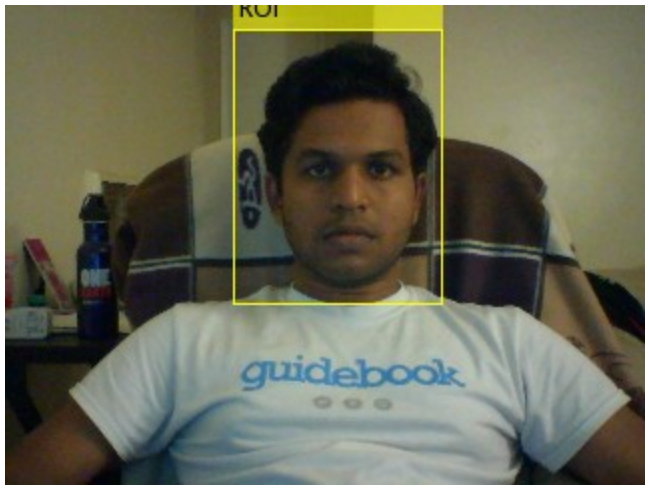
ROI BASED VIDEO COMPRESSION TAILORED FOR CONVERSATIONAL VIDEO COMMUNICATIONS

Pradeep Kumar Govindaraju

Objective

- Scenario - Video Communication
- Detect Human face (Region of Interest) in real time.
- Compress ROI with higher fidelity than other parts.



ROI detection algorithm

- Visual features including motion
- Segmentation algorithms
- Learning / Classification based

Viola Jones Face Detection

To represent ROI - Rectangular bounding box

Image Source : http://siret.ms.mff.cuni.cz/facereco/method/

Forward Transform & Quantization

- Discrete Cosine Transform

DCT Coefficients

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantized coefficients

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	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	0

Quantization table

1	1	2	4	8	16	32	64
1	1	2	4	8	16	32	64
2	2	2	4	8	16	32	64
4	4	4	4	8	16	32	64
8	8	8	8	8	16	32	64
16	16	16	16	16	16	32	64
32	32	32	32	32	32	32	64
64	64	64	64	64	64	64	64

Original Quantized coefficient calculation

QF = (32 x F)/(2 x S x W)

F – DCT coefficient

S – Quantization Scale

W – Quantization table element

Modified Quantized Coefficient Calculation for ROI

QF = (32 x F x sROI)/(2 x S x W)

Forward Transform & Quantization

- Integer Transform
  - We can control Quantization Parameter value.
  - QP value implicitly defines all other parameters like Quantization step
  - So ROI priority is varied by varying the QP (0 – 51).

QP	$\sim Q_{step}$ (n = 0)	$\sim Q_{step}$ (n = 1)	$\sim Q_{step}$ (n = 2)
0	0.625	0.625	0.6423
1	0.6875	0.7031	0.6917
2	0.8125	0.7812	0.7906
3	0.875	0.8984	0.8894
4	1.0	0.9766	0.9882
5	1.125	1.1328	1.1364
6	1.25	1.25	1.2847
...	...		
12	2.5	2.5	2.5694
...	...		
18	5.0	5.0	5.1387
...	...		
48	160	160	164.4384
...	...		
51	224	230	227.6840

Source : The H.264 advanced video compression standard textbook

	Input Image	IT – Total (kilo bits)	DCT – Total (kilo bits)	IT- Comp. Ratio	DCT- Comp. Ratio	IT – ROI PSNR	DCT – ROI PSNR	IT- nonROI PSNR	DCT- nonROI PSNR	IT- Total PSNR	DCT- Total PSNR
1	Image 1	154.32	108.13	0.1580	0.1107	43.594	45.04	35.67	35.64	36.90	36.95
2	Image 1	146.95	92.37	0.150	0.0946	42.03	42.02	35.67	35.62	36.78	36.74
3	Image 1	124.94	98.70	0.1279	0.1010	42.03	42.02	24.71	38.34	26.18	39.14
4	Image 2	148.05	99.74	0.152	0.102	44.803	47.625	32.65	32.03	33.86	33.29
5	Image 2	145.64	89.83	0.15	0.092	44.10	44.13	32.66	32.02	33.88	33.24
6	Image 2	98.17	98.14	0.10	0.10	41.07	41.08	23.10	35.31	24.38	36.23

Image 1
[560 x 420]
(2)

