NEW APPROACHES TO COLOR IMAGE RESTORATION AND ZOOMING OF COMPRESSED VIDEO

Narasimha Kaulgud

- Restoration of color images, degraded by inter-channel blur
- Zooming of still images
- Zooming of compressed video

Mode of Presentation

- Color Image Restoration
  Markov Random Fields, Observation Model, Energy function, Results, Observations and Limitations
- Still Image Zooming
  Existing Methods, Using MRF, MRA, MRA Formulation, Joint method, Color image zooming, Observations and Limitations
- Compressed Video Zooming Motivations, Video coding and compression, Zooming, Motion Estimation, Proposed method, MRME, Performance Measure Frame interpolation, Observations
- Conclusions
- Future Directions

Color Image Restoration:

Removal of degradations (blur and/or noise) from the observed degraded image

Markov Random Field

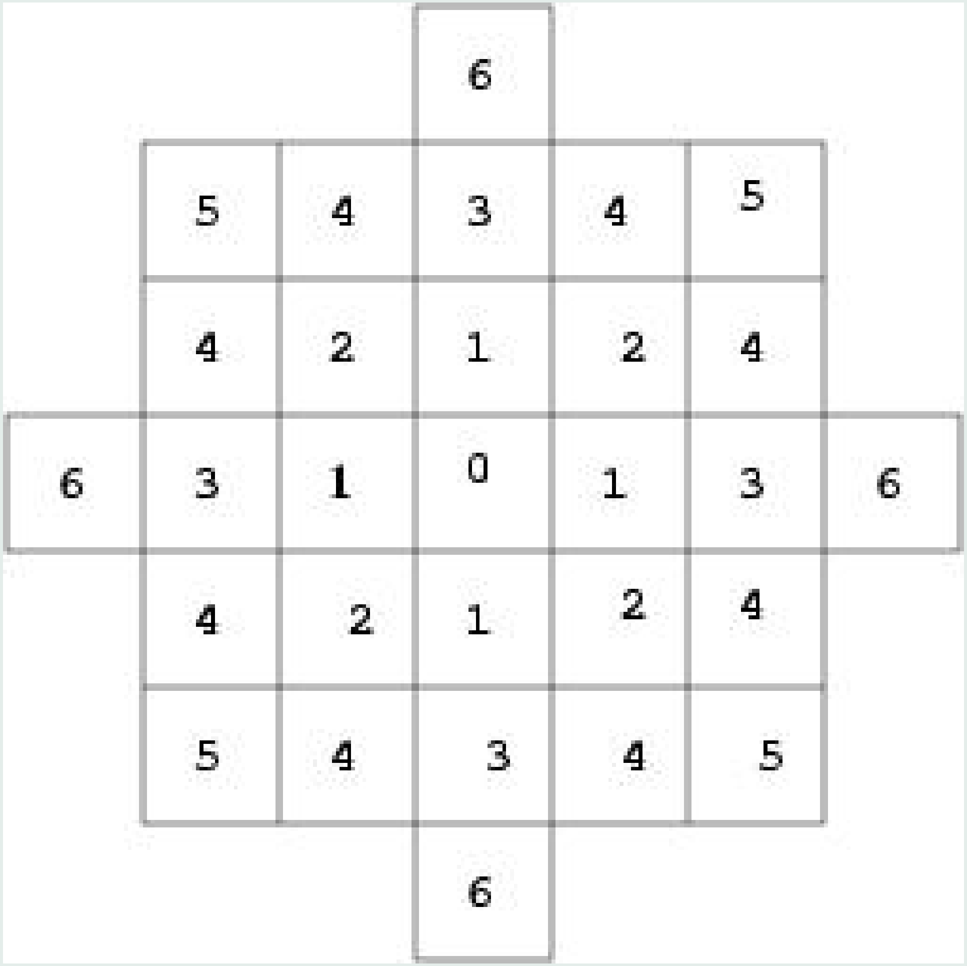
$$P(X_{i,j} = x_{i,j} | X_{k,l} = x_{k,l}), (k,l) \neq (i,j) = P(X_{i,j} = x_{i,j} | X_{k,l} = x_{k,l}) \ , k,l \in \eta_{i,j} \quad (1)$$

$$P(X = x) = \frac{1}{Z} exp^{-U(x)/T} \quad (2)$$

$$U(x) = \sum_{c \in \mathcal{C}} V_c(x) \quad (3)$$

$$\begin{aligned} U(x) \ = \ & \textstyle\sum_{c \in \mathcal{C}} [\mu(x_{i,j} - x_{i,j-1})^2 (1 - v_{i,j}) + (x_{i,j} - x_{i,j+1})^2 (1 - v_{i,j+1}) + \\ & (x_{i,j} - x_{i-1,j})^2 (1 - h_{i,j}) + (x_{i,j} - x_{i+1,j})^2 (1 - h_{i+1,j})] + \\ & \gamma [v_{i,j} + h_{i,j} + v_{i,j+1} + h_{i+1,j}] \end{aligned} \quad (4)$$

$\mathcal{C}$ is set of *cliques*



A neighborhood system

Observation Model

$$Y = \mathbf{H}X + N \tag{5}$$

$$X = [X_{0,0}\ X_{0,1} \ldots X_{M-1,M-1}]^T \tag{6}$$

where,

$$\begin{gathered} X_{i,j} = [x^r(i,j) \quad x^g(i,j) \quad x^b(i,j)\ ]^T \\ 0\ \leq\ i,j\ \leq\ M-1 \end{gathered} \tag{7}$$

Y and N are similarly defined

$$\mathbf{H} = \begin{pmatrix} H_\xi & H_1 & H_1 & 0 & \dots & H_1 & H_1 \\ H_1 & H_\xi & H_1 & H_1 & \dots & 0 & H_1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ H_1 & H_1 & 0 & \dots & H_1 & H_1 & H_\xi \end{pmatrix} \tag{8}$$

where

$$H_\xi = \begin{pmatrix} \bar{H}_\xi & \bar{H}_1 & \bar{H}_1 & 0 & \dots & \bar{H}_1 & \bar{H}_1 \\ \bar{H}_1 & \bar{H}_\xi & \bar{H}_1 & \bar{H}_1 & \dots & 0 & \bar{H}_1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{H}_1 & \bar{H}_1 & 0 & \dots & \bar{H}_1 & \bar{H}_1 & \bar{H}_\xi \end{pmatrix} \tag{9}$$

