDDD
2	Head - Left Hand	2:	ABCCCAAAAADGGGGH I I I I I FBADAAAAAAAAADHHGHHGH I I I I I I I FCAAAAAAADGGGGGH I I I I I I I I FCBAAAAAAADGGGGH I I I I I
3	Head - Right Hand	3:	HGGGH I I I I I HGHCCBAAAAAAADGGGH I I I I I I I HACBCCBAAAADGGGH I I I I I I I HHCCCBAAAAADGGGH I I I I I I I I CCCCCBAAAA
4	Head - Left foot	4:	I I F F I HDAAABAADE I I I C F I I I I I I H EAAADGA EDAA EDBC F I F F I I H EAAAADGA EDAADH EB F I F F I I I EAAAADGC F EBB E I I I F I E
5	Head - Right Foot	5:	DAADH I F I I I I I I I I HGDAAAADDAA E F I CC F FB E I I I EAAADGGAAAAA EHB E I F EH I I I EAAADGDAAAA E EHB F FC EH I I I I EAAAAA
6	Root - Left Hand	6:	BB EBBBBBBB EHHHHHHHHH EBB EHHBBBBBBB EHHHHHHHHHHBB EHHHHBBBBBB EEHHHHHHHHHH EHHHHHBBBBBBB EHHHHHHHHHH
7	Root - Right Hand	7:	EEHHHHHHHHHHHH EBBBBBBBB EEHHHHHHHHHHHHHHHBBBBBBBB EHHHHHHHHH EEHHHHH EBBBBBBB EHHHHHHHHHHHHHHHHBBBBBBB
8	Root - Left Foot	8:	H EB EHH EBBBBBB EEHHHHHHHHHHHBB EHHH EBBB EEB EEHHHHHHHHH EB EHHH EBBB EEEEHHHHHHHHHHHBB EHHHBBBB EEBB EHHHHHH
9	Root - Right Foot	9:	EBB EHHHHHHHHHHB EHH EBBBB EEBB EHHHB EEEHHH EHH EBBBBBBBB EHHHHH EEHHB EHHH EBBB EEB EEHHHHHH EEHHB EHH EBBBB
10	Left Hand - Right Hand	10:	H I I I I I I I I I I HGDAAAAAAAAAADH I I I I I I I I I I HDAAAAAAAAAAADG I I I I I I I I I I HDAAAAAAAAAAADGH I I I I I I I I I HDAAAAAAAAAA
11	Left Hand - Left Foot	11:	I H F C F I I I FCBDGGGGGAAACC FHGDC F I I I I FBDGGGGDAACCCC EGGC F I I I FCBDGGGDAAAACBBDGGC F I I I FCBDGGGGGAAAAA
12	Left Hand - Right Foot	12:	FCCC F I I I I I HDADGGGDAADGDBCCC F I I I I I HGDADGGDAADGDAACC F I I I I I I HDADDADAADGDABCC F I I I I I I HDADGDAAAAD
13	Right Hand - Left Foot	13:	HGGDAAADGGGDBCC F I I I I I I HHGGGDAAADGG ECCCC F I I I I I HGGGDAAAADGG ECCC F I I I I I I HGGGDAAAADGG ECCCC F I I I I I
14	Right Hand - Right Foot	14:	DGGDAAAAAAADGGH FCC F I I I I HGGGDAAAAAAAADGG FCC F I I I I HGGGDAAAACCBDGG FC I F F I I I HGGGDAAAAABADGG FC F I I I I I
15	Left Foot - Right Foot	15:	DAAD E I E I I I I I I I H EDDAAAADDAAD E I BC F F F EH I I HGDAAAAAAAAAA E I I F I F FHH I I EAD EBAADDAA E I I I I I F F I I I I EBAAAAA

Figure 7-7: All QTC_B movement sequences of the teacher during 3.64 s of movement

As stated earlier, Samba dance is a dance with numerous periodic movement patterns, which can be discovered via an analysis of the QTC movement sequences of dancers. One way to visually recognise the periodicity in movement sequences is mapping sequences to dot plots. From a dot plot, certain sequence features (such as ‘repeats’) can be visually identified (Mount, 2007). Dot plots are constructed using two sequences- one written along the top row and the second written along the leftmost column of a two-dimensional matrix. In a dot plot, each dot represents a point at which there is a match between the characters in the corresponding columns. Thus, it is possible to identify a certain number of matches in a sequence in a search window defined *a priori*. Repetitiveness in a single sequence can be assessed by plotting a sequence against itself in a dot plot and sections that share similarities become visible in the form of lines off the main diagonal. Figure 7-8 comprises dot plots of the QTC_B movement sequences for three pairs of body parts (i.e. left hand-right hand, left foot-right foot, and right hand- left foot) for the teacher, student 1, and student 2. To derive the plots, we run a window spanning 10 characters along movement sequences in which 8 characters are matched. Many repetitive sequences of relative movements can be observed in the dot plots of left hand-right hand for all three Samba dancers, whereas almost no repetition is observed in the QTC_B relations of left foot-right foot with a window of the same size. This pattern suggests that dancers focused more on the movement of their hands and less on the movement of their feet. In other words, regularity is more visible in the movement of hands than of feet. The neat straight lines in the left hand-right hand dot plot for the teacher indicate that regular and perfect repetitions of the teacher's movements over time. The lines in the dot plots for students 1 and 2 show various deviations and are not as straight as those of the teacher. These deviations are caused by lag and lead times in the repetition of the same movements by students. Based on these plots, we can roughly infer that the movements of student 1 and 2 are not as regular as the movements of the teacher. Next, we will further examine this irregularity in the students' movement via sequence alignment and attempt to identify them automatically.

Additional information can be retrieved from the histograms of QTC_B movement sequences (Figure 7-9). For instance, the histogram representing left hand-right hand relations in the dancers shows that the QTC_B relations $(-\ +)$ (i.e. character A) and $(+\ -)$ (i.e. character I) occur more frequently than other QTC_B relations (Figure 7-9). This is due to the nature of Samba dance, in which one hand follows the other hand most of the time. The histogram presenting left foot-right foot relations in the dancers reveals the low frequency of QTC_B relation $(-\ -)$ (i.e. character G). This is because the dancers' feet rarely moved towards/away from one another other during the particular dance fragment. However, it is possible for this pattern to be observed more frequently in other types of rhythmic movements. The patterns observed in the histograms of movement sequences may change from one type of dance to another. For example, there are varieties of dances in which the hands move towards/away from each other most of the time. In these cases, the QTC_B relations $(- -)$ and $(+ +)$ would create peaks in the corresponding histogram of movement sequences.

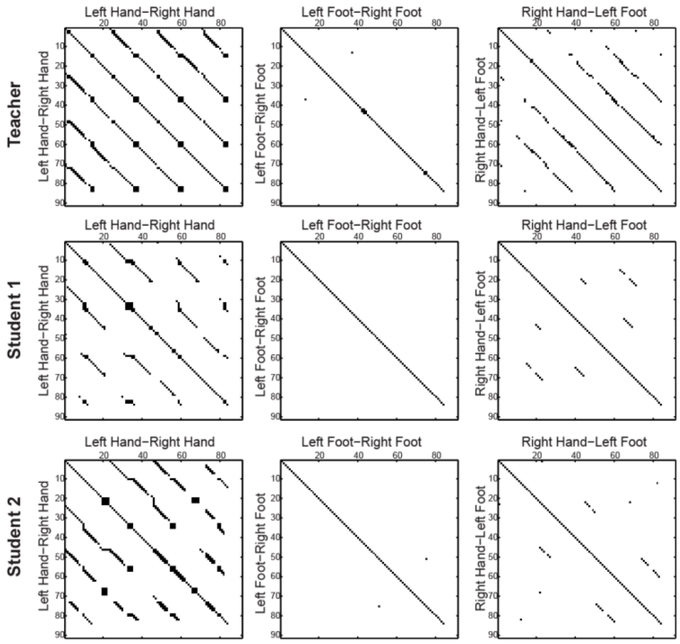


Figure 7-8: Dot plots of QTC_B movement sequences of left hand-right hand, left foot-right foot, and right hand-left foot for three Samba dancers

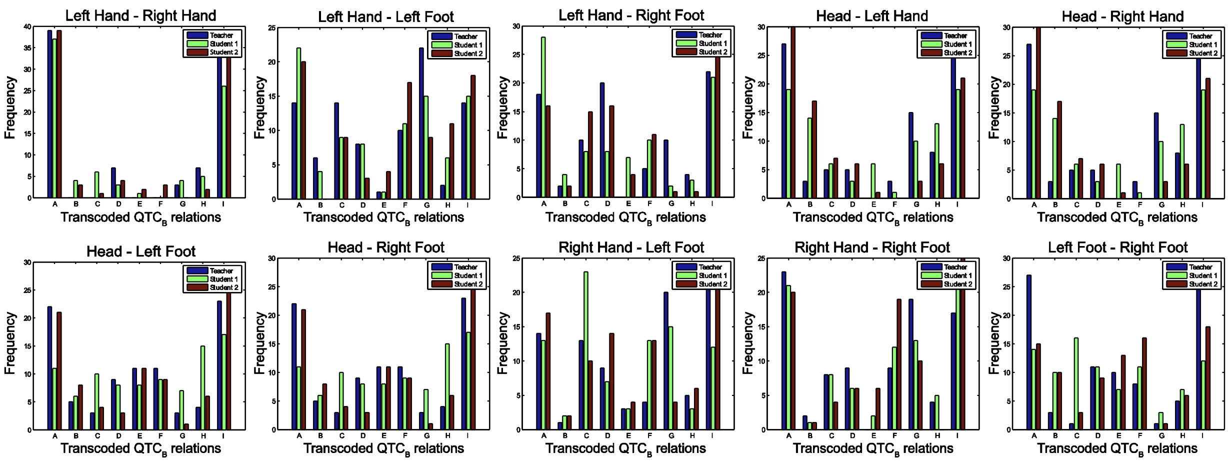


Figure 7-9: Histograms of QTC_B relations

7.4.2 Step 2: Aligning QTC Movement Sequences

In the second step, we align the QTC movement sequences of different body parts of the dancers. Using SAMs, we determine the degree of similarity between the movements of dancers during their performance. Finally, we cluster QTC movement sequences based on similarity and evaluate the overall performance of each dancer.

The main challenge is to optimally align the QTC movement sequences of the students with the movement sequences of the teacher. SAM is applied to identify the parts of the students' performance that matched or mismatched the performance of the teacher. When the difference between the aligned QTC movement sequences of the teacher and the student is sufficiently small, we can conclude that the student has performed her/his movement very well on the basis of the teacher's movements as the choreographic benchmark. Clearly, not all movements of the student's body comply with the benchmark. To visualise and analyse the (dis)similarity between the body movement of the students with respect to that of the teacher, we examine a time interval of 3.64 s of their performances. We deliberate keep the time interval short to make it easier to recognise (dis)similarity in the movement sequences and study the basic movements of dancers.

As mentioned earlier, the alignment of two sequences is based on minimising the distance between them (using a pre-defined scoring matrix). Running SAM on two sequences yields two measures: (i) the distance (or similarity) between two sequences and (ii) the best possible alignment of the two sequences, which is the alignment that minimises the overall distance between the two sequences.

A scoring matrix is developed based on the conceptual distance of QTC_B relations. The conceptual distance is defined as a measure of closeness of two QTC_B relations by counting the number of changes in the symbols of the QTC_B representation $(A\ B)_B$ (Van de Weghe & De Maeyer, 2005). The smallest conceptual distance is zero (i.e. the distance between a QTC_B relation and itself). The conceptual distance between '0' and '+' or '–' is one. The conceptual distance between '–' and '+' equals two (one for '–' to '0' and one for '0' to '+') because direct transition is impossible (Galton, 2001). The overall conceptual distances between two QTC_B relations can then be calculated by summing up the conceptual distance over both relation symbols and multiplying by 2.5 to rescale it to the interval [0 10]. Therefore, a similarity score between two QTC_B relations can be calculated as $(10 - 2.5 * \text{conceptual distance})$. Table 7-1 presents the resultant QTC_B scoring matrix. An exact character match is assigned a similarity score of 10 (maximal similarity) and a mismatch is given a similarity score of 0 (maximum conceptual distance). For example, the conceptual distance between the two QTC_B relations $(-\ +)$ (i.e. character A) and $(-\ 0)$ (i.e. character D) is equal to one. For every conceptual distance unit, the similarity score decreases by 2.5 units from the maximal similarity score of 10. Therefore, the similarity score between A and D is equal to 7.5.

Two parameters that need to be set in the process of sequence alignment are gap opening and gap extension. In this study, *insertion/deletion* penalties for gap opening and for gap extensions are, – 5 and – 3, respectively. In SAMs, dynamic programming algorithms are used in the search for optimal alignment to either maximise a similarity measure or minimise a distance measure based on the predefined scoring matrix (Wilson, 2008).

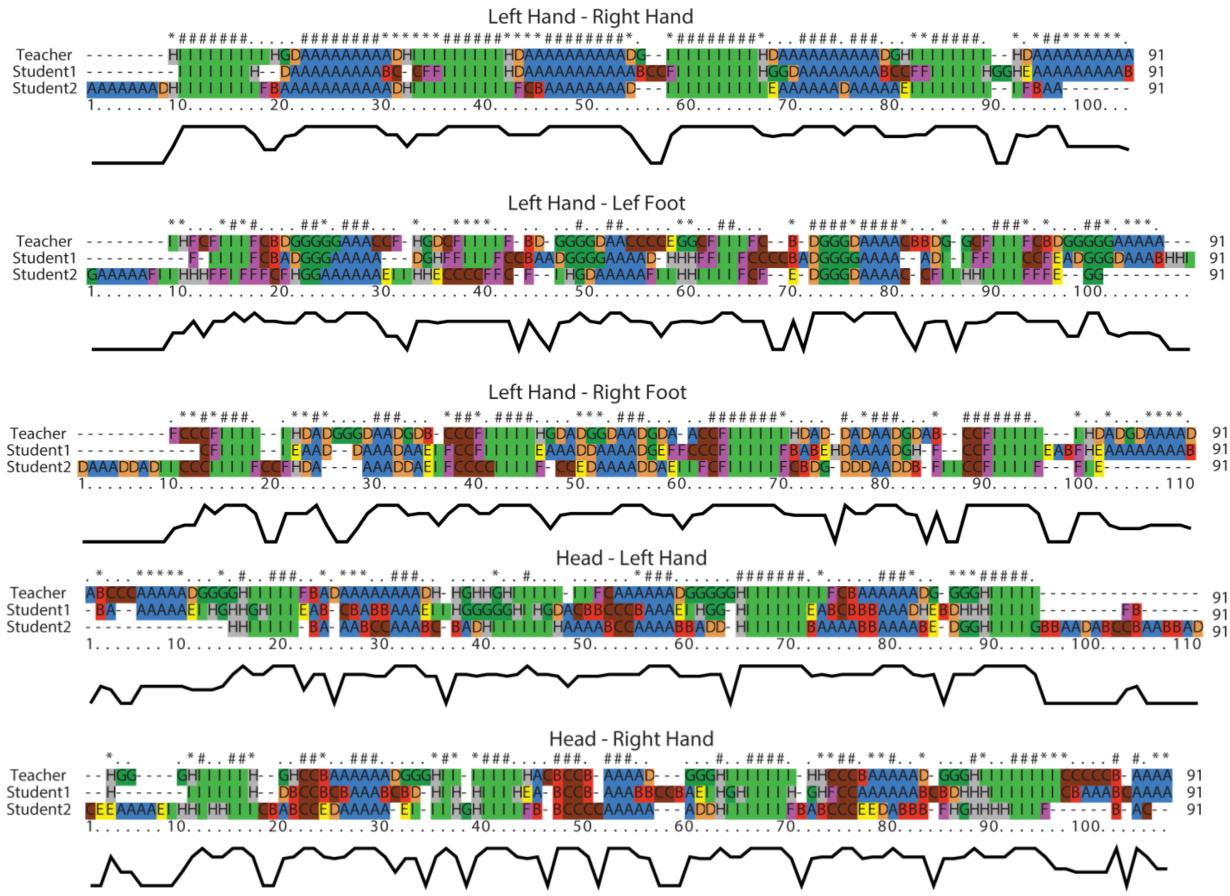
Each dancer has 15 possible QTC_B movement sequences representing 15 interacting pairs of body parts. Because the dancer's torso is used as a reference point for the movement of other body parts, movement sequences involving torso (i.e. root) are not considered in the alignment process. Using the specified similarity scores and penalties (Table 7-1), a multiple alignment of QTC_B movement sequences is generated with the ClustalTXY software package (Wilson, 2008) based on the progressive alignment procedure. At a given time, three corresponding QTC_B movement sequences (i.e. of the teacher and the two students) are aligned followed by a multiple alignment using a global alignment (Needleman & Wunsch, 1970).

Table 7-1: Sequence alignment scoring matrix for QTC_B relations

Similarity Matrix	A	B	C	D	E	F	G	H	I
A	10	7.5	5	7.5	5	2.5	5	2.5	0
B	7.5	10	7.5	5	7.5	5	2.5	5	2.5
C	5	7.5	10	2.5	5	7.5	0	2.5	5
D	7.5	5	2.5	10	7.5	5	7.5	5	2.5
E	5	7.5	5	7.5	10	7.5	5	7.5	5
F	2.5	5	7.5	5	7.5	10	2.5	5	7.5
G	5	2.5	0	7.5	5	2.5	10	7.5	5
H	2.5	5	2.5	5	7.5	5	7.5	10	7.5
I	0	2.5	5	2.5	5	7.5	5	7.5	10

QTC_B Relations	Code
(– +)	A
(0 +)	B
(+ +)	C
(– 0)	D
(0 0)	E
(+ 0)	F
(– –)	G
(0 –)	H
(+ –)	I

Figure 7-10 presents the results of the alignment of QTC_B movement sequences. For clarity, the characters (i.e. transcoded QTC_B relations) have been colour-coded. The line above the aligned sequences is used to mark strongly conserved positions. Four characters to indicate the degree of matches: '#' indicates positions that are 80%-100% identical, '*' indicates positions that are 60% -80% identical, ':' indicates positions that are 40% -60% identical, '.' indicates positions that are 20% -40% identical. The curve below the movement sequences represents the rate of changes in the match and mismatch of characters at each time stamp of movement after sequence alignment. Less fluctuation in curves with highly matched characters at each time stamp indicates more similarity between movement sequences. The results show (the lack of) regularity in dance movement patterns. For instance, the sequences representing the left hand-right hand relations exhibit periodicities in the dancers' movements. This pattern can be observed from the succession of colours and attributed to the way in which dancers paid more attention to the movement of hands than to other parts of the body. Moreover, the relative movements of head and hands show more regularity than the relative movements of head and feet, suggesting that dancers were more successful in adjusting the movement of the upper part of their body relative to the lower part. From the sequences of left foot-right foot relations, it can be observed that the rate of changes in movement patterns is rather high compared to those of the hands.