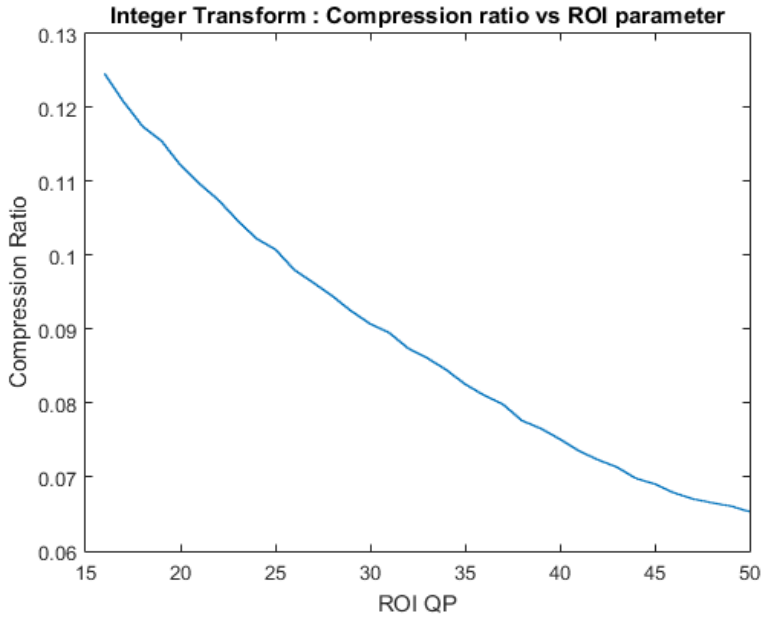
DCT based Compression

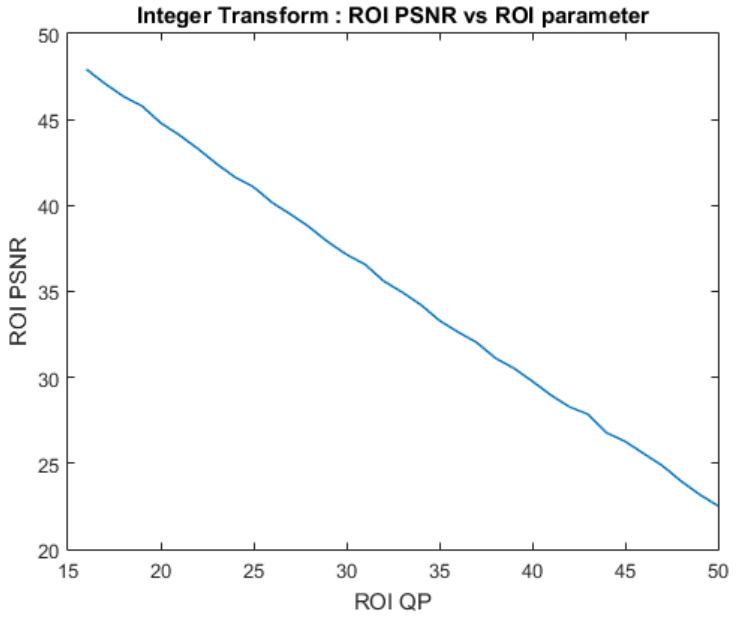
Integer Tranform based Compression

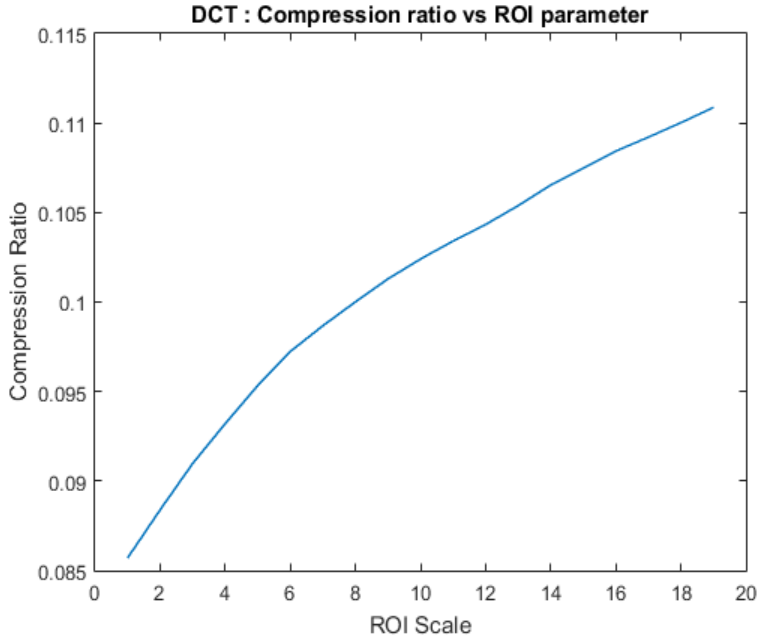
Image 2
[424 x 304]
(5)