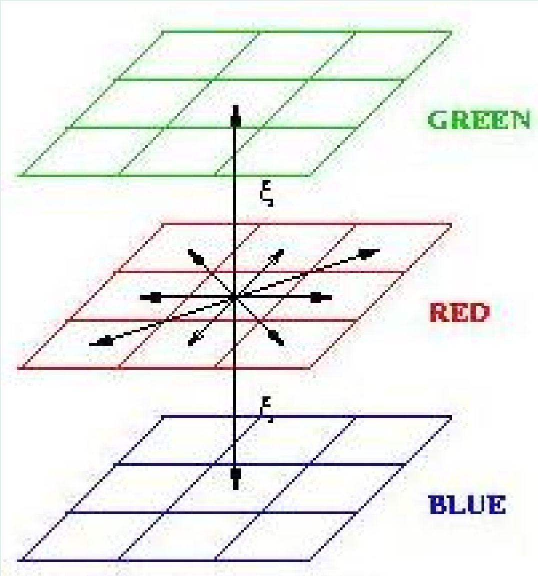
$$\bar{H}_\xi = \begin{pmatrix} 1 & \xi & \xi \\ \xi & 1 & \xi \\ \xi & \xi & 1 \end{pmatrix} \tag{10}$$

$\bar{H}_1$ is:

$$\bar{H}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{11}$$

The structure of H_1 will be same as that of H_ξ with $\bar{H}_\xi$ replaced by $\bar{H}_1$ as :

$$H_1 = \begin{pmatrix} \bar{H}_1 & \bar{H}_1 & \bar{H}_1 & 0 & \dots & \bar{H}_1 & \bar{H}_1 \\ \bar{H}_1 & \bar{H}_1 & \bar{H}_1 & \bar{H}_1 & \dots & 0 & \bar{H}_1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{H}_1 & \bar{H}_1 & 0 & \dots & \bar{H}_1 & \bar{H}_1 & \bar{H}_1 \end{pmatrix} \tag{12}$$



Color Image Interchannel Blurring

Energy Function

a posteriori energy function given by

$$U_p(x) = U(x) + \frac{\|n\|_2^2}{2\sigma_e^2} \tag{13}$$

$$\begin{aligned} U_1(x^c, h^c, v^c) \; &= \textstyle\sum_{i,j} \mu[(x^c_{i,j} - x^c_{i,j-1})^2(1 - v^c_{i,j}) + (x^c_{i,j} - x^c_{i-1,j})^2(1 - h^c_{i,j})] \\ &+\gamma[h^c_{i,j} + v^c_{i,j}] \; for \;\; c = r, g, b \end{aligned} \tag{14}$$

This is called Non-Interaction (NI) or Linear model

$$\begin{aligned} U_2(x, l, v) \; &= \textstyle\sum_{c=1}^{3} \sum_{d=1}^{3} \sum_{i,j} \\ &\mu[(x^c_{i,j} - x^c_{i-1,j})(x^d_{i,j} - x^d_{i-1,j}) \\ &(1 - l^c_{i,j})(1 - l^d_{i,j}) \\ &+(x^c_{i,j} - x^c_{i,j-1})(x^d_{i,j} - x^d_{i,j-1}) \\ &(1 - v^c_{i,j})(1 - v^d_{i,j})] \\ &+\gamma[l^c_{i,j} + v^c_{i,j} + l^d_{i,j} + v^c_{i,j}] \end{aligned} \tag{15}$$

This is called First Order Interchannel Interaction (FOII) model

Results

$$PSNR = 10 \log \frac{255^2}{\|x - \hat{x}\|^2} \qquad (16)$$

	Degraded image			NI			FOII		
ξ	R	G	B	R	G	B	R	G	B
0.4	14.77	14.64	15.05	16.31	16.99	15.10	20.15	17.28	17.82
0.6	14.65	14.51	14.91	18.78	16.28	16.45	20.15	17.28	17.82
0.8	14.57	15.68	16.07	18.88	18.88	19.46	19.89	20.44	21.42
1.0	14.36	15.37	15.80	17.73	18.45	19.29	19.38	19.97	20.72
1.25	14.14	15.19	15.66	17.53	18.17	18.10	18.71	19.44	19.29
1.5	14.12	15.14	15.65	17.09	17.82	17.85	18.30	18.89	19.33

PSNR values for synthetic image

	Degraded image			NI			FOII		
ξ	R	G	B	R	G	B	R	G	B
0.25	23.77	19.95	20.35	24.45	20.52	20.97	24.75	20.78	21.02
0.5	23.41	18.98	19.92	24.14	20.41	20.98	24.65	20.56	21.23
0.75	23.60	19.75	19.99	23.88	20.41	20.96	24.41	20.62	21.07
1.0	23.22	19.70	19.92	23.36	20.37	20.80	23.96	20.50	20.84
1.25	22.93	19.34	19.68	23.14	20.16	20.72	23.54	20.24	20.64
1.5	22.50	19.20	20.07	22.69	19.79	20.16	20.67	20.43	20.56

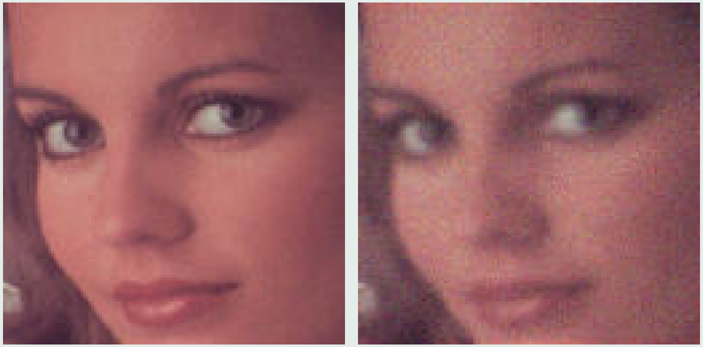
SNR Values for Lisa image

Methodology.	R	G	B
Degraded Img.	22.93	20.51	20.86
Linear	23.63	21.58	22.12
FOII	24.47	23.32	23.08

Lisa image degraded in YIQ coordinates.

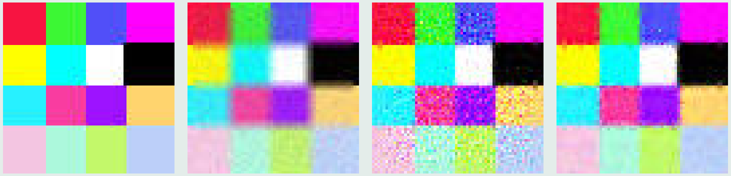
Methodology.	R	G	B
Degraded Img.	22.72	19.90	19.01
Linear	23.51	21.22	21.57
FOII	24.42	23.18	22.94

Lisa image degraded in Ohta's coordinates

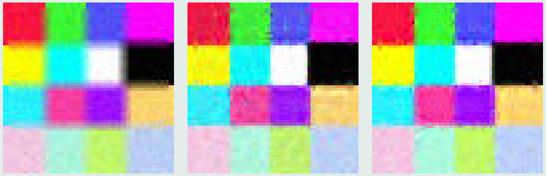


Lisa image: Original and Degraded in YIQ coordinates.

Lisa image: Original and Degraded in YIQ coordinates.