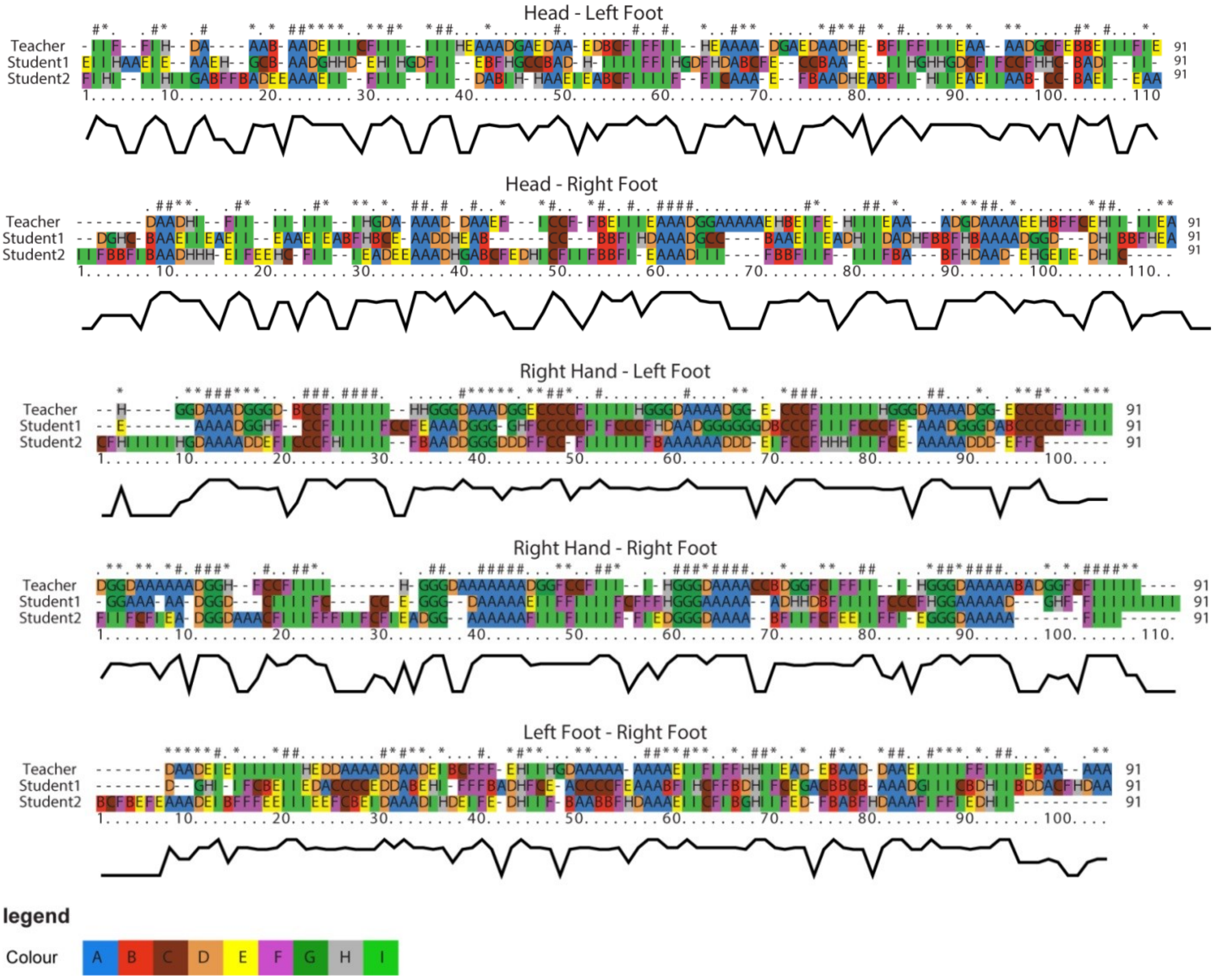


Figure 7-10: Multiple alignment of QTC_B movement sequences of dancers' body movements

Using sequence alignment, repetitive movement patterns for each dancer can be individually assessed as smaller units of the entire performance. For this purpose, the rhythm in the music is used to mark the starting and ending points of the repetitive movement patterns. In our case, the entire performance lasts 91 time units and consists of 3 complete repetitive patterns that each lasts 22 time units. Aligning these repetitive movements allows us to examine the degree of similarity between the performances of dancers across successive beats. Figure 7-11 presents the results of aligning the movement sequences for each pair of body parts in relation to musical beat.

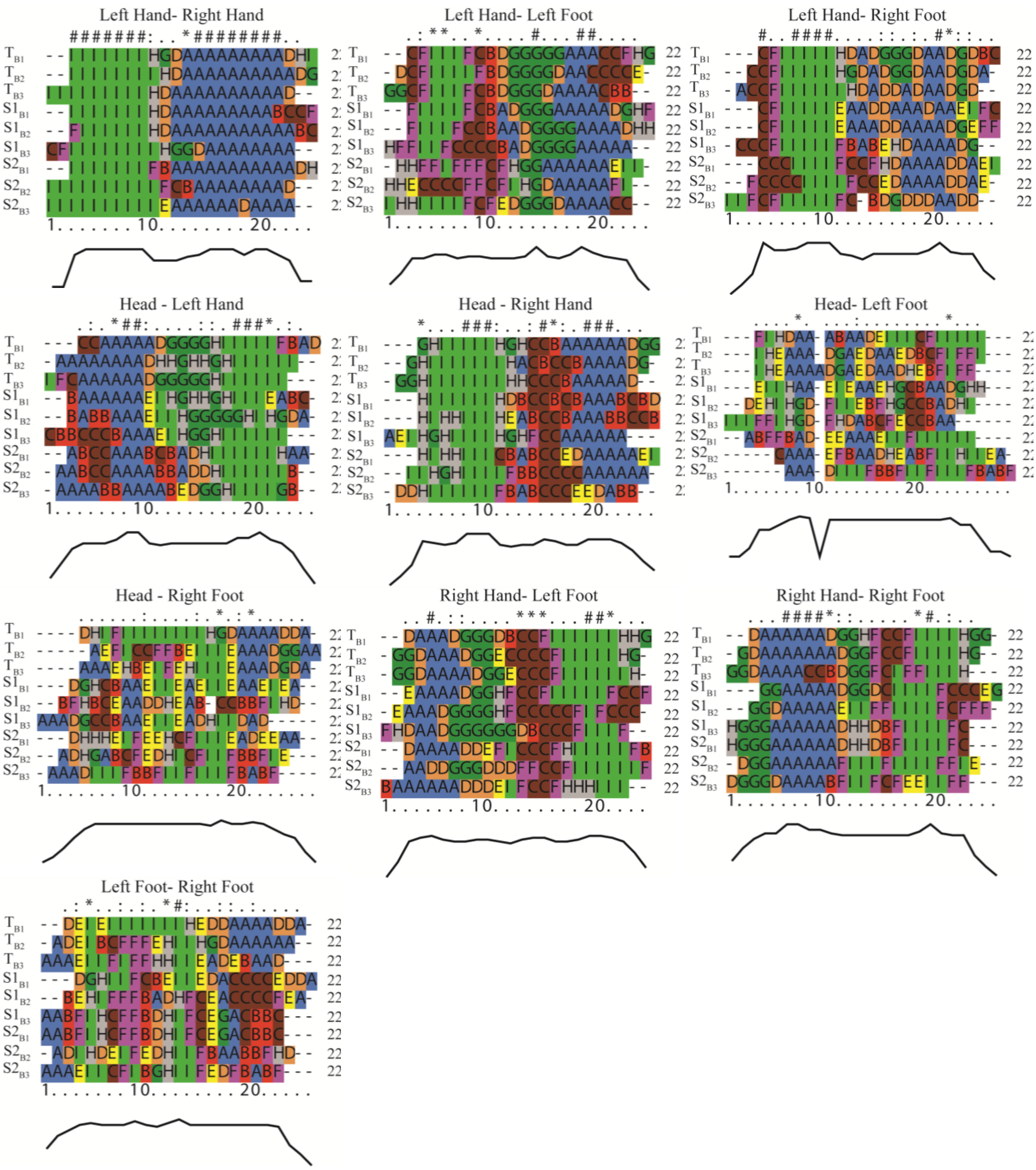


Figure 7-11: Multiple alignment of QTC_B movement sequences based on the beats of the music

In addition to visually characterising the similarities/differences in movement patterns of dancers based on the rhythm of the music (i.e. T_{B1}, T_{B2}, T_{B3}, $S1_{B1}$, $S1_{B2}$, $S1_{B3}$, $S2_{B1}$,

$S2_{B2}$, and $S2_{B3}$)[2], we further present a numerical measure based on alignment scores and represented in the form of hierarchical clusters of movement sequences.

7.4.3 Step 3: Hierarchical Clustering of Movement Sequences

Clustering enables the detection of objects that share similar properties. Clustering is typically application dependent. In this study, we attempt to cluster the dancers' movements based on the relative motions of various body parts. We use a hierarchical clustering method to build a hierarchy of clusters (i.e. the movement sequences of the teacher and the students). Based on the multiple alignment of QTC_B movement sequences shown in Figures 7-10 and 7-11, (dis)similarity matrices are constructed to cluster the sequences. The results of clustering are represented in the form of dendrograms in Figure 7-12. A dendrogram supports the determination of a typology of different movement behaviour of dancers. The results of applying SAM on real dance data suggest that certain movements were harder to follow by the students than other movements. Figure 7-12 shows the agglomerative hierarchical clustering in the form of dendrograms for the sequences as presented in Figure 7-11.

The height of the branch points shows the extent to which clusters differ from one another: the greater the height, the greater the difference. The value 0 represents the minimum distance after aligning the movement sequences, whereas 1 represents the maximum distance. As shown in the dendrograms, distances vary from one pair of body parts to another. In Figure 7-12, for example, the relative motion of the teacher's hands did not differ significantly from that of the students, as demonstrated by the relatively short distance in the left hand-right hand dendrogram. In contrast, the head-left foot dendrogram shows a significant difference between the last two beats of the teacher and the other beats. Based on this method of alignment and clustering, we observe that the performance of student 1 is better than that of student 2. Furthermore, this method allows us to identify the pairs of student body parts that more closely resembled those of the teacher. These results can assist instructors in recognising the strengths and weaknesses in their students' performance in the process of learning dance.

[2] Teacher beat 1 (T_{B1}), Teacher beat 2 (T_{B2}), Teacher beat 3 (T_{B3}), Student 1 beat 1 ($S1_{B1}$), Student 1 beat 2 ($S1_{B2}$), Student 1 beat 3 ($S1_{B3}$), Student 2 beat 1 ($S2_{B1}$), Student 2 beat 2 ($S2_{B2}$), and Student 2 beat 3 ($S2_{B3}$).

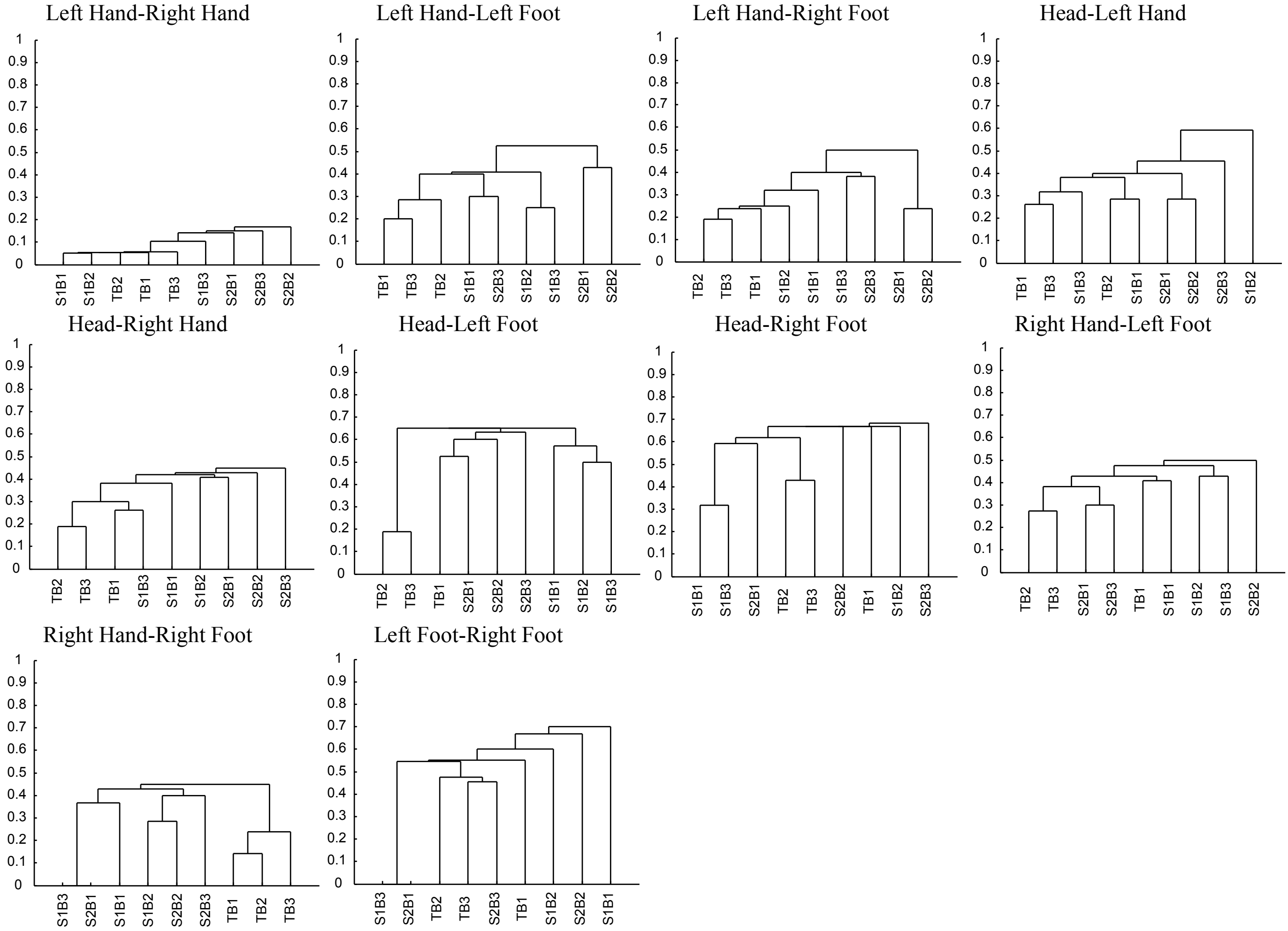


Figure 7-12: Dendrograms based on the results of alignment in Figure 7-11

7.5 Discussion

Much progress has been made regarding the theories, methodologies, and applications for analysing, modelling, and interpreting movement data. Researchers have focused on different aspects in this area, including analysing the sequential aspects within the spatial and temporal dimensions of movement data (e.g., (Delafontaine et al., 2012b; Dodge et al., 2012; Shoval & Isaacson, 2007a; Yuan & Raubal, 2014)). For example, in (Dodge et al., 2012), key parameters that characterise the movement of objects, the so-called movement parameters (MPs) such as speed, acceleration, or direction and derived from the trajectories of objects were taken into account for finding similar trajectories. They compared sequences of class labels as symbolic representation of MPs for the similarity measure. In this section, we compare our approach in this chapter with respect to two well-known techniques, namely RElative MOtion (REMO) and Dynamic Time Warping (DTW).

As a key contribution of this chapter, we addressed the applicability of the SAM approach to analyse movements of MPOs. The method is comparable to, for example, REMO and DTW. REMO is an approach that describes motion patterns by changes in the motion attributes of objects such as the speed and motion azimuth of individual MPOs over time (Laube et al., 2005). The comparison of motion attributes is performed using a two-dimensional matrix in which the horizontal dimension represents successive points in time and the vertical dimension represents different objects. The entries of the matrix correspond to motion attribute values. Consequently, the motion patterns of a group of MPOs are reflected in a single REMO matrix. The DTW is an algorithm for measuring the similarity between two time series that may vary in time or speed (Müller, 2007). Unlike traditional distance measures such as Euclidean distance, DTW can calculate the similarity between two time series that may feature some noise and displacements. We have comprehensively compared REMO with DTW and featured some of the advantages and drawbacks of both techniques with respect to the same case study in the next chapter.

Although the concept of REMO, DTW and SAM are uncomplicated and applicable to many research domains, the understanding of these techniques requires some expert knowledge. Throughout this chapter, we had this idea to highlight the usefulness of qualitative information in the analysis and reasoning movement data. QTC information built based on changing Euclidean distances between two MPOs was cross-pollinated with SAM. This point of view has not been considered in the next chapter.

Unlike DTW, both REMO and SAM can reveal interesting information about motion events retrieved from the interrelation among multiple MPOs. With this difference that QTC considers the relative motion of one object with respect to another object (i.e. relative movements) and REMO allow the identification and quantification of individual motion behaviour, events of distinct group motion behaviour, so as to relate the motion of individuals to groups (Laube et al., 2005). In the DTW approach, we may not investigate movements of multiple objects simultaneously and find such interesting patterns.

The superiority of SAM and DTW over the REMO approach is that REMO is very sensitive to noises, shifts, and distortions in movement data. Thus, drawing analogy between REMO matrices based on such data is challenging. SAM and DTW overcome this limitation and give intuitive distance measurements between time series by handling both global and local shifting of the time dimension. Another advantage of DTW and SAM is the ability to handle time series with different lengths, while this is quite challenging with REMO approach.

From the visual analysis point of view, REMO and SAM support more human's intuition in order to interpret the visual results. Therefore, the high dependency on expert knowledge can be counted as a weakness of DTW approach.

7.6 Conclusions and Future Work

To analyse the similarity in the movements of moving objects, we proposed an innovative approach in which the sequence alignment method (SAM) was used to align and assess qualitative movement sequences. The proposed methodology can be applied to any domain in which an understanding of object movement patterns is important.

In this chapter, the movements of three Samba dancers were analysed to measure the degree of (dis)similarity between the dancers' movements. Characterising similarity/dissimilarity contributes to a better understanding of how dancers move. QTC information was used to form QTC movement sequences. The results of a similarity assessment of QTC movement sequences were presented in the form of dendrograms, in which similar movement sequences were grouped in the same clusters. In other words, our strategy clustered the movement of moving objects based on their movement patterns. Contrary to most existing work that uses very detailed data for movement analysis, we were able to achieve our goals based on a summary of the movement data (i.e. QTC information describing the qualitative vision of the relative movement of MPOs). Thus, this work did not involve any of the existing geometry-based similarity analyses.

We showed that the detected and analysed patterns formed based on the discretised and qualitative movement data were informative and useful. A comprehensive study can reveal the compromised levels of discretisation and qualitativness of data without losing much valuable information.

Extra information, e.g. directional information can also be considered to identify motion patterns using other types of QTC, e.g. QTC Double-Cross (QTC_C). QTC_C provides more detail than QTC_B but increases the problem complexity.

The examples presented in this chapter were based on a relatively short time interval. In future work, we intend to apply the approach to larger trajectory data sets. Another avenue for future work will be developing a prototype that allows instructors to assess the movement of dancers interactively. We hope to report on these extensions in the near future.

References

Abbott, A. (1995). Sequence analysis: New methods for old ideas. *Annual Review of Sociology, 21* (1), 93-113.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26* (11), 832-843.

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., & De Maeyer, P. (2007). The qualitative trajectory calculus on networks. *Spatial Cognition V: Reasoning, Action, Interaction, 4387*, 20-38.

Brakatsoulas, S., Pfoser, D., Salas, R., & Wenk, C. (2005). On map-matching vehicle tracking data. *Proceedings of the 31st International Conference on Very Large Data Bases* (pp. 853-864). Trondheim, Norway.

Brodie, M., Walmsley, A., & Page, W. (2008). Fusion motion capture: A prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing. *Sports Technology, 1* (1), 17-28.

Buchin, K., Buchin, M., Van Kreveld, M., & Luo, J. (2009). Finding long and similar parts of trajectories. *Proceedings of the 17th International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS)* (pp. 296-305).

Chen, L., Ozsu, M. T., & Oria, V. (2005). Robust and fast similarity search for moving object trajectories. *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data* (pp. 491-502). New York : ACM.

Colombo, G., Facoetti, G., & Rizzi, C. (2013). Virtual testing laboratory for lower limb prosthesis. *Computer - Aided Design & Applications, 10* (4), 671-683.

DeCesare, N. J., Squires, J. R., & Kolbe, J. A. (2005). Effect of forest canopy on GPS-based movement data. *Wildlife Society Bulletin, 33* (3), 935-941.

Delafontaine, M., Bogaert, P., Cohn, A. G., Witlox, F., De Maeyer, P., & Van de Weghe, N. (2011a). Inferring additional knowledge from $QTC_{(N)}$ relations. *Information Sciences, 181* (9), 1573-1590.

Delafontaine, M., Chavoshi, S. H., Cohn, A. G., & Van de Weghe, N. (2012a). A Qualitative Trajectory Calculus to Reason about Moving Point Objects. In: S. M. Hazarika (Ed.), *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions* (pp. 147-167). IGI Global.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011b). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Delafontaine, M., Versichele, M., Neutens, T., & Van de Weghe, N. (2012b). Analysing spatiotemporal sequences in Bluetooth tracking data. *Applied Geography, 34* (2012), 659-668.

Dodge, S. (2011). *Exploring Movement Using Similarity Analysis.* University of Zurich, Zurich.

Dodge, S., Laube, P., & Weibel, R. (2012). Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science, 26* (9), 1563-1588.

Fabrikant, S. I., Rebich Hespanha, S., Andrienko, N., Andrienko, G., & Montello, D. (2008). Novel method to measure inference affordance in static small-multiple map displays representing dynamic processes. *The Cartographic Journal, 45* (3), 201-215.

Faloutsos, C., Jagadish, H. V., Mendelzon, A. O., & Milo, T. (1997). A signature technique for similarity-based queries. *Compression and Complexity of Sequences* (pp. 2-20). IEEE Computer Society.