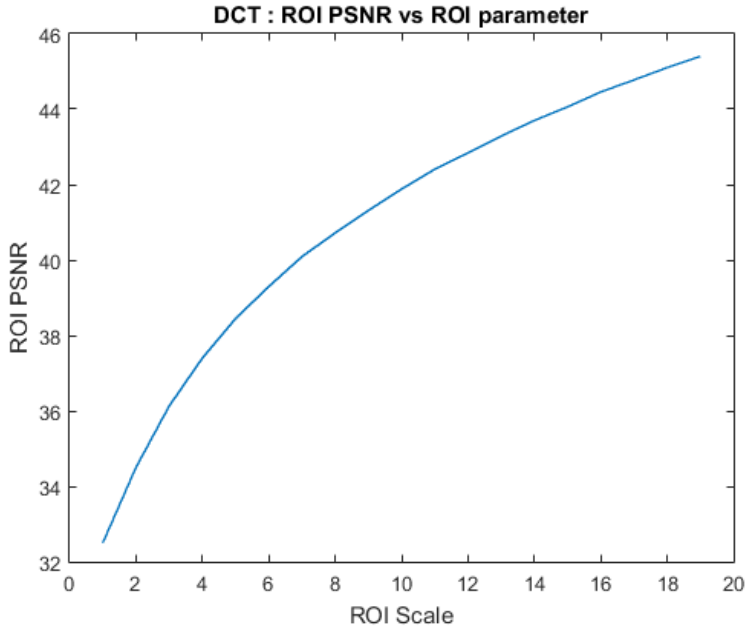
DCT based Compression

Integer Tranform based Compression


Integer Transform : Compression ratio vs ROI parameter
Compression Ratio
ROI QP
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0.06
15
20
25
30
35
40
45
50


Integer Transform : ROI PSNR vs ROI parameter
ROI PSNR
ROI QP
50
45
40
35
30
25
20
15
20
25
30
35
40
45
50


DCT : Compression ratio vs ROI parameter
Compression Ratio
ROI Scale
0.115
0.11
0.105
0.1
0.095
0.09
0.085
0
2
4
6
8
10
12
14
16
18
20


DCT : ROI PSNR vs ROI parameter
ROI PSNR
ROI Scale
46
44
42
40
38
36
34
32
0
2
4
6
8
10
12
14
16
18
20

Motion Compensation[30 frames]

	ROI Priority Parameters	Macroblock Size	Bitrate Achieved	Video Size – Compression ratio	ROI PSNR	Non ROI PSNR	Total PSNR
1	ROI = 20 | Non ROI = 35 (Integer Transform)	16 x 16	1.2893 Mbps	0.068043	20.508	18.589	18.970
2	ROI = 20 | Non ROI = 35 (Integer Transform)	8x8	1.2865 Mbps	0.067894	20.627	18.627	19.038
3	ROI = 20| Non ROI = 35 (Integer Transform)	4x4	1.3147 Mbps	0.069385	20.901	18.671	19.157
4	ROI – 1.05 | Non ROI – 0.06 (DCT)	16x16	0.9905 Mbps	0.052279	20.511	18.462	18.883
5	ROI – 1.05 | Non ROI – 0.06 (DCT)	8x8	0.9939 Mbps	0.052454	20.6294	18.484	18.938
6	ROI – 1.05 | Non ROI – 0.06 (DCT)	4x4	1.0554 Mbps	0.055702	20.8993	18.523	19.048