Synthetic image: Original and Degraded in Ohta coordinates.
Restored using NL and FOII model



Synthetic image: Degraded in YIQ coordinates.
Restored using NL and FOII model

Observations

- Proposed FOII model performs better than NI model, for different values of ξ
- for $\xi = 0$ performance of NI and FOII are similar; FOII giving slightly better SNR improvements.
- FOII works satisfactorily even when ξ in unknown.
- FOII can be considered *partially blind* restoration model

Limitations

Simulated annealing converges very slowly
Not suited for highly textured images

Parrot image: Original and Degraded.
Restored using NL and FOII model

Still Image Zooming

- Generation of high resolution image from the observed low resolution image
- Pad zeros to intermediate values and then pass it through a filter

Some of the Existing Methods

- Linear interpolation
- Pixel replication
- Sinc
- Spline

Still Image Zooming Using MRF

Assume that the given low resolution image Y is modeled as

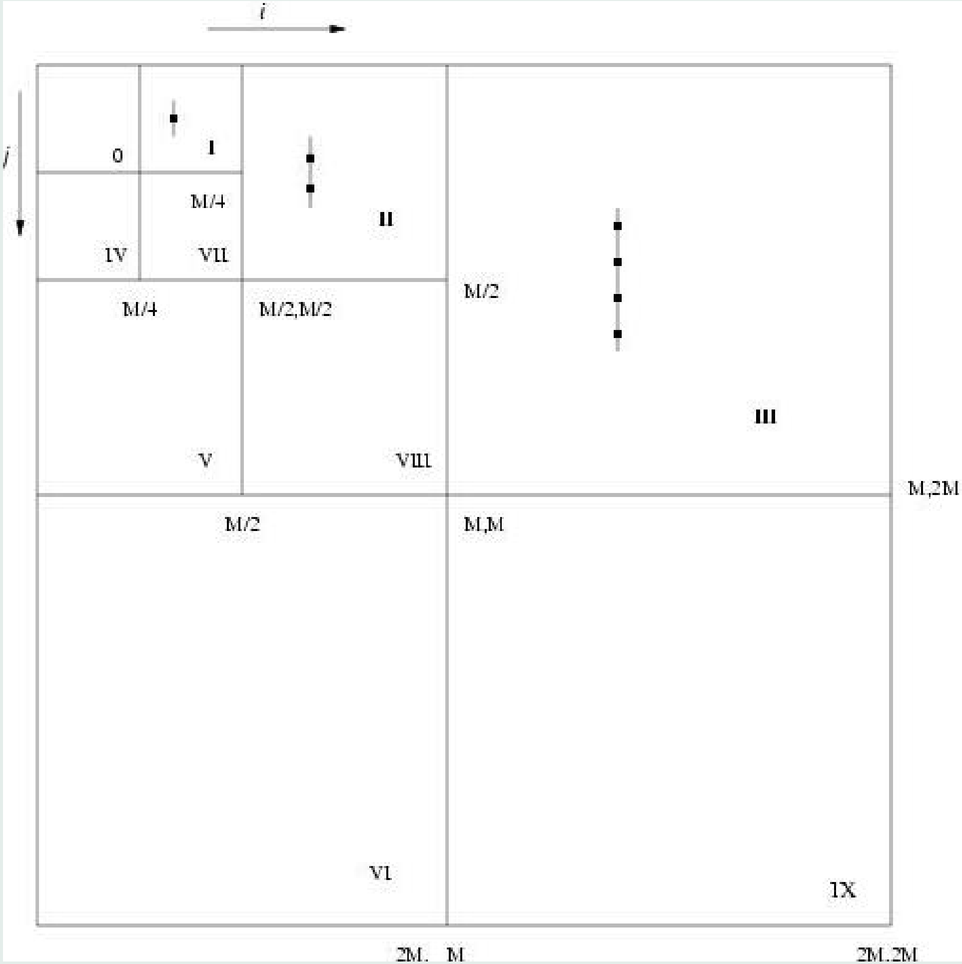
$$Y = \mathbf{D}X + V \tag{17}$$

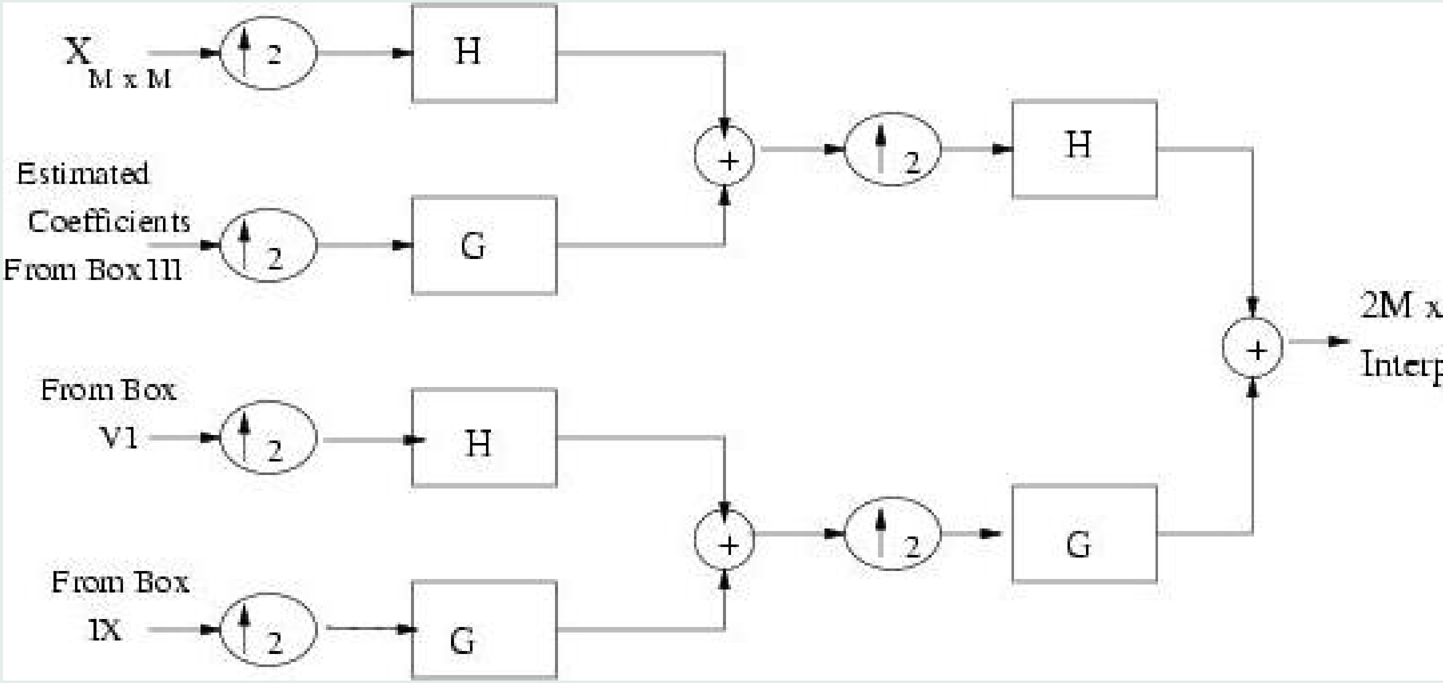
Structure of $\mathbf{D}$ is:

$$\mathbf{D} = \begin{pmatrix} \mathbf{C} & \mathbf{C} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \mathbf{C} & \mathbf{C} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \mathbf{C} & \mathbf{C} \end{pmatrix} \tag{18}$$

$$\mathbf{C} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_1 & c_2 & c_3 & c_4 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_3 & c_4 & 0 & 0 & 0 & 0 & c_1 & c_2 \end{pmatrix} \tag{19}$$