Frank, A. U. (1996). Qualitative spatial reasoning: cardinal directions as an example. *International Journal of Geographical Information Systems, 10* (3), 269-290.

Freksa, C. (1992). Using orientation information for qualitative spatial reasoning. *Lecture Notes in Computer Science* (639), 162-178.

Frentzos, E., Gratsias, K., & Theodoridis, Y. (2007). Index-based most similar trajectory search. *Proceedings of the 23rd International Conference on Data Engineering (ICDE 2007)*(pp. 816-825).

Gagliardo, A., Ioale, P., Savini, M., Lipp, H. P., & Dell'Omo, G. (2007). Finding home: The final step of the pigeons' homing process studied with a GPS data logger. *Journal of Experimental Biology, 210* (7), 1132-1138.

Galton, A. (2001). Dominance diagrams: A tool for qualitative reasoning about continuous systems. *Fundamenta Informaticae, 46* (1-2), 55-70.

Gau, R. J., Mulders, R., Ciarniello, L. J., Heard, D. C., Chetkiewicz, C. L. B., Boyce, M., Munro, R., Stenhouse, G., Chruszcz, B., Gibeau, M. L., Milakovic, B., & Parker, K. L. (2004). Uncontrolled field performance of Televilt GPS-Simplex

(TM) collars on grizzly bears in western and northern Canada. *Wildlife Society Bulletin, 32* (3), 693-701.

Giannotti, F., & Pedreschi, D. (2008). *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery*. Springer.

Guerra-Salcedo, C. M., Janek, L., Perez-Ortega, J., & Pazos-Rangel, R. A. (2005). Predicting performance of swimmers using machine learning techniques. In: N. Callaos & W. Lesso (Eds.), *Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2005)* (Vol. 3, pp. 146-148). Orlando.

Hornsby, K., & Egenhofer, M. (2002). Modelling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence, 36* (1-2), 177-194.

Hvidberg, M. (2006). Tracking human exposure to ultrafine particles in Copenhagen using GPS. *Epidemiology, 17* (6), S38-S38.

Joh, C. H., Arentze, T., Hofman, F., & Timmermans, H. (2002). Activity pattern similarity: A multidimensional sequence alignment method. *Transportation Research Part B: Methodological, 36* (5), 385-403.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Lin, B., & Su, J. (2005). Shapes based trajectory queries for moving objects. *Proceedings of the 13th annual ACM International Workshop on Geographic Information Systems (GIS '05)* (pp. 21-30). Bremen: ACM.

Michael, K., McNamee, A., Michael, M. G., & Tootell, H. (2006). Location-based intelligence - Modelling behaviour in humans using GPS. In: R. Brody (Ed.), *Proceedings of the IEEE International Symposium on Technology and Society (ISTAS 2006)* (pp. 1-8), New York, USA.

Miller, H. J., & Han, J. (2009). *Geographic Data Mining and Knowledge Discovery* (2nd ed.). CRC Press.

Mirabella, O., Raucea, A., Fisichella, F., & Gentile, L. (2011). A motion capture system for sport training and rehabilitation. *Proceedings of the 4th International Conference on Human System Interactions (HSI)* (pp. 52-59).

Morrison, D. A. (2010). Sequence alignment: Methods, models, concepts, and strategies. *Systematic Biology, 59* (3), 363-365.

Mount, D. W. (2007). Dot Matrix Pairwise Sequence Comparison. *CSH Protoc*, pdb top31.

Müller, M. (2007). Dynamic time warping. *Information Retrieval for Music and Motion*, 69-84.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48* (3), 443-453.

Ong, R., Pinelli, F., Trasarti, R., Nanni, M., Renso, C., Rinzivillo, S., & Giannotti, F. (2011). Traffic jams detection using flock mining. In: D. Gunopulos, T. Hofmann, D. Malerba & M. Vazirgiannis (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 650-653). Berlin Heidelberg: Springer.

Orellana, D., Bregt, A. K., Ligtenberg, A., & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. *Tourism Management, 33* (3), 672-682.

Pan, X., Han, C., Dauber, K., & Law, K. (2007). A multi-agent based framework for the simulation of human and social behaviour during emergency evacuations. *AI & Society, 22* (2), 113-132.

Pelekis, N., Kopanakis, I., Marketos, G., Ntoutsi, I., Andrienko, G., & Theodoridis, Y. (2007). Similarity search in trajectory databases. *Proceedings of the 14th International Symposium on Temporal Representation and Reasoning (TIME 2007)* (pp. 129-140).

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic-based on regions and connection. In B. Nebel, C. Rich & W. R. Swartout (Eds.), *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR 92)* (pp. 165-176).

Rosenberg, M. S. (2009). *Sequence Alignment: Methods, Models, Concepts, and Strategies*. University of California Press, Berkeley.

Shoval, N., & Isaacson, M. (2007a). Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers, 97* (2), 282-297.

Shoval, N., & Isaacson, M. (2007b). Tracking tourists in the digital age. *Annals of Tourism Research, 34* (1), 141-159.

Sinha, G., & Mark, D. M. (2005). Measuring similarity between geospatial lifelines in studies of environmental health. *Journal of Geographical Systems, 7* (1), 115-136.

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Van de Weghe, N., Bogaert, P., Cohn, A., Delafontaine, M., De Temmerman, L., Neutens, T., De Maeyer, P., & Witlox, F. (2007). How to handle incomplete knowledge concerning moving objects. In: B. Gottfried (Ed.), *Proceedings of the Workshop on Behaviour Monitoring and Interpretation (BMI 2007)* (pp. 91-101).

Van de Weghe, N., Cohn, A. G., De Tre, G., & De Maeyer, P. (2006). A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems. *Control and Cybernetics, 35* (1), 97-119.

Van de Weghe, N., Cohn, A. G., Maeyer, P. D., & Witlox, F. (2005a). Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems with Applications, 29* (4), 977-983.

Van de Weghe, N., & De Maeyer, P. (2005). Conceptual neighbourhood diagrams for representing moving objects. *Perspectives in Conceptual Modelling, 3770*, 228-238.

Van de Weghe, N., De Tré, G., Kuiijpers, B., & De Maeyer, P. (2005b). The double-cross and the generalization concept as a basis for representing and comparing shapes of polylines. *Lecture Notes in Computer Science, 3762*, 1087-1096.

Van de Weghe, N., Maddens, R., Bogaert, P., Brondeel, M., & De Maeyer, P. (2004). Qualitative analysis of polygon shape-change. *Proceeding of International Geoscience and Remote Sensing Symposium IGARSS '04* (Vol. 6, pp. 4157-4159).

Van Kreveld, M., & Luo, J. (2007). The definition and computation of trajectory and subtrajectory similarity. *Proceedings of the 15th annual ACM International Symposium on Advances in Geographic Information Systems (GIS '07)* (pp. 1-4). New York: ACM.

Van Shaick, J., & van der Spek, S. C. (2008). *Urbanism on Track: Application of Tracking Technologies in Urbanism* (Vol. 1). IOS Press.

Versichele, M., Neutens, T., Delafontaine, M., & Van de Weghe, N. (2012a). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography, 32* (2), 208-220.

Versichele, M., Neutens, T., Goudeseune, S., Bossche, F., & Weghe, N. (2012b). Mobile mapping of sporting event spectators using Bluetooth sensors: Tour of Flanders 2011. *Sensors, 12* (10), 14196-14213.

Vlachos, M., Kollios, G., & Gunopulos, D. (2002). Discovering similar multidimensional trajectories. *Proceedings of the 18th International Conference on Data Engineering* (pp. 673-684). San Jose: IEEE.

Wang, L., Hu, W., & Tan, T. (2003). Recent developments in human motion analysis. *Pattern Recognition, 36* (3), 585-601.

Wilson, C. (2001). Activity patterns of canadian women: Application of ClustalG sequence alignment software. *Transportation Research Record: Journal of the Transportation Research Board, 1777* (1), 55-67.

Wilson, C. (2006). Reliability of sequence-alignment analysis of social processes: Monte Carlo tests of ClustalG software. *Environment and Planning A, 38* (1), 187-204.

Wilson, C. (2008). Activity patterns in space and time: Calculating representative Hagerstrand trajectories. *Transportation, 35* (4), 485-499.

Yanagisawa, Y., Akahani, J.-i., & Satoh, T. (2003). Shape-based similarity query for trajectory of mobile objects. In: M.-S. Chen, P. Chrysanthis, M. Sloman & A. Zaslavsky (Eds.), *Mobile Data Management* (pp. 63-77). Berlin Heidelberg. Springer.

Yuan, Y., & Raubal, M. (2014). Measuring similarity of mobile phone user trajectories- A spatio-temporal edit distance method. *International Journal of Geographical Information Sciences, 28* (3), 496-520.

8

Relative Motion & Dynamic Time Warping

We are drowning in information and starving for knowledge. Rutherford D. Roger

Modified From: Seyed Hossein Chavoshi, Bernard De Baets, Tijs Neutens, Hyowon Ban, Ola Ahlqvist, Guy De Tré, and Nico Van de Weghe (2014). Knowledge Discovery in Choreographic Data Using Relative Motion Matrices and Dynamic Time Warping. Applied Geography, 47 (2014), pp. 111-124

8 RELATIVE MOTION & DYNAMIC TIME WARPING

Abstract: Recent advances in location-aware technologies have led to the exploitation of geospatial methods to uncover valuable information from large movement data sets. The focus of this chapter is on choreographic information. In particular, the goal is to visualise and analyse the motion patterns of Samba dancers during their performance by means of two complementary methods. The first method performs map algebra with *RElative MOtion* (REMO) matrices to study the evolution of motion attributes, such as speed, motion azimuth, and vertical angle over time. The second method applies dynamic time warping (DTW) to time series of motion attributes. The results demonstrate that both methods are useful in numerically comparing the performance of Samba dancers and visually exploring the motion patterns of different body parts.

8.1 Introduction

Motion is the change in position of objects, such as vehicles, animals, hurricanes or oil spills, with respect to time. Recent technological progress in location-aware technologies, including Bluetooth sensors, RFID[3], and GPS[4], have made it possible to track changes in the positions of such objects and gather enormous amounts of movement data, often as a series of discrete observations of moving objects represented as tuples of *x*, *y*, *z* coordinates and time *t*. This enormous amount of positional data has led researchers from different disciplines to develop new ways to explore movement data. Particular attention has been directed towards analysing human movement. For example, Sigal et al. (2010) developed a hardware system to capture synchronised video and ground-truth 3D motion. Chaudhry et al. (2009) presented an activity recognition method that classifies the human activities in video sequences. Nagashima et al. (2012) demonstrated a method for analysing the principal components of human motion in time series and estimating the functional mode of the human motion. Yuan and Raubal (2012) developed a technique for analysing dynamic mobility patterns of mobile phone datasets using Dynamic Time Warping (DTW). Observations of human motion may be related to individual or collective motion behaviour (Andrienko et al., 2008a), and human motion has been studied in various research domains, such as social

[3] Radio Frequency Identification
[4] Global Positioning System

sciences, geomarketing, transportation, political sciences, biomedical analysis, and sports and recreation.

Due to the large amount of movement data that is becoming available through mobile sensors, data mining methods are necessary to uncover useful, hidden information that can be used for different purposes. In this study, we aim at understanding the motion patterns of moving point objects (MPOs) through similarity analysis. Because motion patterns are clearly visible in many rhythmic dances, we used some basic movements of Samba as a case study to investigate two approaches to measure similarity in motion patterns, namely *RElative MOtion* (REMO) and *Dynamic Time Warping* (DTW).

The concept of REMO was introduced in (Laube & Imfeld, 2002; Laube et al., 2005). In short, REMO takes different motion attributes such as the speed and motion azimuth of individual MPOs into account to be compared over space and time. This comparison is performed using a two-dimensional matrix in which the horizontal dimension represents successive points in time and the vertical dimension represents different objects. The entries of the matrix correspond to motion attribute values. Consequently, the motion patterns of a group of MPOs are reflected in a single REMO matrix. Dynamic Time Warping (DTW), on the other hand, is an algorithm for measuring the similarity between two sequences that may vary in time or speed.

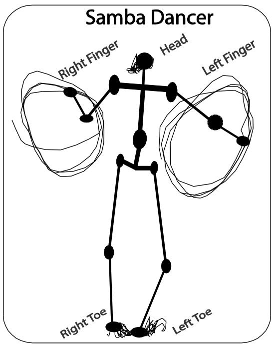
This chapter will apply and refine both methods using dance data. In the next section, we briefly provide a background of the basic concepts of REMO and DTW. In Section 8.3, we apply both approaches to evaluate the performance of novice Samba dancers with respect to that of their teacher. In Section 8-4, we discuss our findings. Finally, the conclusions of this study and suggestions for future work are presented in Section 8.5.

8.2 Background

8.2.1 RElative MOtion (REMO)

Our research seeks to contribute to the exploration of the motion patterns of multiple objects. One relevant concept in this respect is REMO (Laube & Imfeld, 2002; Laube et al., 2005), a method that describes motion patterns by changes in the motion attributes of objects (e.g., change in speed or motion azimuth over time). The REMO representation transforms trajectories of MPOs into a matrix that allows for the matching of motion patterns (Laube et al., 2005). In this matrix, a row represents the motion attribute values of an object over time and a column represents time stamps of the movement. Figure 8-1 shows an example of a REMO matrix based on a part of the Samba dance dataset used throughout this chapter. Here, the movements of five body parts of a dancer over a time interval lasting 13 * 0.04 s form a REMO matrix in which the values of the cells represent the speed of each body part at each time stamp.

According to Laube & Imfeld (2002), the purpose of forming matrices of motion attributes (i.e. REMO matrices) is to recognise interrelations between the motion of groups of point objects. There are various types of motion patterns that can be detected in REMO matrices, such as those that occur over time, across objects, and the combination of both (for further details, see (Laube & Imfeld, 2002)). To better understand the aforementioned motion patterns, we have highlighted some of them in Figure 8-1. For example, the left toe of the Samba dancer was moving with a constant speed of 0.7 m/s over the interval t_7 to t_9, demonstrating *constancy*. In contrast, there is a motion pattern across all body parts of the dancer revealing that they had an identical speed at t_{11}.



Speed (m/s)

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Head	0.1	0.2	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.3	0.2
Right Finger	0.2	0.2	0.3	0.2	0.2	0.3	0.2	0.2	0.1	0.2	0.2	0.3	0.2
Left Finger	0.5	0.2	0.2	0.1	0.2	0.2	0.4	0.2	0.2	0.1	0.2	0.3	0.4
Right Toe	0.5	0.5	0.3	0.3	0.3	0.3	0.2	0.1	0.1	0.5	0.2	0.7	0.7
Left Toe	0.3	0.5	0.2	0.2	0.5	0.5	0.7	0.7	0.7	0.3	0.2	0.3	0.3

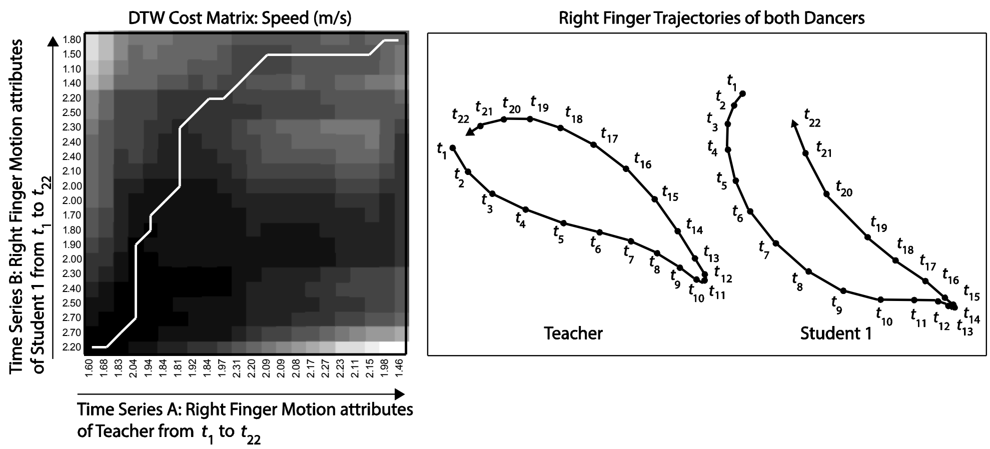
Figure 8-1: REMO matrix based on the movements of five body parts of a Samba dancer over a given time interval

In this chapter, the REMO matrices of an object's motion are considered to be the basis for all further analyses. Instead of categorising motion attributes into discrete classes, as the originally conceived in REMO (Laube & Imfeld, 2002), the original values of motion attributes are retained. A REMO matrix thus represents the performance of a dancer over a time interval of movement using a set number of descriptive attributes for each time stamp.

8.2.2 Dynamic Time Warping (DTW)

A traditional comparison of two time series reveals whether the two time series are similar (Das et al., 1997; Morse & Patel, 2007). Generally, a traditional distance measure such as Euclidean distance can be used to quantify the difference between two time series. However, this measure is not the best choice for comparing time series that may feature some noise and displacements. To overcome this problem, many algorithms have been proposed to measure the similarity between two time series, among which Dynamic Time Warping (DTW) has attracted the attention of many researchers from different disciplines, including speech recognition (Sakoe & Chiba, 1978), pattern recognition (Oveneke et al., 2012; Yuan & Raubal, 2012), handwriting recognition (Qiao & Yasuhara, 2006), and music retrieval (Lijffijt et al., 2010). DTW can calculate the similarity between two time series based on finding an optimal match between them even if they are not identical in size. An illustrative example is when speech recognition tries to compare two recordings of the same phrase and one person talks slower than the other or uses varying pauses between words.

One of the common challenges in analysing dance is a lack of clear synchronisation between dancers' movements over a specific time interval. The main advantage of DTW is that one can obtain a robust measurement from the comparison of two time series even if they are partially out of sync.



(a) (b)

Figure 8-2: DTW cost matrix of the right finger time series of teacher and student 1 (a), and the corresponding trajectories (b)

Figure 8-2 illustrates the procedure of calculating the DTW distance between two given time series of a Samba dance teacher and one of his students, student 1, following Salvador and Chan (2007). Consider two time series *A* and *B* of length 22:

$$A = [a_1, a_2, ..., a_i, ..., a_{22}]$$
$$B = [b_1, b_2, ..., b_j, ..., b_{22}]$$

To determine the DTW distance, the time series are warped non-linearly in the time dimension. A warp path *W* is denoted as follows:

$$W = [w_1, w_2, ..., w_K], \text{ with } \max(|A|,|B|) \le K < |A| + |B|,$$

where *K* is the length of the warp path, the k^{th} element of the warp path is $w_k = (i_k, j_k)$, i_k is an index of time series *A*, and j_k is an index of time series *B*. The warp path starts at the bottom-left corner of Figure 8-2a (i.e. $w_1 = (1,1)$) and finishes at the top-right corner of the matrix (i.e. $w_K = (22,22)$). There is also a constraint on the warp path that forces i_k and j_k to be increasing in the warp path as indicated below.

$$w_k = (i_k, j_k), w_{k+1} = (i_{k+1}, j_{k+1}), \text{ with } i_k \le i_{k+1} \le i_k + 1, j_k \le j_{k+1} \le j_k + 1,$$

The optimal warp path is the minimum distance warp path (MDWP). The distance of a warp path is calculated based on the following equation:

$$Dist(A,B) = \sum_{k=1}^{K} d(a_{i_k}, b_{j_k}), \qquad \text{Eq. 8-1}$$

where $d(a_{i_k}, b_{j_k})$ is the distance between two entities of the time series *A* and *B*. In this study, three motion attributes, namely, speed, motion azimuth, and vertical angle, form time series for each individual body part. For those time series with speed values, $d(a_i, b_j)$ in the DTW procedure is calculated based on the Euclidean distance. Because motion azimuth and vertical angle are both angular data, the following equation is used to calculate the distance to overcome the circularity problem in the angular data (for a detailed explanation, see (Fisher, 1995)):

$$d(a_i, b_j) = 1 - \cos(a_i - b_j) / 2, \qquad \text{Eq. 8-2}$$

Dynamic programming is employed to find the MDWP from the beginning to the $(i, j)^{th}$ cell. To find the minimum distance warp, every cell should be filled in the cost matrix. There is a constraint to find the MDWP: the warp path should either increase by one or remain the same along the i and j axes. Therefore, the value of each cell in the cost matrix is as follows (for a detailed explanation, see (Salvador & Chan, 2007)):

$$C(i,j) = Dist(i,j) + \min(C(i-1,j), C(i,j-1), C(i-1,j-1)), \qquad \text{Eq. 8-3}$$

Figure 8-2 shows an example of a cost matrix and the MDWP traced through it from $C(1, 1)$ to $C(22, 22)$. The warp path highlighted in the cost matrix in Figure 8-2a (i.e. DTW distance between the two series) represents the minimum distance required to find the best match between these two sequences. The regions of low cost are indicated by dark colours and regions of high cost are indicated by light colours in the cost matrix.