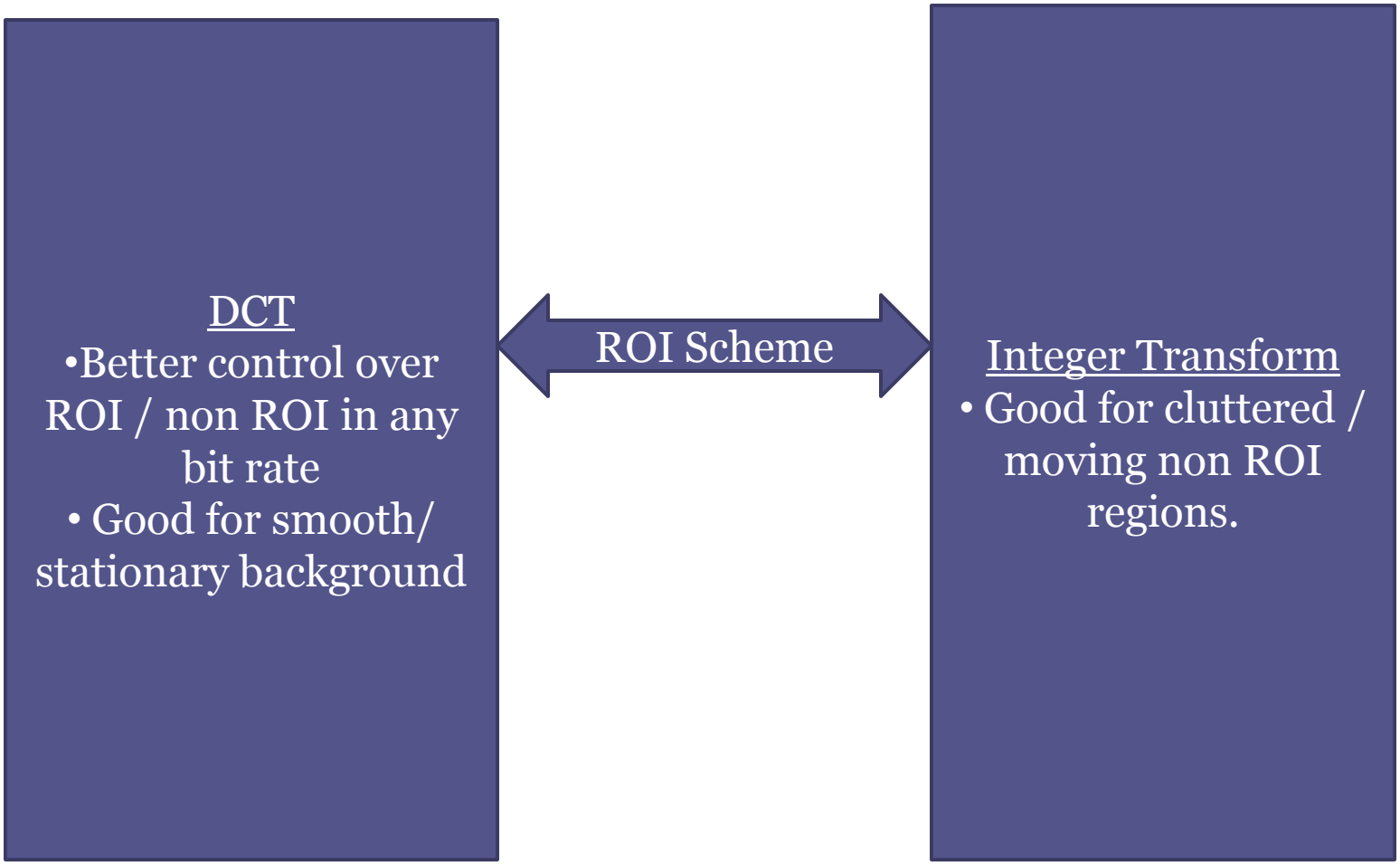
- GOP = IPIPIPIP…
- Cost function for motion estimation : Mean Absolute Difference
- Residue is transformed and quantized with same ROI scheme
- All the Bitrates are calculated based on Entropy.
- Video – Missa [Very smooth non ROI / background]

Motion Compensation [440 frames]

	ROI Priority Parameters	Macrob lock Size	Bitrate Achieved	Video Size – Compression ratio	ROI PSNR	Non ROI PSNR	Total PSNR
1	ROI = 24 | Non ROI = 40 (Integer Transform)	16 x 16	0.999 Mbps	0.049269	20.508	18.589	18.970
2	ROI – 1.05 | Non ROI – 0.1 | (DCT)	16x16	1.061 Mbps	0.052336	19.508	15.521	15.882
3	ROI – 0.78 | Non ROI – 0.078 | (DCT)	16x16	0.842 Mbps	0.041565	18.404	14.981	15.300

- GOP = IPIPIPIP…
- Cost function for motion estimation : Mean Absolute Difference
- Residue is transformed and quantized with same ROI scheme
- All the Bitrates are calculated based on Entropy.
- Video – Salesman [Cluttered background]

Conclusion


DCT
•Better control over ROI / non ROI in any bit rate
• Good for smooth/ stationary background
ROI Scheme
Integer Transform
• Good for cluttered / moving non ROI regions.

Thank You

Questions ?