Still Image Zooming Using MRA




X
M x M
2
H
Estimated
Coefficients
From Box 111
2
G
+
2
H
From Box
V1
2
H
From Box
IX
2
G
+
2
G
+
2M x
Inter

MRA Formulation

Properties used:

- If a wavelet coefficient at a coarser scale is insignificant with respect to a given threshold θ, then all wavelet coefficients of the same orientation in same spatial location at finer scales are likely to be insignificant with respect to that θ.
- In a multiresolution system, every coefficient at a given scale can be related to a set of coefficients at the next coarser scale of similar orientation.

Properties used:

- The approximation signal at a resolution 2^{j+1} contains all the necessary information to compute the same signal at a lower resolution 2^j. This is the causality property.
- An approximation operation is similar at all resolutions. The spaces of approximated functions should thus be derived from one another by scaling each approximated function by the ratio of their resolution values.

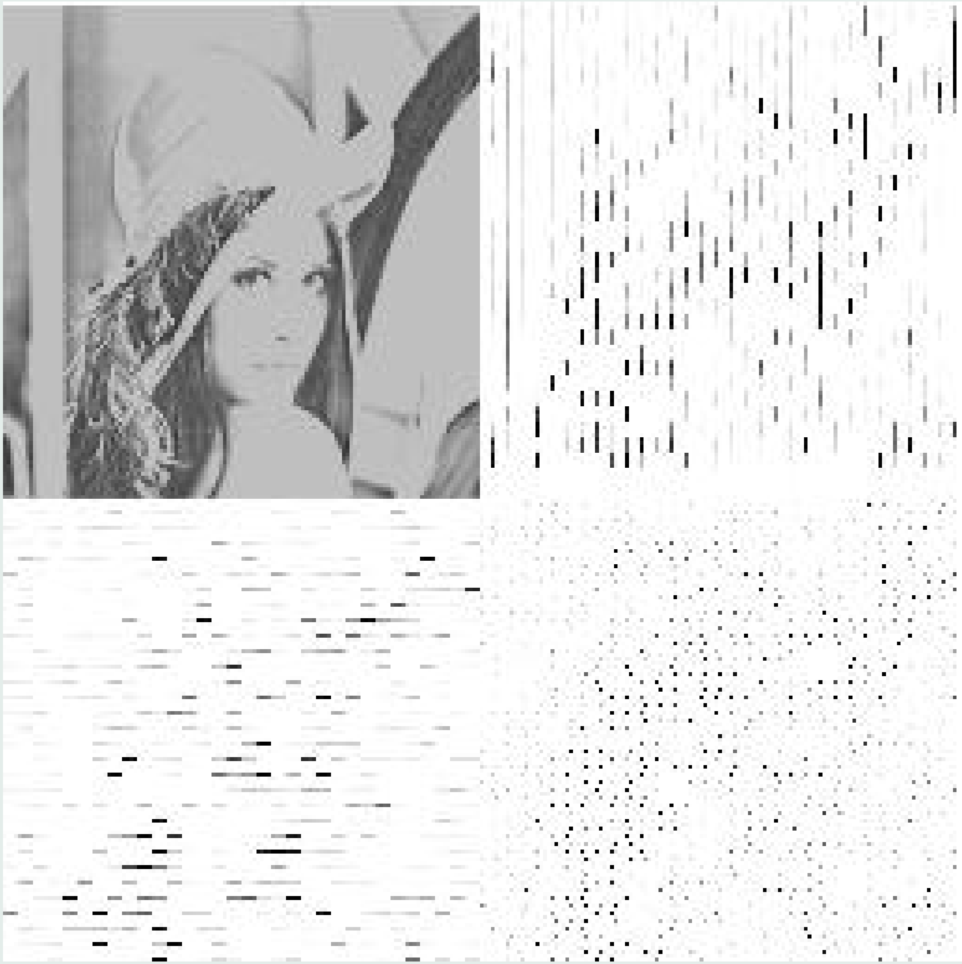
We define $D_{(.)}(i, j)$ as (between boxes I and II):

$$D_1(i, j) = \frac{d_2(i, j)}{d_1(\lfloor i/2 \rfloor, \lfloor j/2 \rfloor)} \tag{20}$$

$$D_2(i, j) = \frac{d_2(i, j + 1))}{d_1(\lfloor i/2 \rfloor, \lfloor (j + 1)/2 \rfloor)} \tag{21}$$

These $D_{(.)}(i, j)$ values are used to estimate coefficients $\hat{d}$ at the finer scale (box III).

$$\begin{aligned} \hat{d}(2i, 2j) &= D_1(i, j) d_2(i, j)(1 - l_{d(i,j)}) \\ \hat{d}(2i, 2j + 2) &= D_2(i, j) d_2(i, j + 1)(1 - l_{d(i,j+1)}) \end{aligned} \tag{22}$$