8.3 Analysing Motion Patterns of Samba Dance

8.3.1 Samba Dataset

The Samba dance data used in this chapter was obtained by using a motion capture (MoCap) system that records the positions of objects over time via reflective markers attached to the objects in combination with infrared cameras. Because in the present study we seek to examine the motion patterns of Samba dancers, only a very basic dataset including the 3D dimensional coordinates of five body parts of each Samba dancer is considered. The data include basic movements of Samba performed by three dancers: one teacher and two students. The data were generated by the Institute for Psychoacoustics and Electronic Music (IPEM), Department of Musicology at Ghent University, Ghent, Belgium. At the beginning, the movements of 27 body parts of each dancer were collected using a motion capture (MoCap) system, but data for only five body parts, namely, the head (H), right finger (RF), left finger (LF), right toe (RT), and left toe (LT), were used in this study. In the MoCap system, the movements of one or more objects can be sampled many times per second. Infrared markers attached to parts of a dancer's body were tracked via a number of infrared cameras. The positional information of the body parts was logged in a local coordinate system. The five body parts were selected to recognise regular and cyclic movements of the Samba dance.

Samba is an old Brazilian style of dance that is lively and rhythmical. The characteristic motion of Samba is bouncing. This is a gentle, rhythmic action felt through the knees and ankles. This bouncing action is quite difficult to master, but is essential to the overall character of the Samba dance. Generally, the Samba dance has a quick beat that requires fast footwork. Moreover, regularity in the motion patterns of the hands and feet can be detected from the REMO matrices discussed in the next subsection. The data feature a time interval of 3.68 s and a total of 92 time stamps with an equal interval of 0.04 s. The Samba dancers were supposed to perform exactly the same movements over any given time interval. However, there may have been some slight differences between the movements of the teacher and those of the students. The aim of this chapter is to explore these differences through REMO and DTW analyses.

As an example, Figure 8-3a illustrates some frames of the movements of the teacher and student 1. The data in the illustrated frames were created based on information gathered from the 27 markers. Figure 8-3b demonstrates the 3D movements of the three dancers by visualising all point data along three orthogonal directions: front view, left, and top view. The blue dots represent the point locations of the teacher's head, left fingers, right fingers, left foot, and right foot. The green dots represent those of student 1 and the pink dots those of student 2. The overlayed dots in the three colours show how the dancers were similar or different in general during the dance. For example, locations of student 2's arms tend to move farther from those of teacher than student 1's arms.

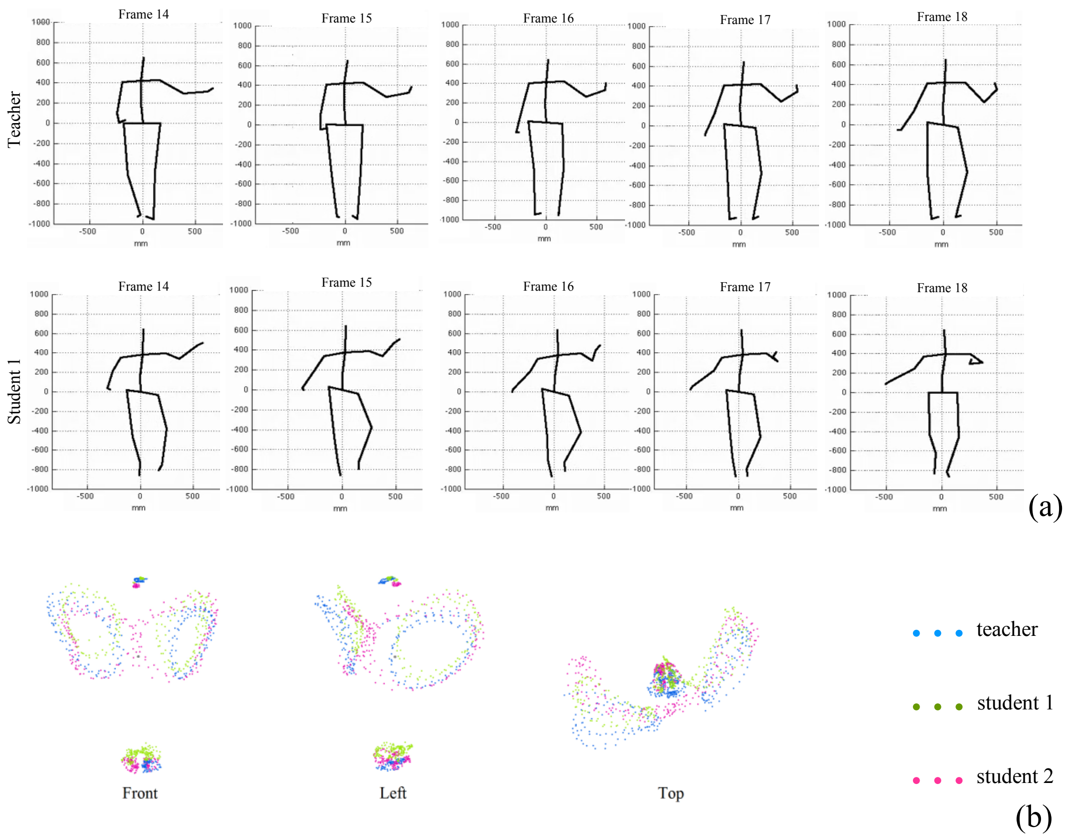


Figure 8-3: Five frames of the movements of the teacher and student 1 (frames 14-18; a), and 3D visualisation of the dancers' movements for the entire dataset (b)

8.3.2 REMO

A REMO analysis can compare the motion attributes of different objects at different time stamps (Laube & Imfeld, 2002). We consider three basic motion attributes, speed, motion azimuth, and vertical angle, to populate the REMO matrices in this study. Each REMO matrix thus consists of 92 columns (time stamps) and five rows (the five body parts). The motion attributes were measured based on the absolute *x*, *y*, and *z* coordinates of temporally consecutive points. The motion attribute *speed* was measured based on the Euclidean distance between two consecutive points. As defined in (Laube & Imfeld, 2002), *motion azimuth* represents the direction of movement in a 2D plane ranging between 0° and 360°. *Vertical angle* represents the direction of movement in a vertical dimension ranging between -90° and 90°. Unlike a crisp categorisation of motion attribute values, such as that performed in (Laube & Imfeld, 2002), which may add some imprecision to the outputs, we analyse and visualise REMO matrices based on the original values of these attributes.

The performance of the teacher is considered the benchmark for comparison with the two students. Figures 8-4, 8-5, and 8-6 illustrate the REMO matrices of the teacher and the students for all three motion attributes. Regular and cyclic motion patterns can be detected in the figures. As mentioned previously, there are more movements in the feet of dancers than the hands due to the nature of this type of dance.

Figure 8-4: REMO matrices of teacher, student 1, and student 2 for speed attribute

Figure 8-4 shows variations in the speed of the dancers' body parts. The figure clearly shows how the hands of student 2 at some time during his performance moved more rapidly than those of the teacher and student 1. We can also see that the dancers did not move their heads (H) much. In addition, we can identify certain repeated motion patterns in the movements of the body parts. For example, the speed of the teacher's right finger (RF) and left finger (LF) were alternately high and low every 10 time stamps (i.e. 10 * 0.04 s).

The REMO matrices of motion azimuth and vertical angle show changes in the direction of movements horizontally and vertically. For example, Figure 8-5 shows that the Samba dancers performed certain movements regularly. However, it also reveals that the motion patterns of students 1 and 2 were not as regular as those of the teacher. The REMO matrix in Figure 8-6 shows that the teacher moved his hands up and down regularly. The movements of the right/left toes captured in Figure 8-6 indicate regular patterns with faster variations than those of the hands. Clearly, one of the advantages of REMO is that it visualises the motion patterns of all objects at once. In the following subsection, we will discuss in further detail how we can measure similarity between the REMO matrices of different dancers/objects.

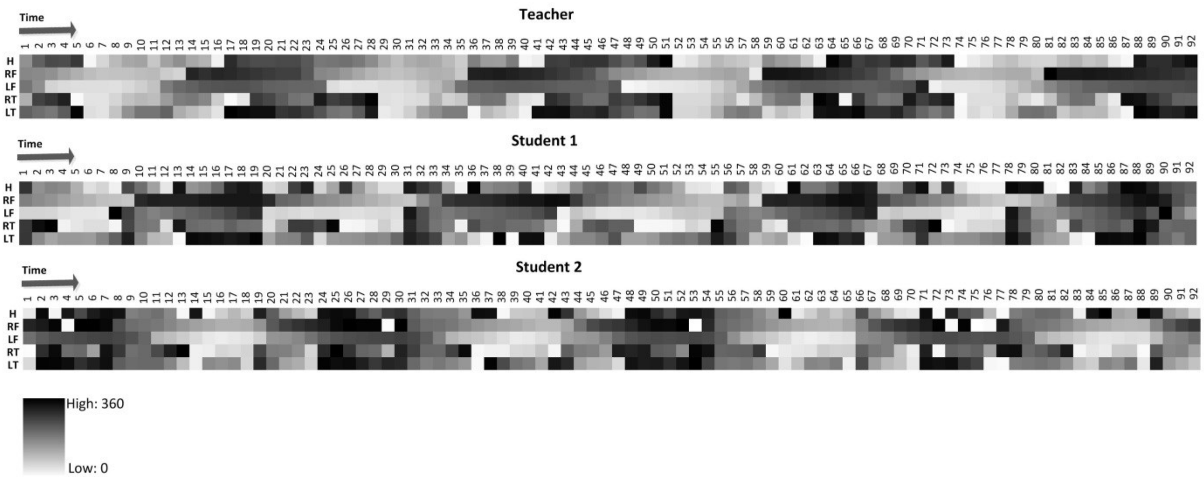


Figure 8-5: REMO matrices of teacher, student 1, and student 2 for motion azimuth attribute

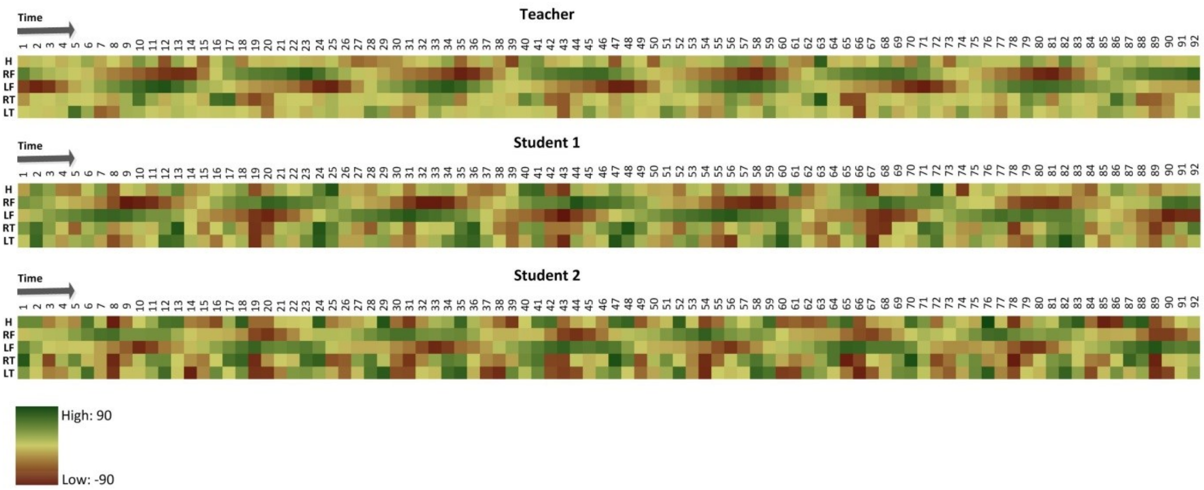


Figure 8-6: REMO matrices of teacher, student 1, and student 2 for vertical angle attribute

8.3.2.1 Map Algebra for Analysing Relative Motion of Dancers

The Samba dancers were supposed to perform analogous movements over a given time interval, with the teacher's performance as a benchmark. To examine the degree of similarity between the REMO matrices discussed in the previous section, we use an existing approach from Geographic Information Science (GIScience), namely, *map algebra* (Tomlin, 1990). In this section, we apply a simple raster subtraction operation to the two REMO matrices. Figure 8-7 shows the results of this operation, where a cell with a negative value in the resulting matrix indicates that the speed of the body part of Samba dancer 1 was lower than that of Samba dancer 2 at the same time stamp.

Samba Dancer 1

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Head	0.1	0.1	0.2	0.3	0.2	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1
Right Finger	0.2	0.2	0.3	0.2	0.2	0.3	0.2	0.2	0.1	0.2	0.2	0.3	0.2
Left Finger	0.5	0.2	0.2	0.1	0.2	0.2	0.4	0.2	0.2	0.1	0.2	0.3	0.4
Right Toe	0.5	0.5	0.3	0.3	0.3	0.3	0.2	0.1	0.1	0.5	0.2	0.7	0.7
Right Toe	0.3	0.5	0.2	0.2	0.5	0.5	0.7	0.7	0.7	0.3	0.2	0.3	0.3

Speed (m/s)

Samba Dancer 2

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Head	0.1	0.1	0.2	0.1	0.1	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1
Right Finger	0.5	0.3	0.3	0.1	0.1	0.7	0.7	0.7	0.4	0.3	0.6	0.7	0.6
Left Finger	0.3	0.5	0.2	0.1	0.1	0.4	0.2	0.2	0.5	0.6	0.7	0.1	0.5
Right Toe	0.5	0.4	0.3	0.2	0.3	0.1	0.5	0.3	0.2	0.6	0.7	0.2	0.4
Right Toe	0.2	0.5	0.2	0.1	0.4	0.5	0.2	0.1	0.5	0.4	0.3	0.2	0.1

Speed (m/s)

Samba Dancer 1 - Samba Dancer 2

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Head	0	0	0	0.2	0.1	0	-0.1	-0.1	-0.1	-0.1	0	0	0
Right Finger	-0.3	-0.1	0	0.1	0.1	-0.4	-0.5	-0.5	-0.3	-0.1	-0.4	-0.4	-0.4
Left Finger	0.2	-0.3	0	0	0.1	-0.2	0.2	0	-0.3	-0.5	-0.5	0.2	-0.1
Right Toe	0	0.1	0	0.1	0	0.2	-0.3	-0.2	-0.1	-0.1	-0.5	0.5	0.3
Right Toe	0.1	0	0	0.1	0.1	0	0.5	0.7	0.2	-0.1	-0.1	0.1	0.2

Speed (m/s)

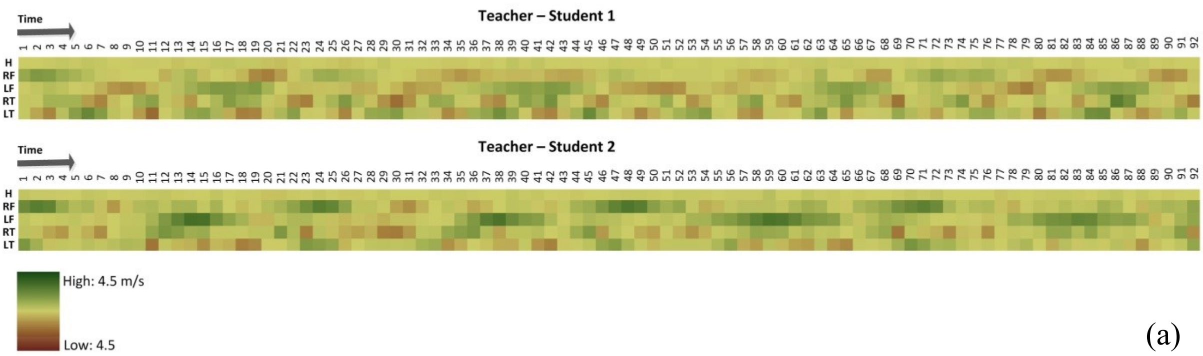
Figure 8-7: Schematic map algebra operation (subtraction) applied to the REMO matrices of two Samba dancers

Clearly, this approach is very sensitive to temporal shifting; hence, the movement data should be first synchronised before applying the map algebra operation. As shown in Figure 8-5, there exist some lags and leads in the performances of students 1 and 2 with respect to the performance of the teacher. To synchronise the lags and leads in the data, we measure the correlation coefficient of the movements of the three dancers. A correlation coefficient can measure the strength and direction of linear relationships between two time series variables. After synchronising the data based on the correlation coefficients, we draw an analogy between the REMO matrices of the teacher and students in the next subsection.

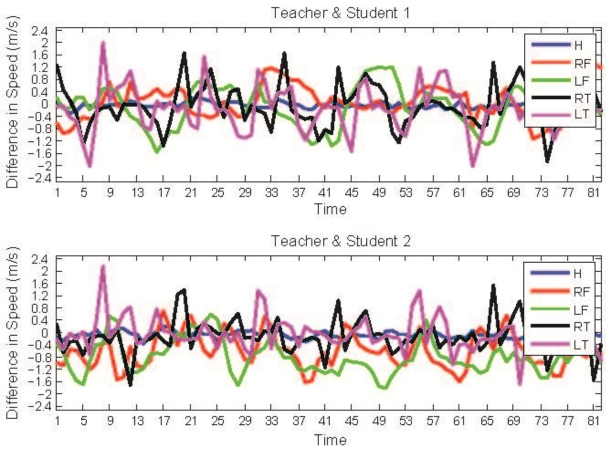
8.3.2.2 Results of Comparing REMO Matrices Using Map Algebra

Ideally, the results of the subtraction between the REMO matrices of the teacher and the students should be minimal (i.e. zero) at every time stamp (i.e. each column in the REMO matrix) and for each body part (i.e. each row in the REMO matrix), though this rarely occurs, particularly for beginners. Figure 8-8 depicts the results of the subtraction of the REMO matrices of the teacher from those of the students in the form of matrices (Figure 8-8a) and graphs (Figure 8-8b) with speed as the motion attribute.

Based on the REMO matrices and the colour bar shown in Figure 8-8a, one may recognise that differences between the movements of the teacher and students are most of the time insignificant, whereas at some moments in time, the differences are noticeable. In Figure 8-8b, we also show the results of the comparison in the form of graphs. Several fluctuations can be observed in Figure 8-8b, which might be because either the students did not perform similar to the teacher or there were slight shifts in the timing of their movements relative to those of the teacher. In Figure 8-8, positive values indicate that the speed of the students' movements was lower than those of the teacher and negative values indicate that the speed of the students' movements was higher than those of the teacher.



(a)



(b)

Figure 8-8: Results of the subtraction of REMO matrices of the teacher from those of the students with speed as the motion attribute in the form of matrices (a), and graphs (b)

Whereas simple subtraction functions well in comparing movement speed, the motion azimuth and vertical angle require a slight modification due to their respective measurement scales. For example, the difference between the two motion azimuth values 5º and 355º is ±350º using a simple operator subtraction, whereas the real absolute difference is 10º. In addition to angles, orientations, and rotations, this issue is also present in other data types such as temporal cycles, e.g., days, weeks, months, and years (Fisher, 1995). To overcome this problem, Eq. 8-2 in Subsection 8.2.2 is applied to the REMO matrices of motion azimuth and vertical angle. In this equation, a_i and b_i are the cell values of the first and second REMO matrices, respectively, and the results of subtraction thus range between 0 and 1. The difference is 0 when the values of the cells between the matrices are identical and 1 when there are 180º differences between the values of the cells. In other words, when two body parts are moving in the opposite direction, the distance is maximum at 1.

Figure 8-9 demonstrates the results of subtracting the teacher's matrices from the students' matrices based on Eq. 8-2 for both motion azimuth and vertical angle. From the bar graphs, we observe high frequencies when the difference in the directions of body part motion is 0º (i.e. same direction) or 180º (i.e. opposite direction). More information regarding the movements of each individual body part can be extracted from the bar graph. For example, in Figure 8-9c, there exists a considerable difference of movement in the left toe (LF) between student 1 and the teacher.

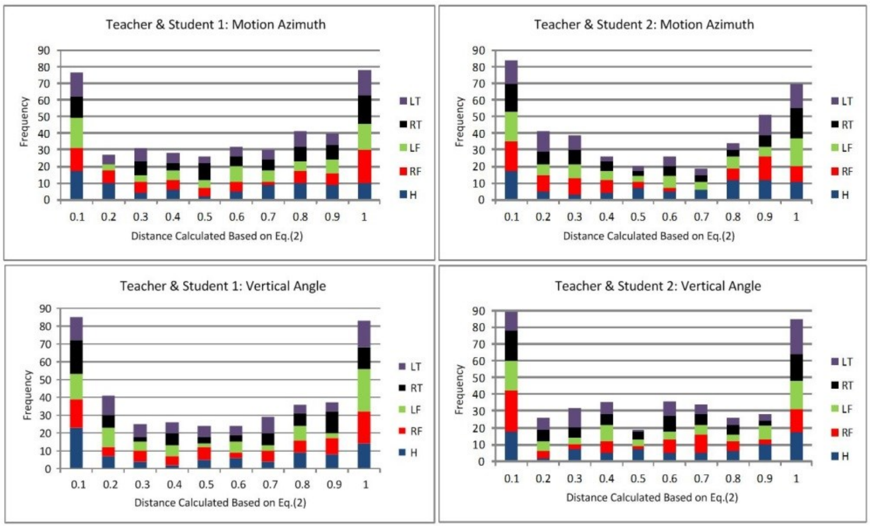
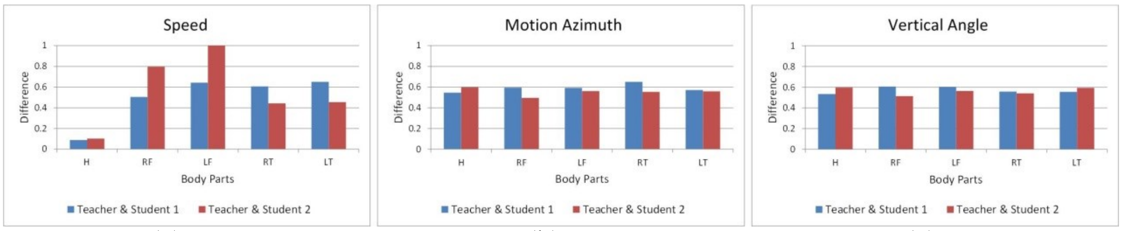


Figure 8-9: Distribution of distances between the REMO matrices of the teacher and students for motion azimuth and vertical angle, calculated based on Eq. 8-2

Figure 8-10 demonstrates a sample assessment of the performances of the Samba dancers measured based on the normalised absolute sum of differences over all time stamps. Figure 8-10a shows that the speeds of the heads of the students were similar to the speed of the teacher's head. The speed of the hands of student 1 is more similar to that of the teacher's hands than that of the hands of student 2. This result can also be observed in the REMO matrices of the students in Figure 8-4. However, the speed of the lower body parts of student 2 (i.e. feet) was more similar to that of the teacher than that of student 1. From a directional point of view, both horizontally and vertically, the students performed quite similarly to each other. For instance, student 2 succeeded in moving his right finger (RF) horizontally and vertically better than student 1. Moreover, student 2 moved his left toe (LF) vertically in a manner more similar to that of the teacher than did student 1(Figure 8-4c).



(a) (b) (c)

Figure 8-10: A sample assessment of the performances of the students

The synchronised data of the entire dance consist of three sets of repeated movement patterns with 22 time stamps each. We evaluate the performances of the students for each set, called "beats" of music: beat 1 ($B_{1;}$ Figure 8-11a, d, and g), beat 2 ($B_{2;}$

Figure 8-11b, e, and h), and beat 3 (B_3; Figure 8-11c, f, and i). We create REMO matrices of B_1, B_2, and B_3 and apply the subtraction operation to them. Figure 8-11 presents the results of the comparison of the teacher and students for all three beats and all three motion attributes. The results show which body parts of the students were moving correctly with respect to the teacher's movements.

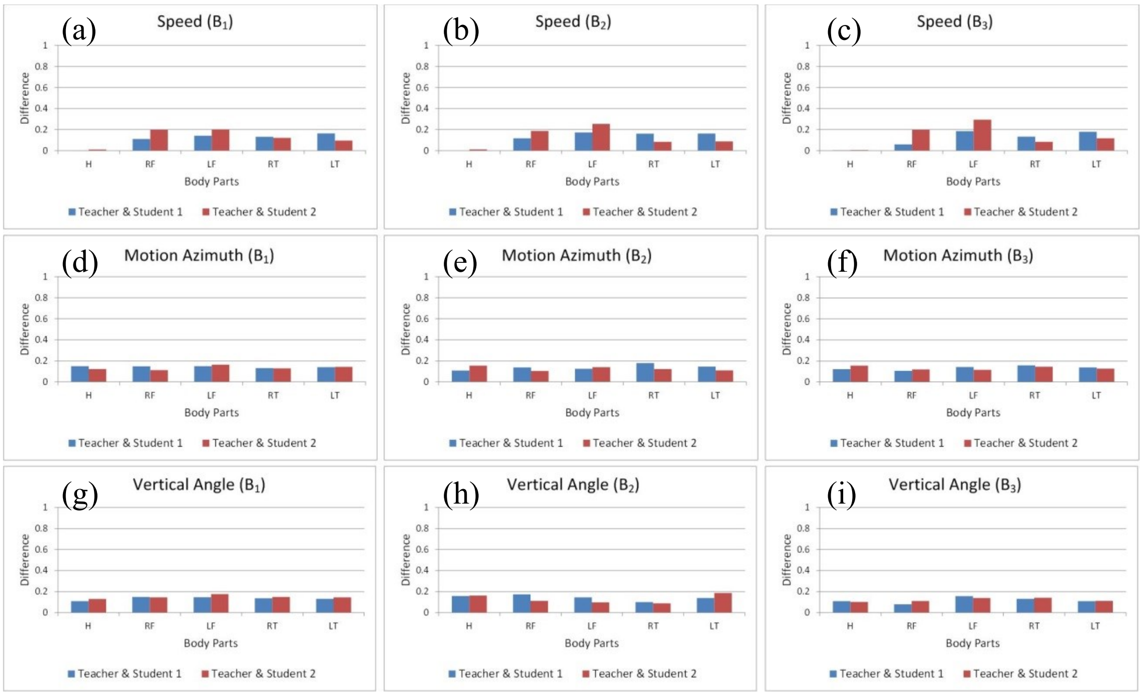


Figure 8-11: Assessment of performances of the students at each beat separately for different motion attributes

As shown in Figure 8-11a, b, and c, student 2 could not match the movements of his left finger (LF) to those of the teacher, whereas student 1 succeeded in moving his right finger (RF) in a manner highly consistent with that of his teacher. Moreover, student 2 showed that he could control the speed of his feet over time, whereas student 1 did not move his right toe (RT) in a manner similar to that of the teacher in the second beat of the dance. Generally, both students showed a similar performance in following the direction of movement of the teacher. As mentioned previously, the method developed in this study is very sensitive to the degree of temporal shifting in the original data. Thus, REMO is mostly suitable in detecting deviations in movement according to the rhythm of the dance, i.e. the degree to which two objects move synchronously.

Analogous situations in geography can be studied in the analysis of animal movements and their individual behavioural responses linked to environmental variables, weather, and ecosystem. For example, Laube and Imfeld (2002) tried to find behavioural patterns in animal observation data. Interesting observations on seasonal range and migration patterns of Porcupine Caribous were obtained from the investigation of the REMO matrices based on GPS observations. In (Laube et al., 2005), an example was given on how to use REMO in order to analyse the behaviour of players during a soccer game.

If we are more interested in the timing than motion pattern, REMO would not provide the best solution. To address such situations, we apply the DTW approach to the same dataset to eliminate the effects of shifting as much as possible in assessing the performances of the dancers.

8.3.3 DTW

In Subsection 8.2.2, the DTW approach was explained in detail. Generally, the approach requires the appropriate data type (i.e. time series data) and expert knowledge of the data. In this subsection, we use DTW to investigate the overall performance of dancers. First, we begin with the time series of speed for each individual body part over the entire dance. Figure 8-12 illustrates DTW cost matrices comparing the teacher with students 1 and 2 for all five individual body parts.

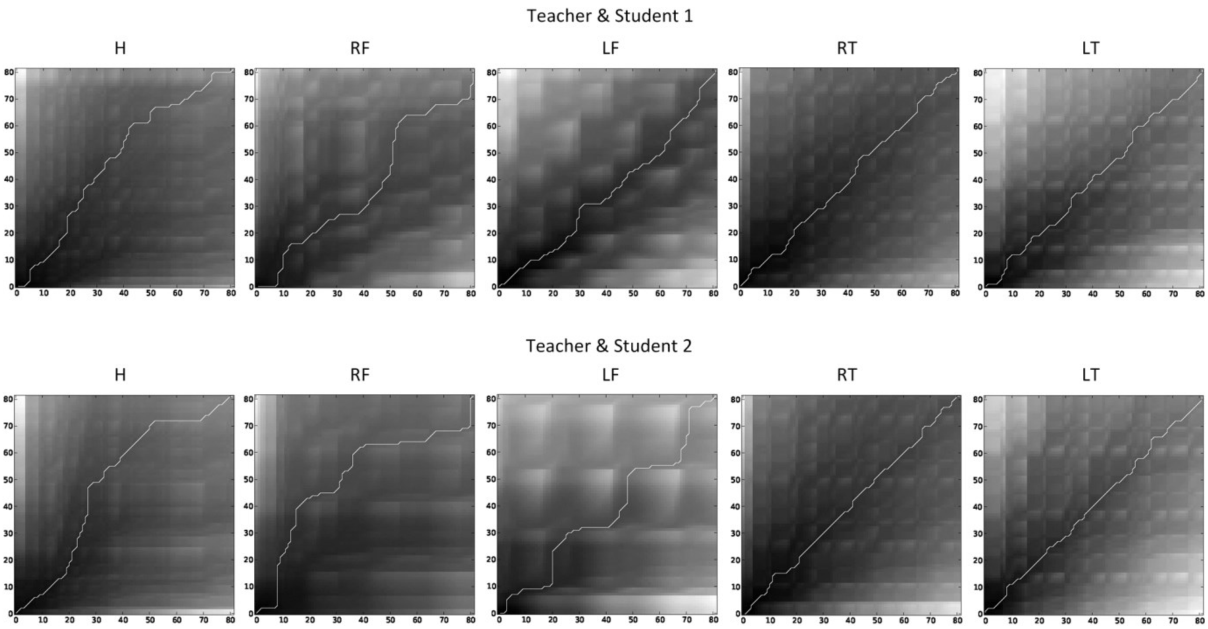


Figure 8-12: DTW cost matrices of all five body parts with speed as motion attribute

In each cost matrix in Figure 8-12 , the vertical axis represents the related time series of each body part of the teacher and the horizontal axis represents the time series of the students. Each matrix also shows a diagonal highlighted that indicates the warp path.

Ideally, the warp path should appear as a perfect diagonal along the matrix when the two time series compared are identical. Any deviation from the perfect diagonal indicates a mismatch, which may be due to temporal shifting. The top-rightmost cell of each DTW cost matrix represents the accumulated value of matches for the two time series. In this chapter, the DTW distance (i.e. the accumulated value of MDWP at the top-rightmost cell of the cost matrix) can be taken as an indicator of the performances of the students. In addition, the cost matrix itself presents informative patterns regarding the movement time series. For example, checker patterns are visible in all of the DTW cost matrices in Figure 8-12, which indicate that some repetitive temporal patterns exist. The longer the repetition period is, the greater the size of the squares becomes. As mentioned previously, the Samba dance has a quick beat that requires fast footwork. Thus, the size of the squares shown in the DTW cost matrices of the RT/LT is smaller compared to that of the RF/LF cost matrices in Figure 8-12.

Similarly, the movements of the lower body parts of the students are noticeably more consistent with respect to those of the teacher than their upper body parts. For example, in Figure 8-12, many fluctuations can be observed in the minimum time warp paths (MTWP) of the RF/LF obtained from the comparison between the time series of the teacher and student 2. The contribution of each of the fluctuations can be analysed through the smaller components of dance (i.e. beats).

For example, Figure 8-13 shows the DTW cost matrix of each beat separately determined from the comparison of the time series of the teacher and student 2 for the RF/ LF with speed as the motion attribute. As shown in Figure 8-13, student 2 moved his right finger (RF) in the second beat and his left finger (LF) in the third beat in a manner similar to that of the teacher because fewer fluctuations can be observed in the MTWPs.

The DTW cost matrices and MTWPs shown in Figure 8-13 are measured based on the Euclidean distance and using dynamic programming for the time series with speed as the motion attribute. However, for the time series with motion azimuth and vertical angle as the motion attributes, as discussed previously in this chapter, Eq. 8-2 is applied to form the DTW cost matrices and MTWPs.

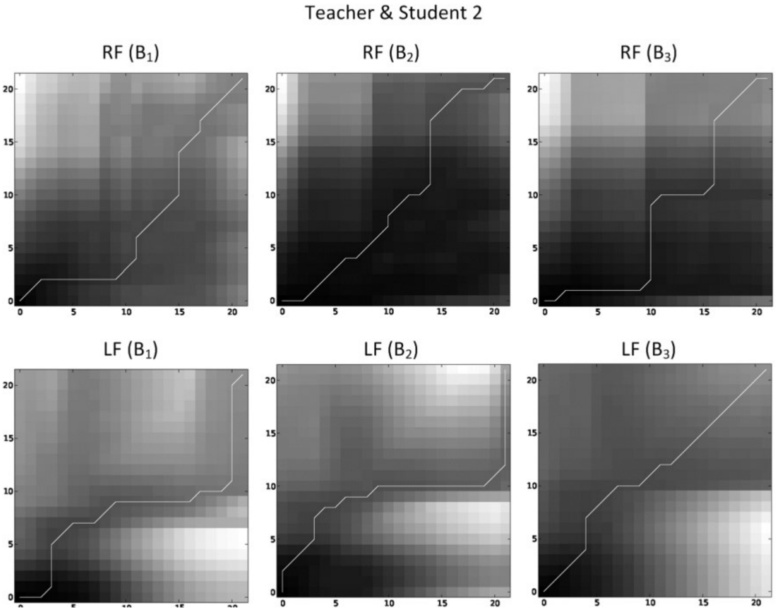


Figure 8-13: DTW cost matrix of RF/LF separately for each beat with speed as motion attribute for time series of teacher and student 2

For example, Figure 8-14 shows the DTW cost matrices of all body parts for time series with motion azimuth as the attribute, and Figure 8-15 shows the DTW cost matrices of the RF/LF for different beats.

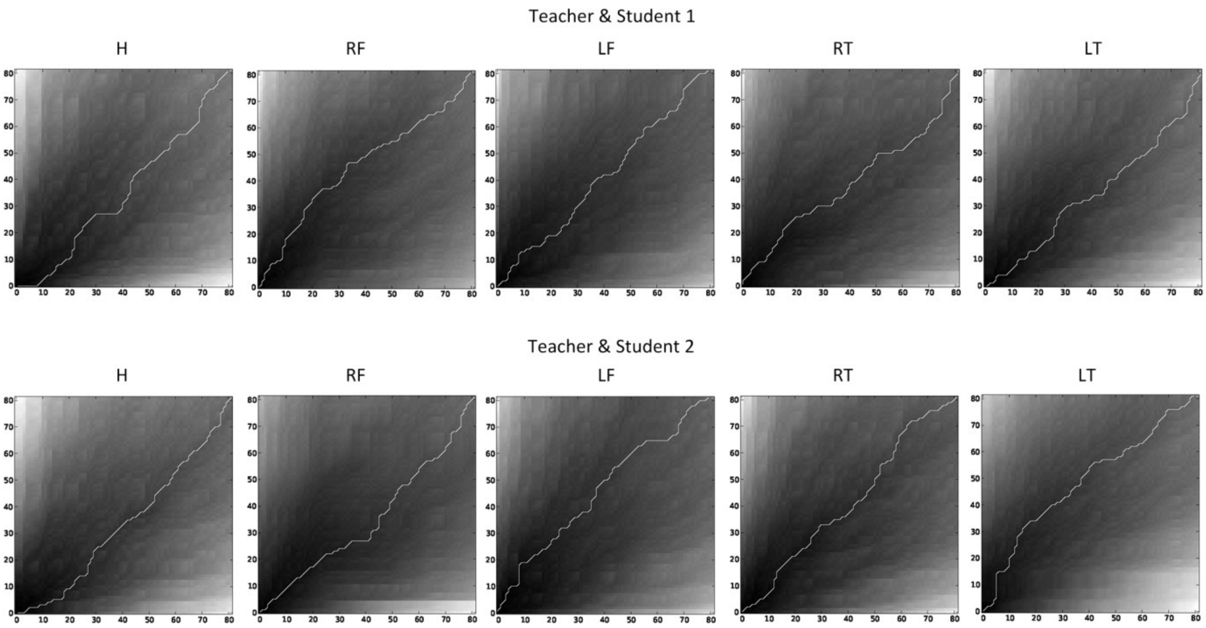


Figure 8-14: DTW cost matrices of all five body parts with motion azimuth as motion attribute

In comparing the matrices, for example, it may be concluded that the movement of the right finger of student 2 at the first beat was more similar to that of the teacher than the movement of his left finger (see Figure 8-15).