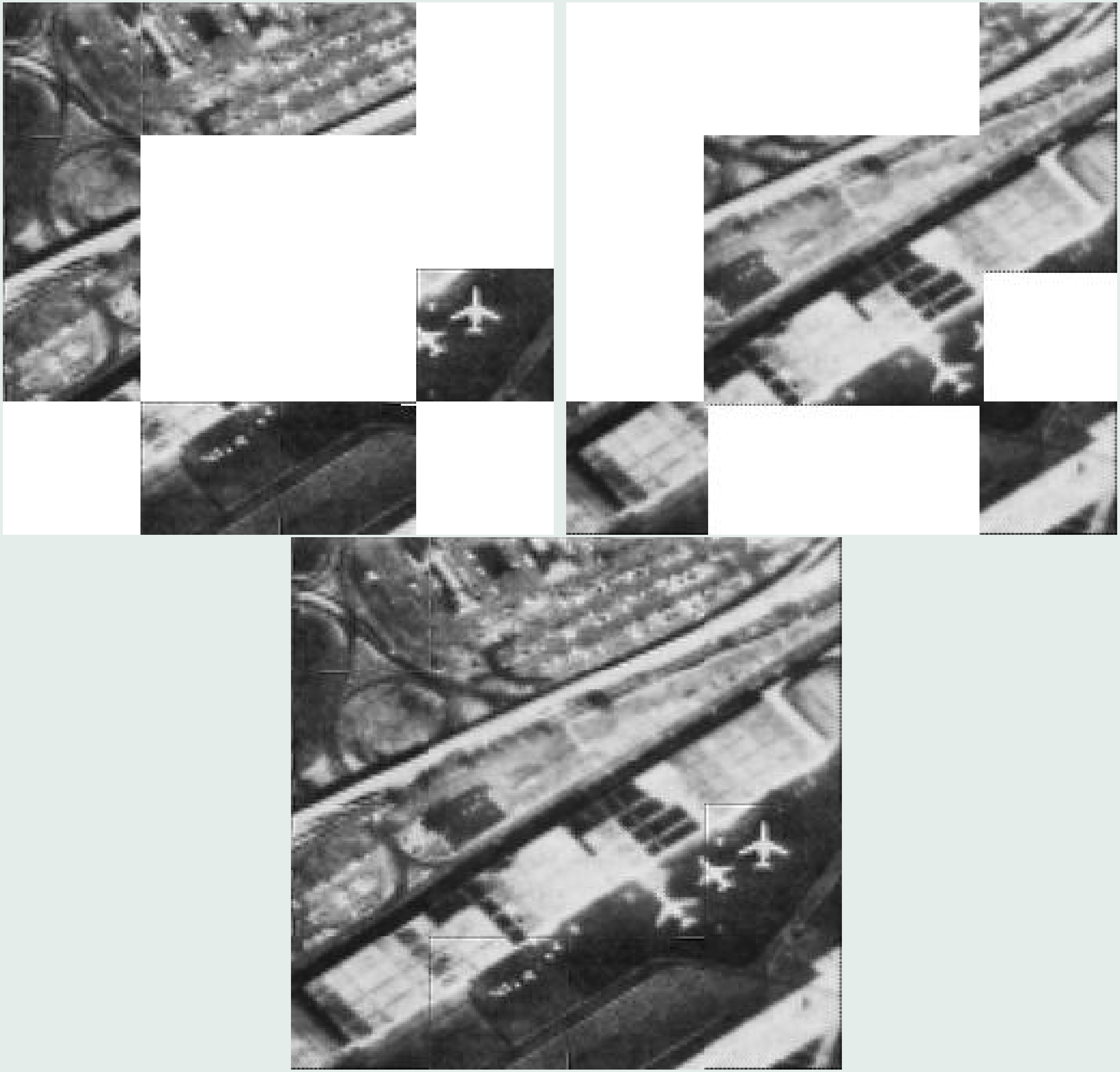


Estimated wavelet coefficients for Lena image

MRA, Spline, Scaling function and MRF based zoomed boat image

Joint MRA And MRF Method

- Combine the MRA and MRF approaches.
- Estimate variance for blocks of data
- Estimate the mean of these variance(MOV)
- Use MOV as the measure of smoothness and interpolate smooth part using MRF and "rough" parts using MRA



MRF, MRA and Joint approach

PSNR Values:

Image	*Spline*	*Sinc*	*MRF*	*MRA*	*Joint*	*Scal. Fn.*
Boat.	24.97	24.49	25.91	29.21	26.92	25.72
Airport	23.82	22.88	24.48	26.98	25.55	24.44
Lena	25.73	24.14	26.69	29.80	28.18	23.49
bird	30.89	20.00	29.78	33.25	31.80	21.41
Einstein	28.28	19.18	27.76	30.17	28.87	24.51

Color image Zooming

- Get the YIQ component of the color image
- Interpolate the Y component using MRA
- I and Q components are interpolated using linear interpolation
- Convert back to RGB

Original Suzie image

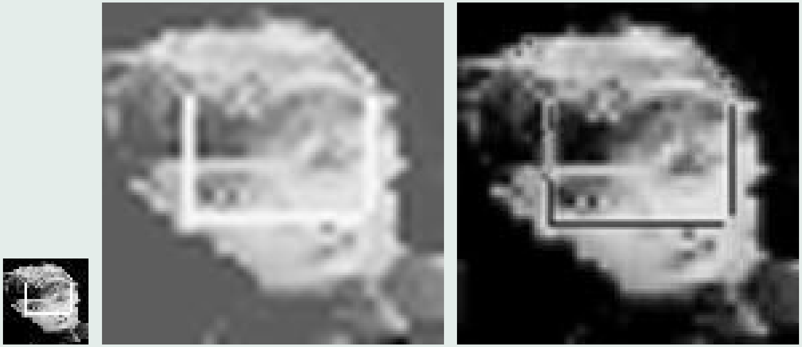
Zoomed Suzie image using MRA and spline

Observations

- MRA gives sharper images with a little blocky images
- DAUB4 was found to be optimal. Haar gives more blocky images and higher Daub smoothens the edges.
- Visually scaling function based gives better results. But this is compute intensive

Limitations

- MRA method does not work satisfactorily for images with sudden transition between black and white. Results in spurious edges.
- Performance is not satisfactory for zooming beyond $4\times$.



Face image: Original, Spline and MRA interpolated($4\times$)