The minimum warp paths shown in the DTW cost matrices are considered criteria for comparing the performances of the different body parts of the dancers. The overall assessment and evaluation of the performances at each beat are illustrated in Figure 8-16 and Figure 8-17, respectively. The results are not on the same scale as those presented in Figures 10 and 11 and should be interpreted independently. Figure 8-16 shows that student 1 kept the speed of his hands as similar to that of his teacher than did student 2, whereas both students moved their feet at a pace similar to that of their teacher. From a directional point of view, specifically with respect to vertical angle, one may recognise a considerable difference between the movements of the left toe of student 1 and student 2. In addition, Figure 8-17 illustrates the assessment of the performances of the students with respect to smaller dance components, i.e. beats

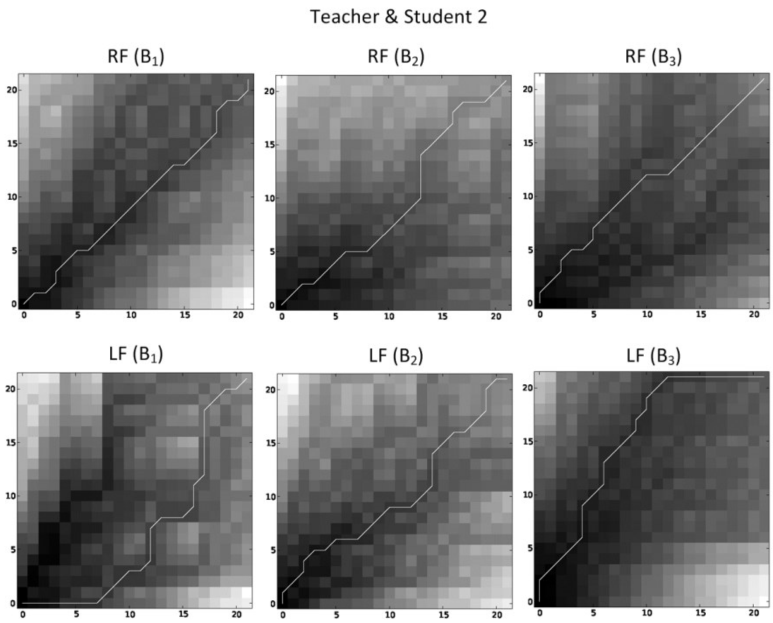


Figure 8-15: DTW cost matrix of RF/LF separately for each beat with motion azimuth as motion attribute

The following results could therefore be important in terms of handling noise caused by the temporal shift in movements. In many cases, the results from the DTW approach confirm the results from the REMO approach unless temporal shifting has a large impact on the output. The DTW results in this chapter suggest that the approach may be useful in studies of geographic phenomena that show repetitive and/or changing patterns over either long or short time periods, such as currents, earthquake, and tsunami (Gurgel et al., 2011; Kennedy & Crozier, 2010; Lipa et al., 2012; Shimamura et al., 2011; Wu et al., 2010)

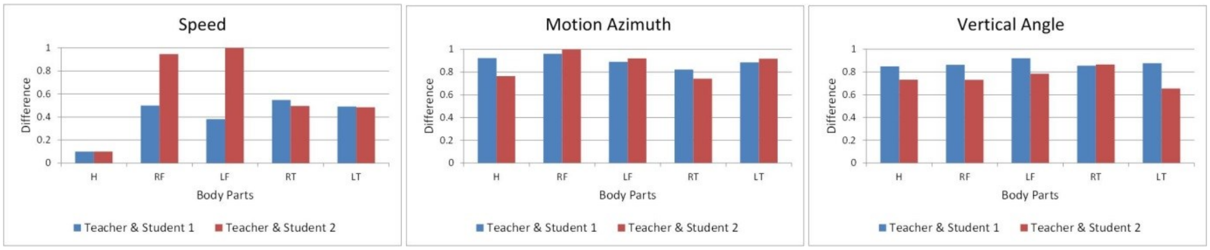


Figure 8-16: Overall assessment of performances of the students with respect to their teacher's performance

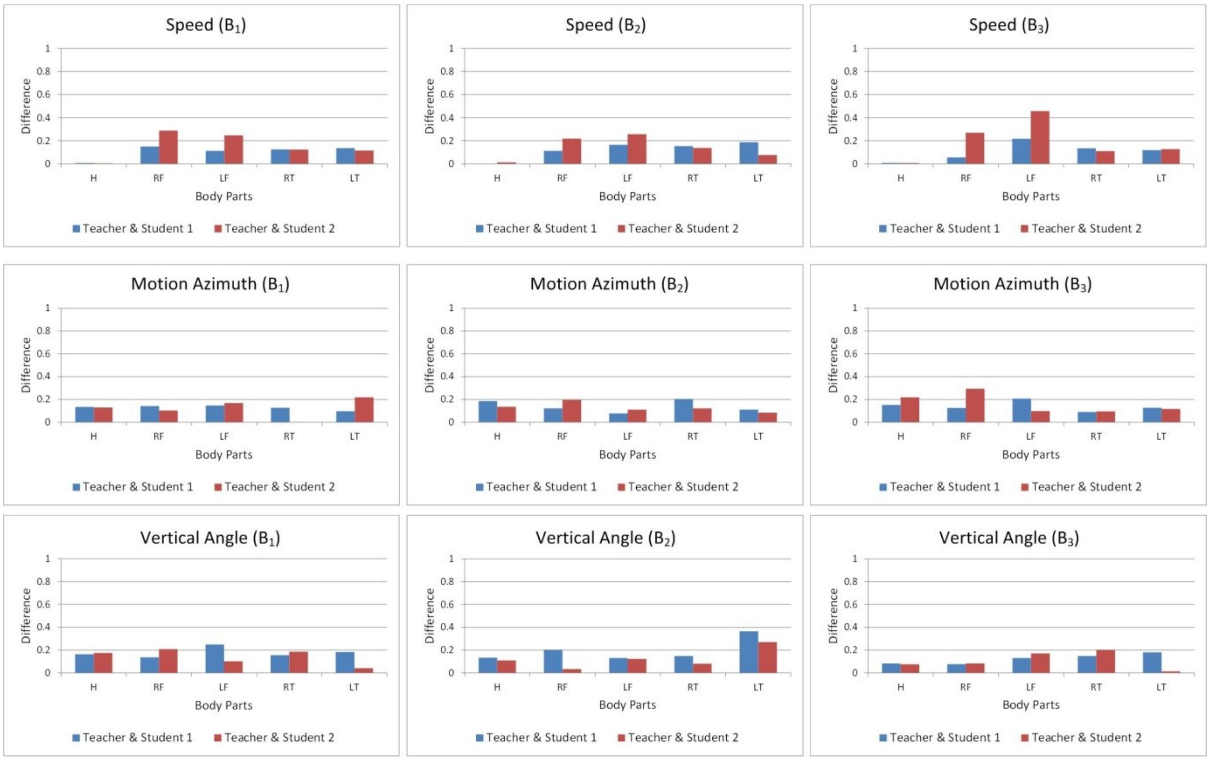


Figure 8-17: Assessment of performances of the students at each beat separately for different motion attributes

The DTW results in this study suggest that the approach may be useful in studies of other phenomena such as evacuations, earthquakes, etc. (Gurgel et al., 2011; Kennedy & Crozier, 2010; Lipa et al., 2012; Shimamura et al., 2011; Wu et al., 2010). For example, the DTW approach might be well integrated with other existing methods in mass event management to examine the behavioural patterns of tracked people and extract insightful information from sequences within such data (for example see (Delafontaine et al., 2012; Versichele et al., 2012)). The DTW approach could also be useful to deal with specific natural phenomena such as earthquakes. For example, Kaya et al., (2011) developed a novel change detection algorithm based on Discrete Cosine

Transform (DCT) and DTW to assess the damages in Haiti earthquake by using very high resolution imagery. In the proposed technique, the DTW algorithm was applied to correct spatial drifts caused by the change of sensors in the pre- and post-earthquake images and thereby minimize the issues caused by poorly registered images. In another area of the Earth sciences namely Geophysics, DTW is used to estimate relative time (or depth) shifts between two seismic images in seismic data processing (Hale, 2013). It is quite ubiquitous in seismic data processing to assess relative shifts in time (or depth) between seismograms.

8.4 Discussion

As a key contribution of this chapter, we addressed the applicability of the REMO and DTW approaches to analyse dance movements. With respect to the case study, both methods have some advantages and drawbacks.

Laube (2005) comprehensively identified the strengths and weaknesses of REMO. REMO is a true integration of space and time, simple and understandable, applicable to many research domains, and extensible. Besides these advantages, REMO suffers from some weaknesses such as its dependency on expert knowledge and discretisation of continuous data. Dynamic Time Warping (DTW) is a widely used method for warping two temporal signals and it is applicable in different domains including speech recognition, signature recognition, robotics, manufacturing, medicine, and shape matching.

The most commonly used approach to instruct dance skills is the demonstration-performance method (Maes et al., 2012). In addition, Ahlqvist et al (2010) introduced a viable tool for exploration, analysis, and knowledge construction from dance data sets in which the demonstrations and the illustrative animations were advantageous to identify differences and similarities in movement patterns. In this chapter, we showed that both REMO and DTW could be used in a dance educational context to explore and investigate the basics of dance movements. However, the understanding of the proposed visualisations requires some expert knowledge. For example, changing the order of entities in the layout of the REMO matrices may add some difficulties in understanding REMO representations. Therefore, automatic recognition algorithms can drastically enhance the determination of the quality of a student's performance in response to the music and in relation to the performance of the teacher in complex dance sequences.

The superiority of DTW over the Euclidean distance has been fully addressed in literature (e.g. (Aach & Church, 2001; Bar-Joseph et al., 2002; Chen et al., 2005b)). In this study, we applied a simple mathematical operation (i.e. subtraction) to draw analogy between REMO matrices. This approach is very sensitive to small distortions in the time axis. Assume two identical REMO matrices but one is shifted slightly along the time axis. The results of the comparison based on the subtraction operation say that the REMO matrices are very different. DTW overcomes this limitation and gives intuitive distance measurements between time series by ignoring both global and local shifting of the time dimension (Salvador & Chan, 2007).

Another advantage of DTW is the ability to handle time series with different lengths, while this is quite challenging with the REMO approach. This property is highly important to examine movements that suffer from data imperfection. Trajectories may contain quite some noise, usually caused by sensor failures and errors in detection techniques. Although outliers are less apparent in trajectories of objects captured by Infrared cameras compared to trajectories captured by GPS, it is challenging whether all captured data are free of any error and usable. For example, in the Tango dance, pairs of dancers perform very close to each other and this may result in some gaps in the tracked data because not all Infrared markers attached to the body parts of dancers are tracked properly.

As stated in (Laube et al., 2005), the REMO analysis helps us to investigate movements of many individual moving objects concurrently and thus allows detection of short- and long-term patterns such as convergence, divergence, and repetition as well as inter-object relationships. In the DTW approach, we may not investigate movements of multiple objects simultaneously and find such interesting patterns.

In the present work, we rate the quality of students' performances stimulating the students to improve their performances. However a common drawback to the presented techniques is the lack of an immediate feedback indicating how well students imitate teacher's movement. This is even more significant when motion time series are assessed together with the corresponding music and where rhythm and timing are dependently considered.

In spite the fact that segmentation of complex time series into smaller units eases perception and learning processes (Brown et al., 2006; Zacks & Swallow, 2007), the usefulness of the REMO and DTW methods with an $O(N^2)$ time and space complexity are limited only to small time series.

Generally, when performing a dance, specific motion patterns in synchrony with the music may be established (Maes et al., 2012). It is of particular interest to examine how the changing sampling rate (i.e. granularity) of time series may influence the results. In order to shed light on this important issue, we investigate the results of REMO and DTW when data is down sampled to a lower resolution. In addition, we illustrate how the use of original values in the time series instead of discretising them may lead to better results to understand movement patterns.

In Figure 8-17a, we illustrate the REMO matrix of the teacher motion azimuth for the first 50 time units of the movement data. Figure 8-17b represents discrete instances of change in motion azimuth rather than trying to perceive the entire motion processes. In this case, we transform continuous movement data (i.e. motion azimuth time series) to discrete classes. Although we concur with Laube's opinion that analysing with delimited change is easier than analysing the processes themselves, a comprehensive study in this regard can reveal the compromised levels of discretisation without losing much valuable information for that specific application. By looking at Figure 8-17b, in general, we may state that the formation of patterns on the REMO matrix is quite the same as the one detected in Figure 8-17a. But if we look more closely at the REMO matrix in Figure 8-17b, we may realise some significant differences due to the classification of the motion azimuth values. This issue not only affects the results obtained from the REMO approach, but also the results of the DTW approach. For example, the DTW results of the left finger (i.e. LF) of the teacher and student 2 are represented in Figure 8-17. In Figure 8-17b less detailed information appears compared to Figure 8-17a. In addition, the warp path in Figure Figure 8-17b is often straight and close to the main diagonal implying convincing performance by student 2, while even slight differences in the performances of student 2 compared to the teacher are reflected in the warp path attained from the original data (Figure 8-17a). In general, the low-resolution data limits the power of analysis and changes the results of the analysis. For example, the results of the REMO and DTW approaches for a lower sampling rate are illustrated in Figure 8-17c and Figure 8-17c. In Figure 8-17c, it is quite challenging to detect movement patterns and interrelations among moving objects. We also see that low-resolution has a significant impact on the results of DTW (Figure 8-17c). All these issues are open research problems and should be comprehensively investigated in the future.

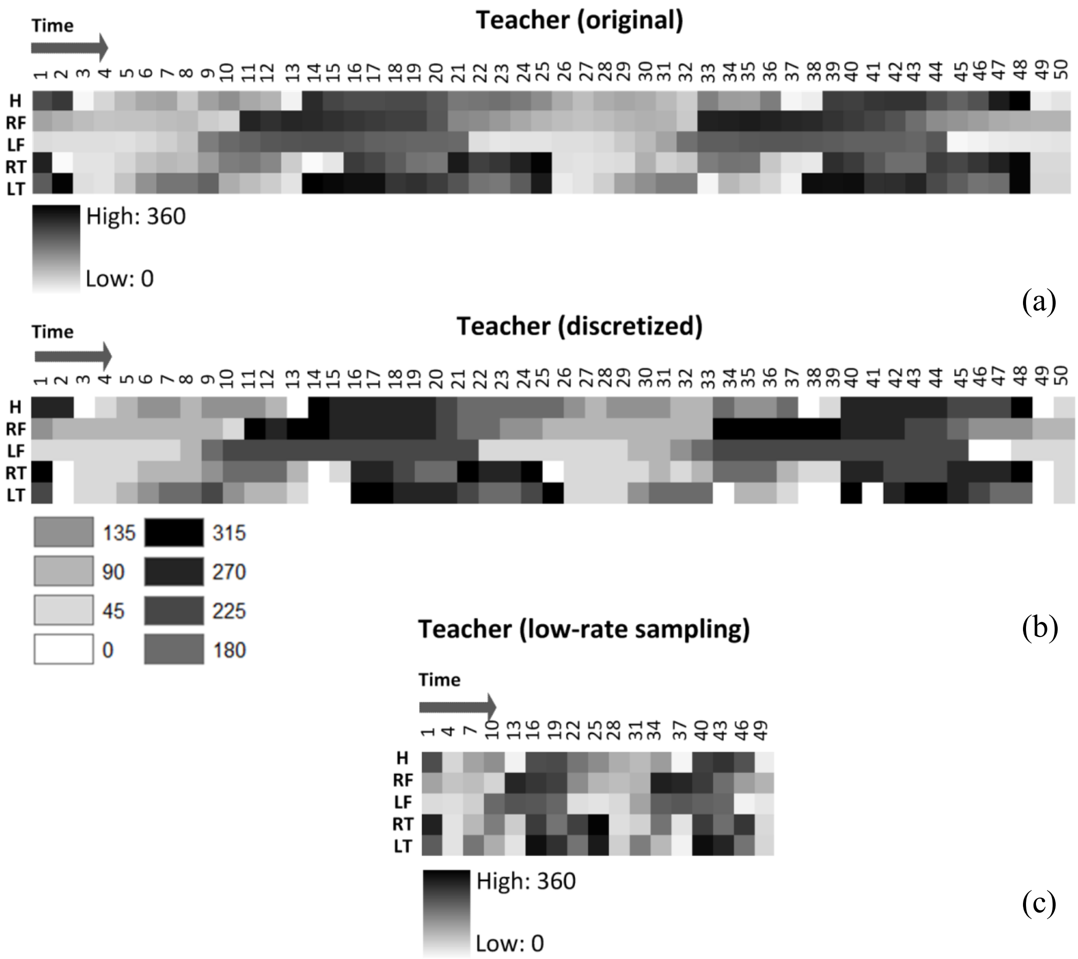


Figure 8-18: The REMO matrices of the teacher for (a) original data (b) discretised data (c) low-rate sampled data

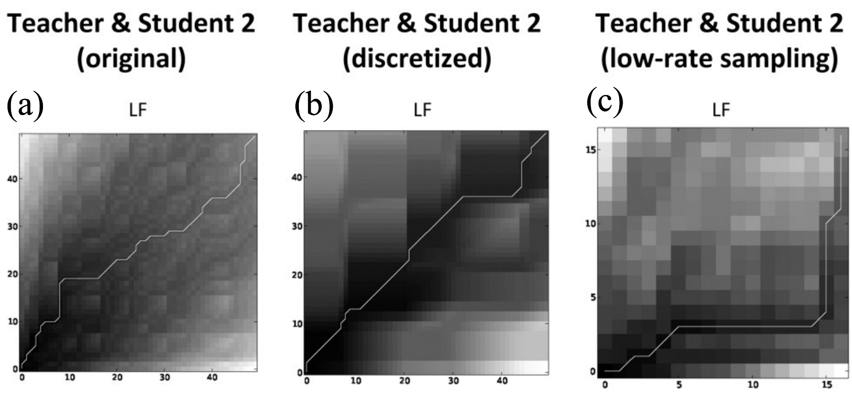


Figure 8-19: The DTW results of comparison left finger'movements of teacher and student 2 for (a) original data (b) discretized data (c) low-rate sampled data

8.5 Conclusions and Future Work

Similarity analysis of movements is recognised as an important task in many domains. This chapter demonstrated how to employ two well-known approaches, REMO and DTW, in the context of dance movements. Detecting and analysing patterns in movements is useful in a number of scenarios. The case study presented in this work involved several basic movements of Samba dancers that presented many regular patterns. REMO was applied to visualise and analyse similarity, particularly with respect to timing and synchronous motion patterns. In this chapter, three motion attributes, speed, motion azimuth, and vertical angle, were studied. Map algebra was employed to compare and measure similarity among REMO matrices. One of the characteristics of this method was its sensitivity to any displacement or shift in the base dataset. Therefore, DTW was proposed as a complementary method for analysing similarity among the motion patterns of moving objects. DTW calculates the similarity between two time series based on determining the optimal match between the series even if they are not identical in size.

Our contribution consisted in providing a sample case study to address, first, individual responses to spatio-temporal events and, second, higher-resolution data of individual movement using modular units exhibiting motion, i.e. body parts; in contrast, most existing studies use a single data unit for an individual. Using such higher-resolution data may be useful in gathering more information regarding individual behaviour. The proposed methodology can also be applied to a large number of dancers to examine their improvement throughout the course of a learning program.

References

Aach, J., & Church, G. M. (2001). Aligning gene expression time series with time warping algorithms. *Bioinformatics, 17* (6), 495-508.

Ahlqvist, O., Ban, H., Cressie, N., & Shaw, N. Z. (2010). Statistical counterpoint: Knowledge discovery of choreographic information using spatio-temporal analysis and visualization. *Applied Geography, 30* (4), 548-560.

Andrienko, G., Andrienko, N., Kopanakis, I., Ligtenberg, A., & Wrobel, S. (2008a). Visual analytics methods for movement data. In: F. Giannotti & D. Pedreschi (Eds.), *Mobility, Data Mining and Privacy - Geographic Knowledge Discovery* (pp. 375-410). Springer.

Bar-Joseph, Z., Gerber, G., Gifford, D. K., Jaakkola, T. S., & Simon, I. (2002). A new approach to analysing gene expression time series data. In: G. Myers, S.

Hannenhalli, D. Sankoff, S. Istrail, P. Pevzner & M. Waterman (Eds.), *Proceedings of the 6th Annual International Conference on Computational Biology* (pp. 39-48). Washington: ACM.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). The neural basis of human dance. *Cerebral Cortex, 16* (8), 1157-1167.

Chaudhry, R., Ravichandran, A., Hager, G., & Vidal, R. (2009). Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09).* (pp. 1932-1939).

Chen, L., Ozsu, M. T., & Oria, V. (2005). Robust and fast similarity search for moving object trajectories. *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data* (pp. 491-502). New York : ACM.

Das, G., Gunopulos, D., & Mannila, H. (1997). Finding similar time series. In: J. Komorowski & J. Zytkow (Eds.), *Principles of Data Mining and Knowledge Discovery* (pp. 88-100). Berlin Heidelberg: Springer.

Delafontaine, M., Versichele, M., Neutens, T., & Van de Weghe, N. (2012). Analysing spatiotemporal sequences in Bluetooth tracking data. *Applied Geography, 34*, 659-668.

Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.

Gurgel, K. W., Dzvonkovskaya, A., Pohlmann, T., Schlick, T., & Gill, E. (2011). Simulation and detection of tsunami signatures in ocean surface currents measured by HF radar. *Ocean Dynamics, 61* (10), 1495-1507.

Hale, D. (2013). Dynamic warping of seismic images. *Geophysics, 78* (2), S105-S115.

Kaya, G. T., Kaya, H., & Ersoy, O. K. (2011). Change detection in very high resolution imagery based on dynamic time warping: An implementation for Haiti earthquake damage assessment. Proceedings of the *International Workshop on the Analysis of Multi-temporal Remote Sensing Images (Multi-Temp'11)* (pp. 13-16). Trento: IEEE.

Kennedy, R. J., & Crozier, W. W. (2010). Evidence of changing migratory patterns of wild Atlantic salmon Salmo salar smolts in the River Bush, Northern Ireland, and possible associations with climate change. *Journal of Fish Biology, 76* (7), 1786-1805.

Laube, P. (2005). *Analysing Point Motion - Spatio-Temporal Data Mining of Geospatial Lifelines.* University of Zurich, Zurich.

Laube, P., & Imfeld, S. (2002). Analysing relative motion within groups of trackable moving point objects. In: M. Egenhofer & D. M. Mark (Eds.), *Geographic Information Science* (pp. 132-144). Berlin Heidelberg: Springer.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Lijffijt, J., Papapetrou, P., Hollmén, J., & Athitsos, V. (2010). Benchmarking dynamic time warping for music retrieval. In: F. Makedon (Ed.), *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments* (pp. 59). Samos: ACM.

Lipa, B., Isaacson, J., Nyden, B., & Barrick, D. (2012). Tsunami arrival detection with High Frequency (HF) radar. *Remote Sensing, 4* (5), 1448-1461.

Maes, P.-J., Amelynck, D., & Leman, M. (2012). Dance-the-Music: An educational platform for the modelling, recognition and audiovisual monitoring of dance steps using spatiotemporal motion templates. *EURASIP Journal on Advances in Signal Processing, 2012* (1), 1-16.

Morse, M. D., & Patel, J. M. (2007). An efficient and accurate method for evaluating time series similarity. *Proceedings of the ACM SIGMOD International Conference on Management of Data* (pp. 569-580). Beijing: ACM.

Nagashima, H., & Katsura, S. (2012). Human-motion analysis of grasping/manipulating motion including time-variable function using principal component analysis. *IEEE/SICE System Integration (SII)* (pp. 798-803).

Nowell, L. T. (1997). *Graphical encoding for information visualization: Using icon colour, shape, and size to convey nominal and quantitative data.* Virginia Polytechnic Institute and State University, Virginia.

Oveneke, M., Enescu, V., & Sahli, H. (2012). Real-time dance pattern recognition invariant to anthropometric and temporal differences. *Advanced Concepts for Intelligent Vision Systems* (pp. 407-419). Springer.

Qiao, Y., & Yasuhara, M. (2006). Affine invariant dynamic time warping and its application to online rotated handwriting recognition. *Proceedings of 18th International Conference on Pattern Recognition (ICPR 2006)* (Vol. 2, pp. 905-908). Hong Kong: IEEE.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing, 26* (1), 43-49.

Salvador, S., & Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis, 11* (5), 561-580.

Shimamura, K., Matsuzawa, T., Okada, T., Uchida, N., Kono, T., & Hasegawa, A. (2011). Similarities and differences in the rupture process of the M-4.8 repeating-earthquake sequence off Kamaishi, northeast Japan: comparison between the 2001 and 2008 events. *Bulletin of the Seismological Society of America, 101* (5), 2355-2368.

Sigal, L., Balan, A., & Black, M. J. (2010). Humaneva: synchronised video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *International Journal of Computer Vision, 87* (1-2), 4-27.

Tomlin, C. D. (1990). *Geographic Information Systems and Cartographic Modelling*. Prentice Hall.

Versichele, M., Neutens, T., Delafontaine, M., & Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography, 32* (2), 208-220.

Wu, Y. P., Wang, C. H., & Shen, Y. P. (2010). Spatial-temporal distribution of water vapor transportation over Tarim Basin during 1948–2009. *Journal of Glaciology and Geocryology, 32*, 1074-1084.

Yuan, Y., & Raubal, M. (2012). Extracting dynamic urban mobility patterns from mobile phone data. In: N. Xiao, M.-P. Kwan, M. Goodchild & S. Shekhar (Eds.), *Geographic Information Science* (pp. 354-367). Berlin Heidelberg: Springer.

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science, 16* (2), 80-84.

9

General Discussion and Conclusions

Stay hungry, stay foolish. Steve Jobs

9 General Discussion and Conclusions

The aim of this chapter is to summarise and discuss the contributions of this thesis in the light of the research questions.

RQ 1: How do we enhance the practical usefulness of QTC?

RQ 2: Is it possible to use QTC in the context of knowledge discovery from movement data?

RQ 3: How do we appropriately employ visualisation techniques in the analysis of movement data?

Within each of the previous chapters, a detailed description of the achievements was presented. In this chapter we supply a link between these findings and criticise the proposed methodologies. Issues and challenges for future research will be discussed.

9.1 Summary

In recent years, different domains (particularly GIS) paid much attention to the analysis of movement data, in particular moving objects. Although analysis of trajectories is generally considered as the starting phase to get insight into the behaviour of moving objects, analysis of interactions between MPOs can be an appropriate alternative approach. Given that the Qualitative Trajectory Calculus (QTC) is a unique qualitative spatio-temporal calculus to handle interactions among moving objects, this study made an effort to show the power of QTC in knowledge discovery from movement data.

Earlier theories introducing topological relationships between objects, such as the Region Connection Calculus (RCC) (Randell et al., 1992), Projective 9^+-Intersection Model (Billen & Kurata, 2008), and the 9-Intersection Model (Egenhofer & Franzosa, 1991), did not take into account the reasoning about continuously disjoint MPOs until the emergence of QTC (Van de Weghe, 2004). The idea behind QTC was to formalise the interaction between moving objects on the basis of the changing distance between them over time. In (Bogaert, 2008; Delafontaine, 2012; Van de Weghe, 2004), the properties of QTC were extensively defined and examined. Inspired by the former results, in this study, we highlighted a set of advantages of QTC in knowledge discovery from movement data over other existing techniques which focus more on geometric shapes of moving object trajectories rather than interrelations between objects. The proposed methods were illustrated in different case studies from different domains such as dance and squash games. In each chapter we discussed the structure of movement data and its properties that might have an effect on the analysis method and outcomes. In addition, since one of the characteristics of a beneficial knowledge discovery technique is to provide visual exploration contexts to intuitively interpret the retrieved information, each of the previously presented methodologies was accompanied with specific visualisation techniques to enhance the power of inferences. Visualisation offers an explicit image of outcomes and brings extra insight that cannot be easily obtained through traditional analysis techniques.

9.2 Discussion and General Conclusions

Each of the chapters in this dissertation focuses on different aspects of one or more research questions. Different contributions are discussed in relation to each research question. In Chapter 2, we presented a brief background on the principles of QTC. Then, an introduction to the concept of the Continuous Triangular Model (CTM) was given in Chapter 3. Chapters 4-8 included four papers that made up the core of the thesis. Here, we point out some of the strengths and weaknesses of the methodologies. By addressing these issues we can construct a blueprint for building an effective knowledge discovery technique.

To begin with, Chapter 4 described a methodology to identify, visualise and interpret repetitive movement patterns in groups of MPOs. This chapter dealt with the use of the CTM to visually explore movement patterns. In (Qiang et al., 2013), we demonstrated how to display linear data in different intervals in the CTM, constituting a basis for a multi-scale analysis. We also pointed out some of the advantages of CTM over other traditional multi-scale visualisation approaches. For example, in the CTM, moving statistics during intervals of different lengths can be displayed in one diagram, which offers an explicit overview of patterns in different scales.

In addition, the CTM is accompanied by a novel indexing technique. For example, in Chapter 4, we could reversely identify the reference point to those compared movement patterns representing a specific amount of similarity. Without delving deeper into the domain of choreography, we successfully showed that the integration of QTC and CTM is well-suited to study those dynamic phenomena in which repetitive behaviour is intrinsic.

In Chapters 5 and 6, we explained how to efficiently index and map movement patterns. We used a fractal-based indexing structure named sequence signature (SESI) to project movement patterns resulting from the MPO interactions. Similar to every other tree-data structure, such as quadtree, in the SESI, a two-dimensional space was recursively subdivided into smaller partitions (i.e. square cells). Each cell belongs exclusively to a specific movement pattern and embodied properties are related to that movement pattern such as frequency and duration. In more detail, Chapter 5 studied the use of SESI to visually summarise the relative movements of two MPOs, while we further developed the idea of SESI to explore and visualise the relative movements of multiple MPOs in Chapter 6.

The use of SESI increased our understanding of movement patterns. We could identify those patterns that had significant impact on the grouping interactive behaviour of objects in the case studies under investigation. Moreover, a structure to compare patterns during different time periods of movement was provided by which similar interactions (i.e. patterns) were detected and clustered.

Given the event-based approach introduced in Chapters 5 and 6, major patterns were typically short in length and therefore easily identifiable in the SESIs. A weak point of the methodology is nevertheless the inability to distinguish movement patterns in the higher lengths of SESIs due to the complex iterative nature of SESIs. Although the use of most basic QTC, i.e. QTC_B, involving a fair degree of abstraction of complex interactive movements of objects, was considered in the development of SESIs, readability of SESIs upon other types of QTC, such as QTC Double-Cross (QTC_C), will be even more challenging.

In Chapter 7, we investigated the extent to which the interactions of one pair of MPOs resembles the interactions of another pair. Generally, interaction is defined as a kind of action in which two or more objects have an effect upon one another (Andrienko et al., 2008b). We have designed our approach to study such patterns. For this purpose, the Sequence Alignment Method (SAM) and Edit Distance were applied to the sequences of QTC relations, representing the interactions among multiple MPOs during different time intervals of movement. The results confirmed the potential usage of SAM in exploring such patterns.

In Chapter 8, we explained and applied two well-defined methods, namely RElative MOtion (REMO) and Dynamic Time Warping (DTW). In the first method, we investigated the evolution of motion attributes of MPOs, such as speed and motion azimuth over time based on the REMO matrices. Laube (2005) comprehensively identified some of the strengths and weaknesses of REMO. For example, REMO is a true integration of space and time, simple and understandable, applicable to many research domains, and extensible. Besides these advantages, REMO suffers from some weaknesses such as its dependency on expert knowledge and discretisation of continuous data. In Chapter 8, REMO matrices provided effective representations to proficiently detect various types of patterns such as convergence, divergence, and repetition. The second approach comprised DTW of time series of the same motion attributes. Generally, a traditional distance measure such as Euclidean distance is not the best choice for comparing time series that may feature some noise and displacements. To overcome this problem, we considered DTW distance as an indicator to investigate the time series of dancers which may include noise and displacements.

Mining periodic patterns is crucial to model regularities in movement, predicting future trajectories and detecting outlying behaviour (Li, 2012). A number of techniques have been proposed in data mining literature to discover periodic patterns (e.g. (Han et al., 1999; Nishi et al., 2013; Yang et al., 2003; Yang et al., 2002)), but less in movement data (e.g. (Gudmundsson & Wolle 2010; Guochen et al., 2014; Turdukulov et al., 2014)). In this study, we delved deeper into periodicity and object relationships in movement data. Searching for partial periodic patterns in spatio-temporal databases is equally important to finding complete periodic patterns. This is more significant in movement-related datasets in which not all of the possible periodic patterns are complete, due to several factors such as failures in data collection procedure. In this regard, in Chapter 4, we could differentiate complete periodic patterns from partial ones in the CTM based on the definition of the proposed similarity measure. While, in Chapters 5 and 6, we paid less attention to this issue. There, the notion of similarity was merely the exact string-matching. Therefore, partially periodic patterns were neglected and thus not delineated in the SESIs. Inspired by prior studies, such as (Han et al., 1999; Ma & Hellerstein, 2001), the proposal of new periodicity detection algorithms that efficiently deal with partial periodic patterns is of vital importance. Indexing and representing partial periodic patterns in the SESIs assists us in fully comprehending the movement behaviour of objects.

In Chapters 7 and 8, we were able to differentiate both complete and partial periodic patterns by employing well-known data mining techniques for time series data. In fact, both SAM and DTW techniques are robust in the presence of shifting noise and, they may thus present a remarkable opportunity to discover knowledge contained in incomplete movement datasets.

The scope of the analysis of movement patterns is very wide. To understand movement patterns which may differ in various applications, we essentially require domain-specific expert knowledge. We claim that visual representation of patterns can assist us in the knowledge discovery procedure even if there is no abundant prior knowledge of data. Different visualisation techniques were proposed in this study. They were considered to be simple but informative so as to inspire one's analytical interest and facilitate analytical thinking. Prior knowledge about the movement data might support interpreting represented patterns, but not necessarily. The analytical usability of the visualisation techniques were demonstrated through some case studies. However, we need more evidence to confirm that non-expert users can effectively use the proposed representations.

Based on the definition by Mackinlay (1986), visual encoding is a projection of raw data records into graphical attributes. Graphical characteristics such as shape, size, and colour play important roles in increasing understandability of the visualised data. In this regard, many studies have been done to examine the perceptual effectiveness of various visual encodings (e.g. (Nowell, 1997)). We argue that the comprehension of the visualised patterns based on the employed encoding techniques in this thesis might need some expertise. A study can be conducted to determine the perceptual effectiveness of each visual encoding technique based on the nature of the attributes of the encoded data.

Clustering, or in other words grouping entities in such a way that entities in the same cluster are more similar to each other than to those in other clusters, finds numerous applications in diverse domains, such as financial markets, medical sciences, earth sciences, and, also, GIS. In a related study (Li, 2012), clustering of moving objects is categorised into two groups: moving object cluster discovery and trajectory clustering (see for example (Chen et al., 2005a; Gaffney et al., 2007; Pelekis et al., 2012; Vlachos et al., 2002)). In the former, the aim is to find clusters of objects with similar movement patterns or behaviour, whereas the latter puts an emphasis on the geometry to cluster trajectories (Li, 2012). We mainly contributed to the first group where the investigation of movement patterns is more significant.

Reliability and robustness of the results of clustering is extremely dependent on the proposed similarity measure (i.e. distance function) in the clustering procedure. Many of the existing distance functions, such as Euclidean distance and Longest Common Subsequences (LCSS), are sensitive to noise, shifts and scaling of data which are usually caused by sensor failures, errors in detection techniques, different sampling rates, etc. (Chen et al., 2005a). Comparing the results of clustering achieved in this research with the results of other distance functions, such as Edit Distance on Real sequence (EDR) (Chen et al., 2005a), which are most robust against the data imperfections is worthwhile.

Movement data themselves are complex. Accordingly, this may increase the computational complexity of models and methods of handling such data. Hence, we need to increase the general performance of the proposed methods through the use of efficient algorithms. Furthermore, the complexity of data conversion should also be considered. In all cases in this study, raw movement data were converted into other forms so they could easily be used in the study. In Chapter 4 for example, movement data of multiple point objects were transformed into QTC matrices. Although the proposed QTC matrix representation offered an elegant solution to analyse movement of multiple objects, such an amount of data increased computational costs exponentially. There are some possibilities to reduce computational costs, for example by employing data-reduction techniques in spatio-temporal pre-processing phases, without information loss (Rodríguez et al., 2003; Rodríguez et al., 2004).

9.3 Directions for Future Work

Chapters 4-7 augmented the potential applicability of QTC in knowledge discovery from movement data that were missed in earlier works. We then verified and confirmed that those achieved results were comparable with results obtained from the purely quantitative approaches applied in Chapter 8. We hope to address some of the below issues in future work.

In Chapter 4, the observed patterns in the CTM resulted from complex movements of multiple MPOs and implied interesting phenomena. Strategies taken to analyse and interpret such information are highly application dependent. Even though the visual comparison and interpretation of CTMs is worthwhile, based on the considerations and issues described in (Qiang et al., 2013), map algebra can insteadbe engaged in the process of manipulating and analysing CTMs. By applying map algebra to the CTM of dancers, patterns can be compared at different scales and more precisely answer the question ‘whether the performance of student 1 or student 2 is more similar to that of the teacher’ according to all possible intervals within the considered time frame.

Two types of granularity, namely the sampling granularity and the analysis granularity, were distinguished by Laube (2005). Choosing appropriate sampling and analysis granularity influences the interpretation of results. For example, in dance, as the main case study of this thesis, changing the granularity by milliseconds may significantly impact the outcome. While in other domains, such as ecology and in particular animal movement analysis, considering a granularity of hours might be adequate enough to reveal the most important aspects of the animal behaviour. Given that those patterns detected at a specific granularity may not be detected at other granularities, changing granularity may be used as an exploratory strategy to disclose interesting patterns in SESIs as well as in other approaches used throughout this thesis.

Movement data can be imperfect due to any combination of inaccuracy, imprecision, and vagueness. Uncertainties are often present with such kinds of data. Tackling vagueness and uncertainties associated with movement data is often complicated. In his doctoral dissertation, Van de Weghe (2004) presented the adequacy of QTC in dealing with incomplete knowledge about moving objects. There are possibilities to take this subject into account with the approaches proposed in this thesis. For example, to identify periodic movement patterns, in Chapters 4, 5 and 6, we proposed some similarity measures in which QTC conceptual animations (i.e. QTC relations hold between pairs of moving objects) were compared. However, QTC conceptual animations are not always complete, for example, due to lack of equal-interval observations or other error sources in the collection of movement data. Hence, some improvements are still needed with respect to the proposed similarity measures to handle such imperfect data. This enhancement would be of vital importance to effectively understand real-life cases.

Another crucial issue is the applicability of the proposed approaches to large sets of MPOs in geographic contexts. We applied our methodologies to several real-life cases. The examples throughout this thesis were kept intentionally as simple, clear and unambiguous as possible to stress major aspects of the reasoning power of the QTC calculus. The datasets used in all chapters of this thesis included rather small numbers of MPOs. It is direction for future research to study movement patterns of large number of objects. For example, in Chapter 8, we may examine the performances of hundreds of dancers based upon the data collected from all body parts. Note that this thesis did not attempt to technically describe specific applications. However, we illustrated that the given applications and the analysis tasks may impose limitations on the use of the proposed approaches.

We discussed and illustrated the adequacy of QTC_B in the procedure of knowledge discovery from movement data. In spite of increasing challenges and complexities, a dedicated extension to formally incorporate other types of QTC in this procedure would be a significant, complementary step. Moreover, extra attention should be paid to highlight the added value of QTC versus other formalisms, such as those of Dylla et al. (2007), Wolter et al. (2007), Hallot & Billen, (2008), Hornsby & King (2008), Golinska-Pilarek & Munoz-Velasco (2012), and Muñoz-Velasco et al. (2014) in this procedure.

In this thesis, we paid less attention to building up query-based analyses of movement data. Today, we witness a significant growth in the development of visual and dynamic query tools to be integrated with GIS and other analytical systems. We intend to support query tools in the proposed techniques. Such query tools play a crucial role in exploratory data analysis and enhance human-computer interactivity (Andrienko & Andrienko, 2006). In this respect, we may take some inspiration from the work in this area. For example, the use of the Triangular Model (TM) was fully investigated in visualising and analysing time intervals (Qiang et al., 2012). They offered not only a compact visualisation of the distribution of intervals, but also provided an innovative temporal query mechanism that relies on geometries in the two-dimensional space. Evolved from the TM, the CTM proposed in Chapter 4 may take advantage of such queries after making some modifications.

The proposed techniques in Chapters 5 and 6 may allow users to benefit from a frequently used query technique called brushing (Monmonier, 1989). With this technique, users are able to directly and flexibly select the movement patterns in the SESIs at different lengths they find interesting. Additionally, they may explore SESIs by successively making queries concerning the frequency and duration of movement patterns at different lengths and observe the visual answers in the SESIs. The effectiveness of such techniques increases when more complex queries can be handled such as those involving imperfect knowledge of movement data.

Weaver (2008) developed multidimensional query techniques for visual analysis of spatio-temporal information on the basis of *Improvise*. *Improvise* is a fully-implemented Java software architecture and user interface that enables users to build and browse highly-coordinated visualisations interactively (Weaver, 2006). For this purpose, he added the REMO algorithms (Laube et al., 2005) as new data transformation modules to the *Improvise* library (Weaver, 2004). The results showed the suitability of this approach to facilitate a flexible interactive querying framework across multiple data dimensions. As a part of Chapter 8, we employed the REMO approach to explore movement patterns of dancers. Accordingly, we may take advantage of this multidimensional query technique to increase the effectiveness of visualisations for data exploration.

We have restricted our detailed approaches to the analysis of those phenomena which are perceived as point objects. Indeed, not all moving objects in the real-world can be assumed as point objects. Future studies will consider how to handle movement of spatially extended objects. Inspired by (Van de Weghe et al., 2004), the movement of even more complex objects, of which the shape changes over time, can be studied as a further possible extension of this work.

From the start, this research was a multidisciplinary project, involving sports, dance, and transport. We believe that this research on exploring movement data can be employed in many other domains as well where discovering unexpected patterns, trends and relationships hidden in massive movement data are critical such as in social sciences, surveillance, and security. It is one of our interests to investigate behavioural patterns in animal observation data. Fine-grained data on animal movements are rapidly growing. Consequently, many substantial repositories of animal tracking data such as *Movebank*[5] are available to apply advanced methods for studying the movement behaviour of animals in their natural environment. In this regard, there is an ongoing collaboration with the animal ecology research group at the Edmund Mach Foundation[6]. We tend to integrate QTC into existing analytical approaches in order to

[5] https://www.movebank.org/
[6] http://cri.fmach.eu/

extract general patterns from an otherwise indiscernible collection of animal trajectories and, thus, create new insights into the territoriality and ecosystems of animals.

Recent advances in motion capture technologies have made it possible to record video images and motion data from patients with movement disorders. As a result, these latest technologies provide tremendous insight into the patient's health status and treatment progress. We believe that each of the techniques presented in this thesis might be integrated with existing analytical approaches, such as gait analysis, to increase our understanding of rehabilitation procedures. For example, SESI, representing the movement patterns of patients, may assist therapists to better detect abnormalities and improvements in the movement skills of patients, and therefore judge the patient's status, treatment and rehabilitation.