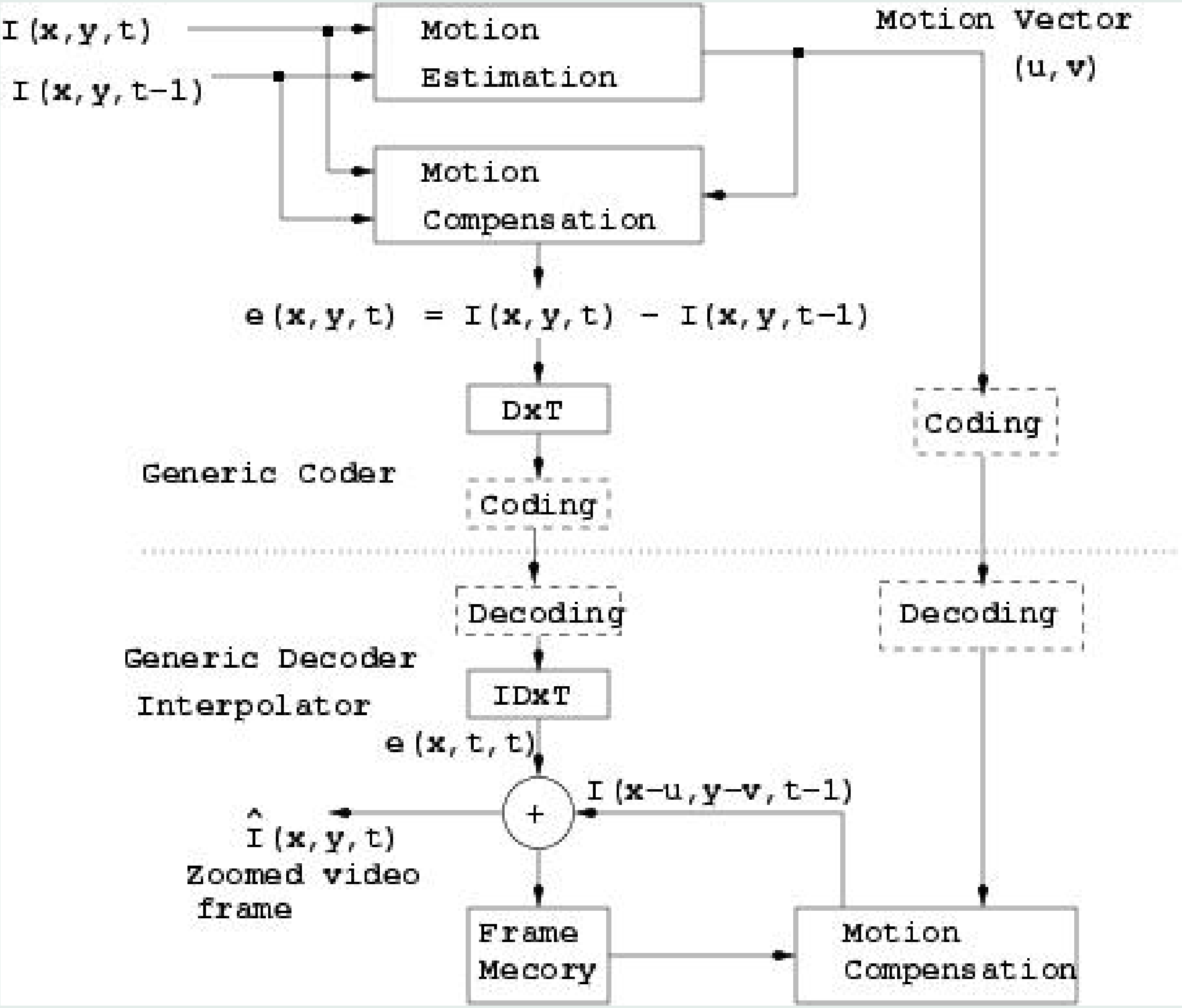
Compressed Video Zooming:

Zoom the given video in compressed domain, by interpolating motion vectors.

- Motivation
- Video Coding and Compression
- Proposed technique
- Extension to MRME
- Results and discussions

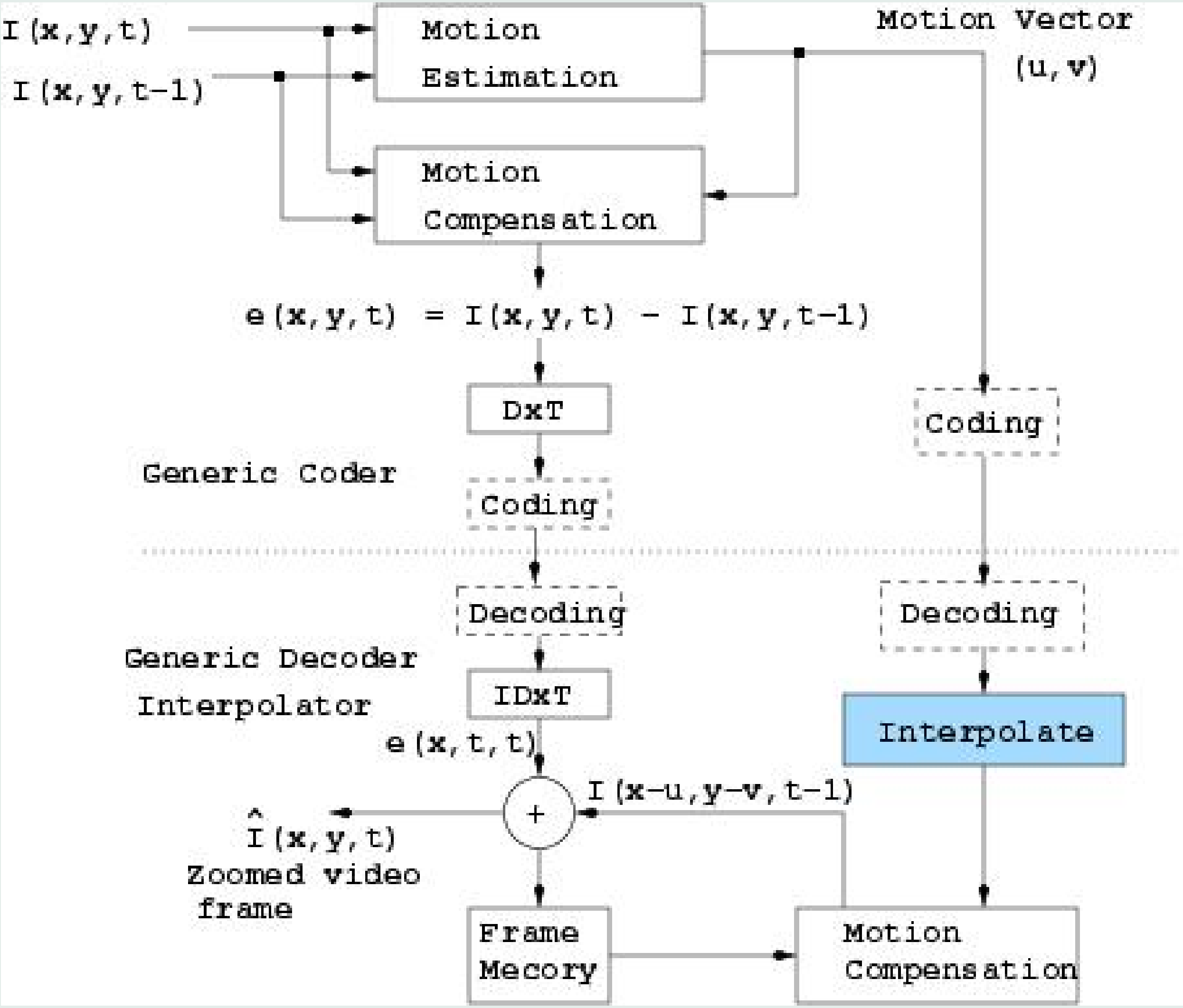
Motivation

- Bit rate
- Channel capacity
- Picture-in-picture TV
- HDTV



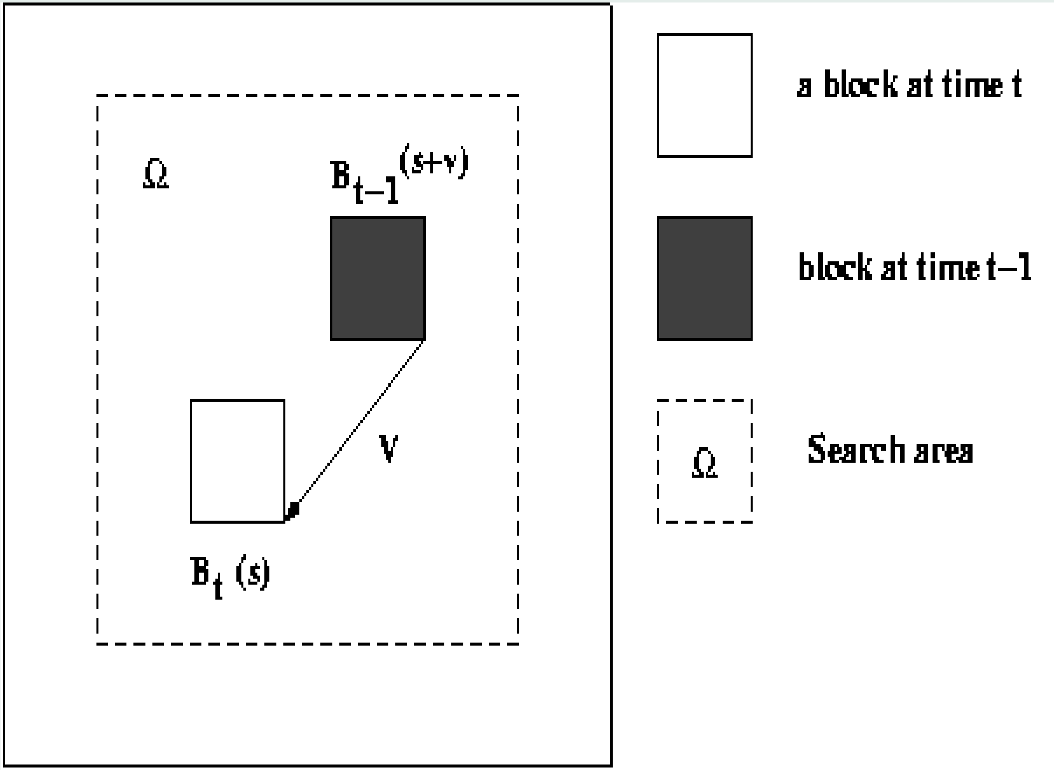
Generic Video Coder/Decoder

Compressed Video Zooming

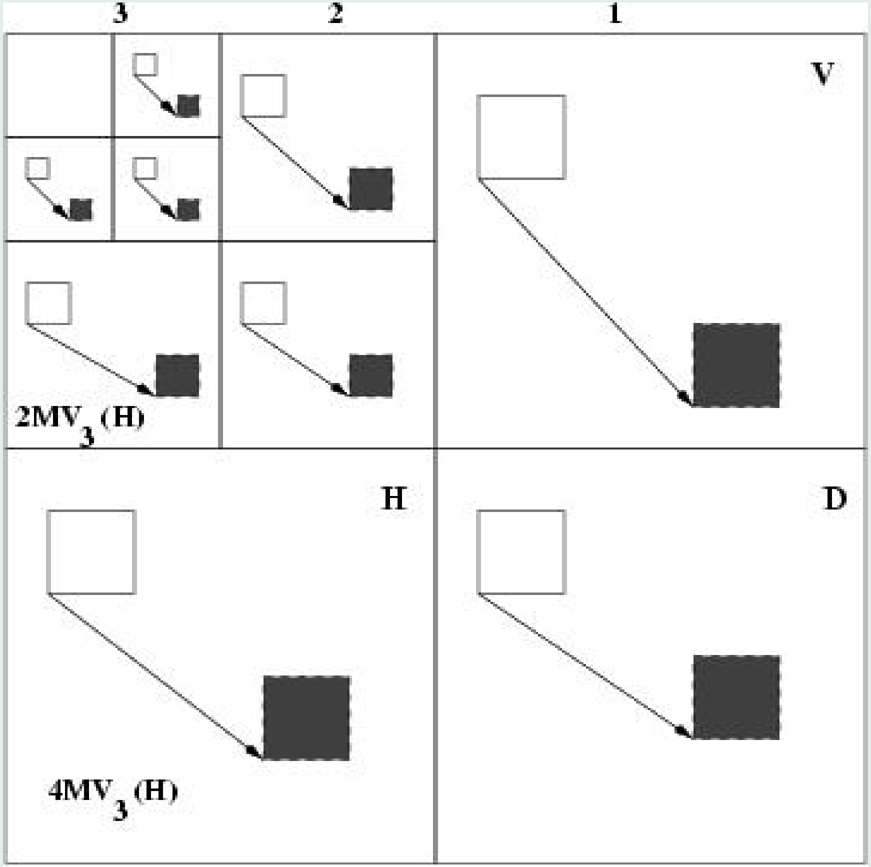


Proposed Method (shaded block)

Motion Estimation

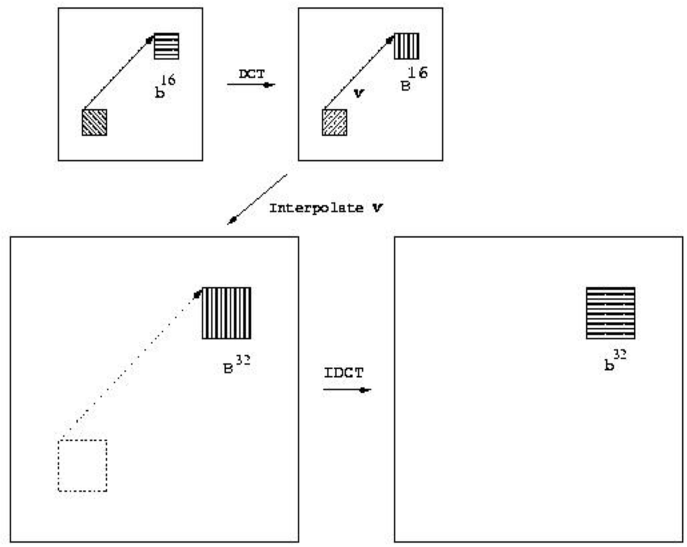


Motion Estimation



DWT Motion Estimation

Proposed Method



Motion Vector Interpolation

Define

$$^{k}\mathbf{B}_p = \sum_i \sum_j {}^{k}B_p^{16}(i,j) \qquad\qquad i,j \in k \tag{23}$$

Motion vector w such that $w = (w_x, w_y)$ is calculated as

$$w = {}_{w\in\Omega}^{arg\ min} |^{k}\mathbf{B}_p -^{k} \mathbf{B}_{p-1}| \tag{24}$$

$$^{k}\epsilon_p = {}^{k}\mathbf{B}_p -^{k} \mathbf{B}_{p-1} \tag{25}$$

The new block locations are now evaluated as

$$k\hat{B}_p^{16} = (^{k}B_{p-1}^{16} +^{k} \epsilon).\hat{w} \tag{26}$$

where ϵ is the *error* and
$\hat{w}$ is the new *interpolated* motion vector value