Undoubtedly, making use of appropriate analysis tools to analyse movement data allows us to effectively understand such complex phenomena (Andrienko et al., 2007). In this respect, few qualified data mining software and prototype tools are available for the flexible and scalable analysis of a mass of moving object, such as *MoveMine* (Li et al., 2010), *SECONDO* (Güting et al., 2010), *V-Analytics*[7], *QTCAnalyst* (Delafontaine et al., 2011), *GeoTM* (Qiang et al., 2012), *T-Pattern Miner*[8], and *Weka*[9]. During this research, we have also implemented some different small-scale tools to demonstrate the practical use of QTC in reasoning and analysing the movement of objects but not uniformly. This is direction for future research to design and develop a QTC prototype to interactively mine and explore other movement patterns such as flocks, leadership, avoidance, pursuit/evasion, and chasing, along with novel visualisation techniques.

References

Andrienko, G., Andrienko, N., & Wrobel, S. (2007). Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter, 9* (2), 38-46.

Andrienko, N., & Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Berlin Heidelberg: Springer-Verlag.

Andrienko, N., Andrienko, G., Wachowicz, M., & Orellana, D. (2008). Uncovering interactions between moving objects. In: T. J. Cova, H. J. Miller, K. Beard, A. U. Frank & M. F. Goodchild (Eds.), *Proceedings of the 5th International Conference on GIScience* (Vol. 5266, pp. 16-26). Park City, Utah.

[7] http://geoanalytics.net/V-Analytics/
[8] http://sourceforge.net/projects/t-patterns/
[9] http://www.cs.waikato.ac.nz/ml/weka/index.html

Billen, R, & Kurata, Y. (2008). Refining topological relations between regions considering their shapes. *Geographic Information Science*, (pp. 20-37). Berlin Heidelberg: Springer.

Bogaert, P. (2008). *A Qualitative Calculus for Moving Point Objects Constrained by Networks.* Ghent University, Ghent.

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., and De Maeyer, P. (2007). The Qualitative Trajectory Calculus on Networks. *Lecture Notes in Artificial Intelligence*, 4387, 20-38.

Chen, L., Ozsu, M. T., & Oria, V. (2005). Robust and fast similarity search for moving object trajectories. *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data* (pp. 491-502). New York : ACM.

Delafontaine, M. (2012). *Modelling and Analysing Moving Objects and Travelling Subjects: Bridging Theory and Practice.* Ghent University, Ghent.

Delafontaine, M., Cohn, A. G., & Van de Weghe, N. (2011). Implementing a qualitative calculus to analyse moving point objects. *Expert Systems with Applications, 38* (5), 5187-5196.

Dylla, F., Frommberger, L., Wallgrün, J. O., Wolter, D., Nebel, B., & Wölfl, S. (2007). SailAway: Formalizing navigation rules. *Proceedings of the Artificial and Ambient Intelligence Symposium on Spatial Reasoning and Communication (AISB'07)* (pp. 470-474).

Egenhofer, M. J., & Franzosa, R. D. (1991). Point-set topological spatial relations. *International Journal of Geographical Information System, 5* (2), 161-174.

Gaffney, S. J., Robertson, A. W., Smyth, P., Camargo, S. J., & Ghil, M. (2007). Probabilistic clustering of extratropical cyclones using regression mixture models. *Climate Dynamics, 29* (4), 423-440.

Golinska-Pilarek, J., & Munoz-Velasco, E. (2012). Reasoning with qualitative velocity: Towards a hybrid approach. *Lecture Notes in Computer Science*, *7208*, 635-646.

Gudmundsson, J., & Wolle, T. (2010). Towards automated football analysis: Algorithms and data structures. *Proceedings of the 10th Australasian Conference on Mathematics and Computers in Sport.*

Guochen, C., Hio, C., Bermingham, L., Kyungmi, Lee., & Lee, I. (2014). Mining frequent trajectory patterns and regions of interest from Flickr photos. *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS)* (pp. 1454-1463).

Güting, R. H., Behr, T., & Düntgen, C. (2010). SECONDO: A platform for moving objects database research and for publishing and integrating research implementations. *Proceedings of IEEE Data Eng. Bull* (pp. 56-63).

Hallot, P., and Billen, R. (2008). Life and motion configurations: A basis for spatio-temporal generalized reasoning model. *Advances in Conceptual Modelling–Challenges and Opportunities* (pp. 323-333).

Han, J., Dong, G., & Yin, Y. (1999). Efficient mining of partial periodic patterns in time series database. *Proceedings of the 15th IEEE International Conference on Data Engineering* (pp. 106-115).

Hornsby, K. S., & King, K. (2008). Modelling motion relations for moving objects on road networks. *Geoinformatica, 12* (4), 477-495.

Laube, P. (2005). *Analysing Point Motion - Spatio-Temporal Data Mining of Geospatial Lifelines.* University of Zurich, Zurich.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Li, Z. (2012). *Mining Periodicity and Object Relationship in Spatial and Temporal Data.* University of Illinois, Urbana-Champaign.

Li, Z., Ji, M., Lee, J.-G., Tang, L.-A., Yu, Y., Han, J., & Kays, R. (2010). MoveMine: Mining moving object databases. *Proceedings of the International Conference on Management of Data (ACM SIGMOD'10)* (pp. 1203-1206).

Ma, S., & Hellerstein, J. L. (2001). Mining partially periodic event patterns with unknown periods. *Proceedings of 17th International Conference on Data Engineering* (pp. 205-214).

Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics (TOG), 5* (2), 110-141.

Monmonier, M. (1989). Geographic brushing: Enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis, 21* (1), 81-84.

Muñoz-Velasco, E., Burrieza, A., Ojeda-Aciego, M. (2014). A logic framework for reasoning with movement based on fuzzy qualitative representation, *Fuzzy Sets and Systems*, *242*, 114-131

Nishi, M. A., Ahmed, C. F., Samiullah, Md., Jeong, B-S. (2013). Effective periodic pattern mining in time series databases. *Expert Systems with Applications*, *40* (8), 3015-3027

Nowell, L. T. (1997). *Graphical encoding for information visualization: Using icon colour, shape, and size to convey nominal and quantitative data.* Virginia Polytechnic Institute and State University, Virginia.

Pelekis, N., Andrienko, G., Andrienko, N., Kopanakis, I., Marketos, G., Theodoridis, Y. (2012). Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems, 38*(2), 343-391.

Qiang, Y., Chavoshi, S. H., Logghe, S., De Maeyer, P., & Van de Weghe, N. (2014). Multi-scale analysis of linear data in a two-dimensional space. *Information Visualization*, *13*(3), 248-265.

Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P., & Van de Weghe, N. (2012). Interactive analysis of time intervals in a two-dimensional space. *Information Visualization, 11* (4), 255-272.

Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In: B. Nebel, W. Swartout & C. Rich (Eds.), *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR)* (Vol. 92, pp. 165-176).

Rodríguez, M. A., Egenhofer, M. J., & Blaser, A. D. (2003). Query pre-processing of topological constraints: Comparing a composition-based with neighborhood-based approach. In: T. Hadzilacos, Y. Manolopoulos, J. Roddick & T. Y. (Eds.), *Proceedings of the 8th International Symposium on Advances in Spatal and Temporal Databases (SSTD)* (pp. 362-379).

Rodríguez, M. A., Van de Weghe, N., & De Maeyer, P. (2004). Simplifying sets of events by selecting temporal relations. In: M. Egenhofer, C. Freksa & H. Miller (Eds.), *Proceedings of the 3th International Conference on Geographic Information Science (GIScience)* (pp. 269-284).

Turdukulov, U., Calderon Romero, A. O., Huisman, O., & Retsios, V. (2014). Visual mining of moving flock patterns in large spatio-temporal data sets using a frequent pattern approach. *International Journal of Geographical Information Science*, (DOI:10.1080/13658816.2014.889834), 1-17.

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Van de Weghe, N., Maddens, R., Bogaert, P., Brondeel, M., & De Maeyer, P. (2004). Qualitative analysis of polygon shape-change. *Proceeding of International Geoscience and Remote Sensing Symposium IGARSS '04* (Vol. 6, pp. 4157-4159).

Vlachos, M., Kollios, G., & Gunopulos, D. (2002). Discovering similar multidimensional trajectories. *Proceedings of the 18th International Conference on Data Engineering* (pp. 673-684). San Jose: IEEE.

Weaver, C. E. (2004). Building highly-coordinated visualizations in Improvise. *Proceedings of the IEEE Symposium on Information Visualization, INFOVIS* (pp. 159-166).

Weaver, C. E. (2006). *Improvise: A User Interface for Interactive Construction of Highly-Coordinated Visualizations.* University of Wisconsin-Madison, Wisconsin.

Weaver, C. E. (2008). Cross-dimensional visual queries for interactive animated analysis of movement. *Proceedings of the Workshop in Geospatial Visual Analytics at GIScience Conference* (pp. 1-4).

Wolter, D., Dylla, F., Frommberger, L., Wallgrün, J. O., Nebel, B., & Wölfl, S. (2007). Qualitative spatial reasoning for rule compliant agent navigation. *Proceedings of the 20th International Florida Artificial Intelligence Research Society Conference* (pp. 673-674).

Yang, J., Wang, W., & Yu, P. S. (2003). Mining asynchronous periodic patterns in time series data. *IEEE Transactions on Knowledge and Data Engineering, 15* (3), 613-628.

Yang, R., Wang, W., & Yu, P. S. (2002). InfoMiner: Mining partial periodic patterns with gap penalties. *Proceedings of the International Conference on Data Mining ICDM'03* (pp. 725-728).

SAMENVATTING (DUTCH SUMMARY)

Door de technologische ontwikkeling in positionerings- en trackingsystemen, zoals GPS, mobiele positionering met Bluetooth en Wi-Fi, en videotracking, zijn er enorme hoeveelheden verplaatsingsgegevens beschikbaar voor analyse. Met zulke gegevens zijn er reeds een groot aantal studies gedaan met betrekking tot de analyse van trajecten, de extractie van bewegingspatronen en experimentele visuele analyses (zie bijvoorbeeld (Andrienko & Andrienko, 2007; Andrienko & Andrienko, 2012; Bak et al., 2012; Dodge, 2011; Giannotti & Pedreschi, 2008; Imfeld, 2000; Laube et al., 2005; Mountain, 2005)).

Ondanks deze inspanningen, is er weinig aandacht besteed aan het kwalitatief redeneren over bewegende objecten. Kwalitatieve formalismen, geschikt om kwalitatieve ruimtelijke en temporele relaties tussen entiteiten uit te drukken, worden algemeen aanvaard als een nuttige benadering om abstractie van de echte wereld te maken en op die manier de complexiteit van het denken over bewegende objecten te vereenvoudigen. De kwalitatieve traject calculus (QTC), die werd geïntroduceerd door Van de Weghe (2004), vertegenwoordigt een krachtige calculus voor de beoordeling van de interactie tussen disjunct bewegende puntobjecten (MPOs).

In dit proefschrift trachten wij kennis te verwerven uit verplaatsingsgegevens met gebruik van QTC. We halen zinvolle informatie uit bewegende objecten databases, ontdekken interessante patronen, en interpreteren ze op een plausibele manier. De volgende drie onderzoeksvragen (OV) werden behandeld:

OV 1: Hoe kunnen we de praktische bruikbaarheid van QTC verbeteren?

OV 2: Is het mogelijk om QTC te gebruiken bij het ontdekken van kennis uit verplaatsingsgegevens?

OV 3: Hoe kunnen we visualisatietechnieken gebruiken bij de analyse van verplaatsingsgegevens?

In de rest van deze samenvatting bespreken we hoe deze onderzoeksvragen in het proefschrift behandeld werden. In dit proefschrift hebben we vooral geprobeerd om ons begrip van het verplaatsingsgedrag van een of meer bewegende objecten uit te breiden. De inhoud van het proefschrift is onderverdeeld in hoofdstukken die werden gepubliceerd, aanvaard of ingediend ter beoordeling in internationale peer-reviewed tijdschriften of boeken op het moment van schrijven. Om ervoor te zorgen dat deze hoofdstukken zelfstandig van elkaar gelezen kunnen worden, zijn er een aantal onvermijdelijke overlappingen in de afzonderlijke hoofdstukken zoals het literatuuronderzoek en de beschrijving van het basisconcept van de Kwalitatieve Traject Caculus.

Na de algemene inleiding in Hoofdstuk 1 waarin de motivatie voor het onderzoek geschetst wordt, wordt in Hoofdstuk 2 een overzicht gegeven van de theoretische basis van de Kwalitatieve Traject Calculus (Qualitative Trajectory Calculus of afgekort QTC) en hoe deze calculus kan worden geïmplementeerd en uitgebreid om ruwe bewegende objecten te vertegenwoordigen en te beredeneren. QTC wordt geïntroduceerd als een tijdruimtelijke kwalitatieve calculus waarbij een veranderende afstand tussen twee objecten wordt bijgehouden doorheen de tijd. Tussen de verschillende types van QTC zoals QTC-Basis (QTC_B), QTC-Dubbel-Kruis (Double-Cross) (QTC_C), en QTC-Network (QTC_N), richten we ons op het meest eenvoudige type (d.w.z. QTC_{B1}). In QTC_{B1}, worden complexe bewegingen vereenvoudigd via relaties tussen paren van op elkaar inwerkende gescheiden puntobjecten. In totaal zijn er negen QTC_{B1} basisrelaties tussen twee disjuncte bewegende objecten. In dit proefschrift tonen we de mogelijkheden tot redeneren met zulke eenvoudige relaties en de extractie van kennis uit bewegingen.

In Hoofdstuk 3 wordt een innovatieve visuele representatie voor tijdreeksen, namelijk het Continue Triangulaire Model (Continous Trangular Model of afgekort CTM), geïntroduceerd. In CTM kunnen alle subintervallen van een tijdreeks in een tweedimensionaal continue gebied. Elk punt vertegenwoordigt een subinterval van de tijdreeks en de waarde die het punt vertegenwoordigt wordt verkregen via een bepaalde functie (bijvoorbeeld het gemiddelde of de som) over de tijdreeks binnen het subinterval. Het CTM geeft dus een expliciet overzicht van tijdreeksen op alle verschillende schalen. Naast tijdreeksen kan CTM ook worden toegepast op lineaire gegevens.

In Hoofdstuk 4 gebruiken wij het concept van QTC en CTM voorgesteld in Hoofdstukken 2 en 3. Dit hoofdstuk stelt een methode voor bestaande uit drie fases voor de identificatie, visualisatie en interpretatie van repetitieve bewegingspatronen in groepen van bewegende puntobjecten. Bewegingen van lichaamsdelen van dansers worden beschreven door opeenvolgingen van QTC_B matrices, die op hun beurt worden gebruikt om de herhalende bewegingspatronen te identificeren (*OV 1)*. Vervolgens wordt een vergelijkende analyse gemaakt om de mate van gelijkenis tussen paren van sequenties te bepalen (*OV 2)*. Ten slotte wordt CTM toegepast om de mate van gelijkenis tussen alle paren van sequenties weer te geven. In dit hoofdstuk konden wij zien hoe CTM de visuele analyse van bewegingspatronen vergemakkelijkt (*OV 3*).

De ontdekking van kennis uit trajecten van bewegende objecten is een belangrijk en uitdagend probleem in veel onderzoeksdomeinen. In Hoofdstuk 5 stellen wij een nieuwe benadering voor voor het identificeren, vertegenwoordigen en clusteren van bewegingspatronen. Sequenties van QTC relaties van de beweging van objecten worden geïndexeerd en weergegeven in een Sequence Signature (SESI) (*OV 1)*. Een SESI is een fractaalfunctie die een visuele samenvatting geeft van de bewegingspatronen van twee MPOs (*OV 2&3)*. We vergelijken SESIs met behulp van een afstandsfunctie, waardoor het mogelijk is om clusters van kwalitatief onderscheiden trajectparen te identificeren. De voorgestelde methode wordt geïllustreerd aan de hand van twee real-world voorbeelden van MPO interactie: auto's op een snelweg en squash spelers. Deze eenvoudige voorbeelden tonen het nut van onze benadering aan voor het blootleggen van bewegingspatronen die zijn verborgen in tijdruimte traject databases.

In Hoofdstuk 6 breiden wij het concept van SESI verder uit om kennis uit beweging van meer dan twee MPOs teverwerven. De toepasbaarheid van de voorgestelde methodologie wordt geïllustreerd aan de hand van een praktisch voorbeeld, de samba dans. Daarbij worden de bewegingen van dansers tijdens verschillende tijdsintervallen vergeleken. De resultaten tonen aan dat de voorgestelde methode effectief kan worden gebruikt om interacties van meerdere MPOs in verschillende domeinen te analyseren.

Om de gelijkenissen in beweging tussen bewegende objecten te analyseren wordt in Hoofdstuk 7 een innovatieve aanpak voorgesteld waarbij de sequentie-aligneringsmethode (Sequence Alignment Method, afgekort SAM) wordt gebruikt voor het aligneren en beoordelen van QTC-sequenties. QTC-informatie werd gebruikt om QTC-sequenties te vormen (*OV 1)*. SAM vergemakkelijkt de identificatie en visualisatie van bewegingspatronen van interacties tussen bewegende objecten (*OV 2&3)*. De voorgestelde methode kan worden toegepast in elk domein waarbij inzicht in bewegingspatronen belangrijk is. In dit hoofdstuk werden de bewegingen van drie samba danseressen geanalyseerd om de mate van (on)gelijkheid tussen de bewegingen van de dansers te meten. Gelijkheid / ongelijkheidsanalyse draagt bij tot een beter begrip van hoe de dansers bewegen. De resultaten van de gelijkheidsanalyse van QTC-sequenties worden gepresenteerd in dendrogrammen, waarbij de sequenties worden gegroepeerd in clusters.

In Hoofdstuk 8 worden in plaats van QTC twee andere kwantitatieve methodes gebruikt om beweging te analyseren. De eerste methode gebruikt kaartalgebra met relatieve beweging (RElative MOtion afgekort REMO) matrices om de evolutie van bewegingsattributen, zoals snelheid en bewegingsazimut, in de tijd te bestuderen. De tweede methode is «Dynamic Time Warping, afgekort DTW» om tijdreeksen van bewegingsattributen te analyseren. De resultaten tonen aan dat beide methoden nuttig zijn in het numeriek vergelijken van de bewegingen van samba dansers en om visuele bewegingspatronen te herkennen (*OV 2&3)*.

Het laatste hoofdstuk van dit proefschrift bevat een uitgewerkte algemene discussie met betrekking tot de initiële onderzoeksvragen. Hierbij wordt ingegaan op problemen en toekomstige uitdagingen.

References

Andrienko, N., & Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. *Cartographica, 42* (2), 117-138.

Andrienko, N., & Andrienko, G. (2013). Visual analytics of movement: An overview of methods, tools and procedures. *Information visualization, 12* (1), 3-24

Bak, P., Marder, M., Harary, S., Yaeli, A., & Ship, H. J. (2012). Scalable detection of spatiotemporal encounters in historical movement data. *Computer Graphics Forum, 31* (3), 915-924.

Dodge, S. (2011). *Exploring Movement Using Similarity Analysis.* University of Zurich, Zurich.

Giannotti, F., & Pedreschi, D. (2008). Mobility, data mining and privacy: A vision of convergence. In: F. Giannotti & D. Pedreschi (Eds.), *Mobility, Data Mining and Privacy-Geographic Knowledge Discovery (pp. 1-11)*. Berlin Heidelberg: Springer

Imfeld, S. (2000). *Time, Point and Space - Towards a Better Analysis of Wildlife Data in GIS.* University of Zurich, Zurich.

Laube, P., Imfeld, S., & Weibel, R. (2005). Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science, 19* (6), 639-668.

Mountain, D. M. (2005). *Exploring Mobile Trajectories: An Investigation of Individual Spatial Behaviour and Geographic Filters for Information Retrieval.* City University, London.

Van de Weghe, N. (2004). *Representing and Reasoning about Moving Objects: A Qualitative Approach.* Ghent University, Ghent.

Biographical sketch

Seyed Hossein Chavoshi (°1983, Iran) started his education of Geomatics and Surveying in 2002 at the Institute of Surveying and Mapping, National Geographical Organization of Iran, where he received a Bachelor degree in 2006. After this program, he accomplished a Master program of geographic information science (GIS) at the University of Tehran. In 2008, he started working on a PhD project funded by the Special Research Fund (BOF). His PhD Research focuses on the analysis of movement patterns. His research outcomes have been published in or submitted to a number of international journals and conference proceedings.