For DCT, ${}^{k}B_p^{32}$ is evaluated as:

$$ {}^{k}B_p^{32} = \begin{bmatrix} {}^{k}\hat{B}_p^{16} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{27} $$

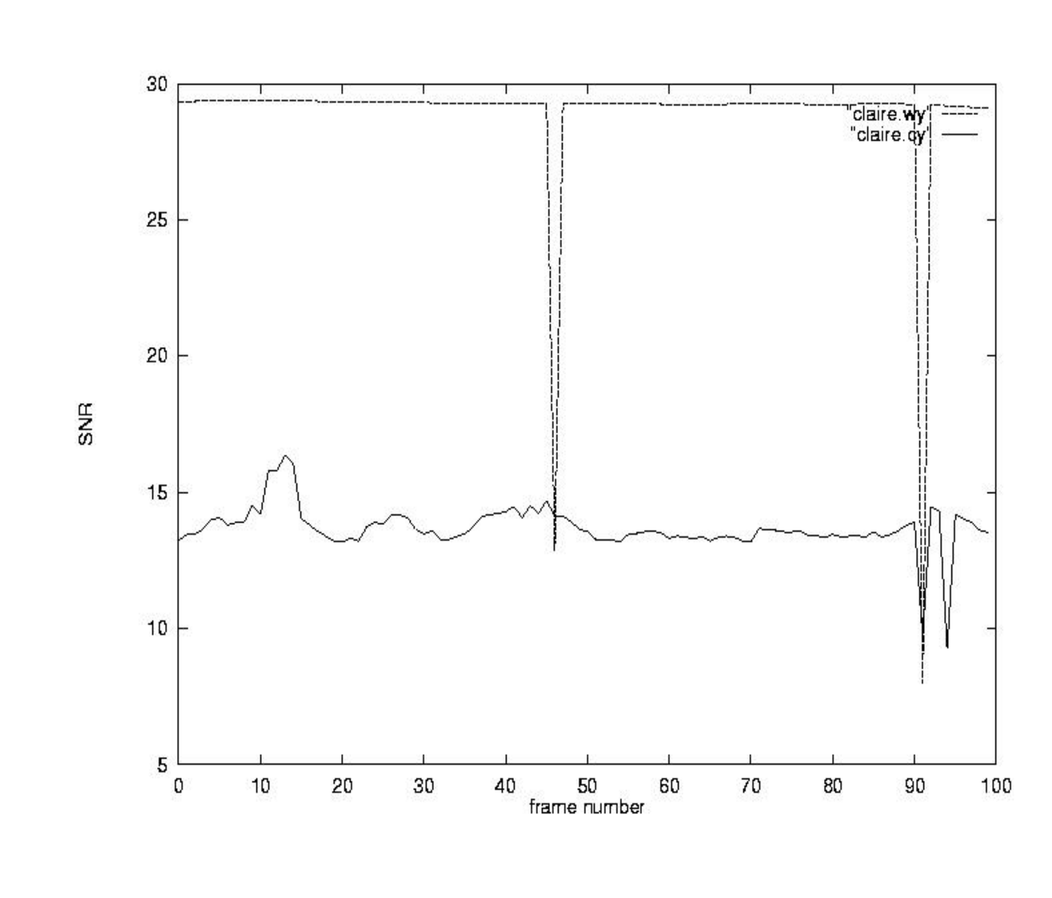
For DWT, ${}^{k}B_p^{32}$ is evaluated as:

$$ {}^{k}B_p^{32} = \begin{bmatrix} {}^{k}\hat{B}_p^{16} & {}^{k}\hat{B}_p^{16}(V) \\ {}^{k}\hat{B}_p^{16}(H) & {}^{k}\hat{B}_p^{16}(D) \end{bmatrix} \tag{28} $$

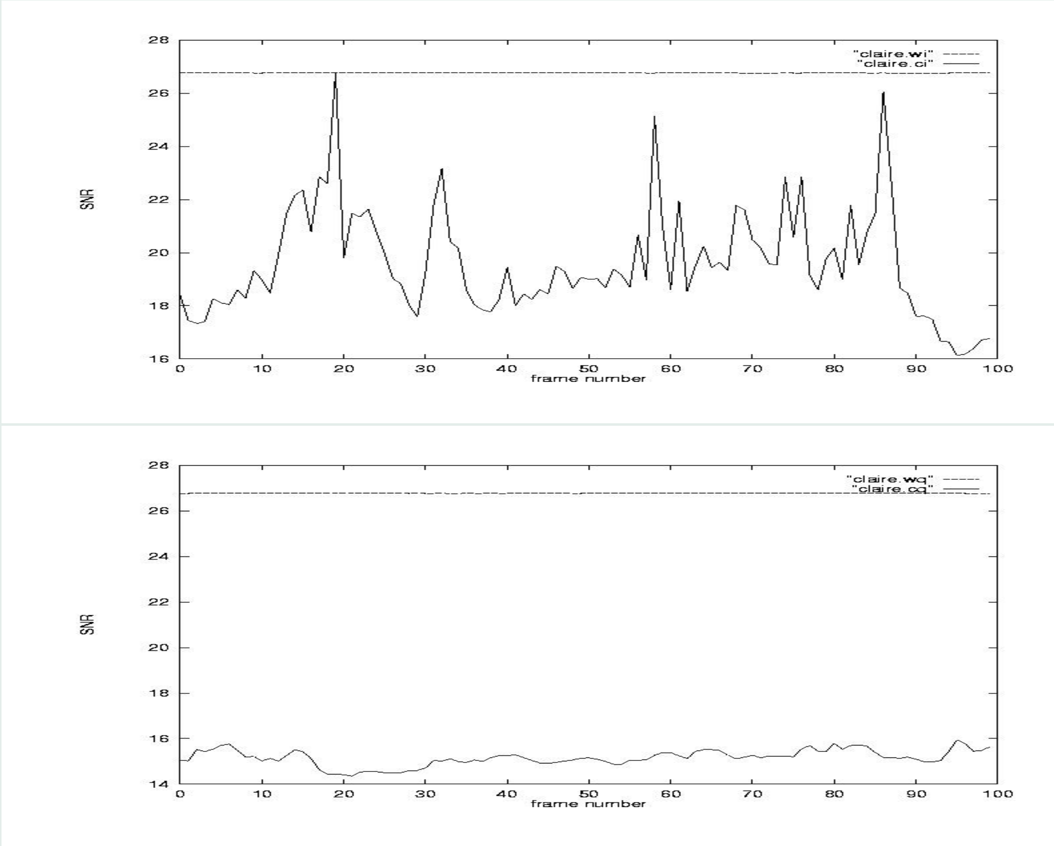
Performance Measure

- Given video is zoomed using DCT
- Each frame is divided into 8×8 block and these blocks are extended to 16×16 directly by zero-padding, without using motion information
- We use this video (represented as $\tilde{X}$) as the reference performance is measured for each frame p as

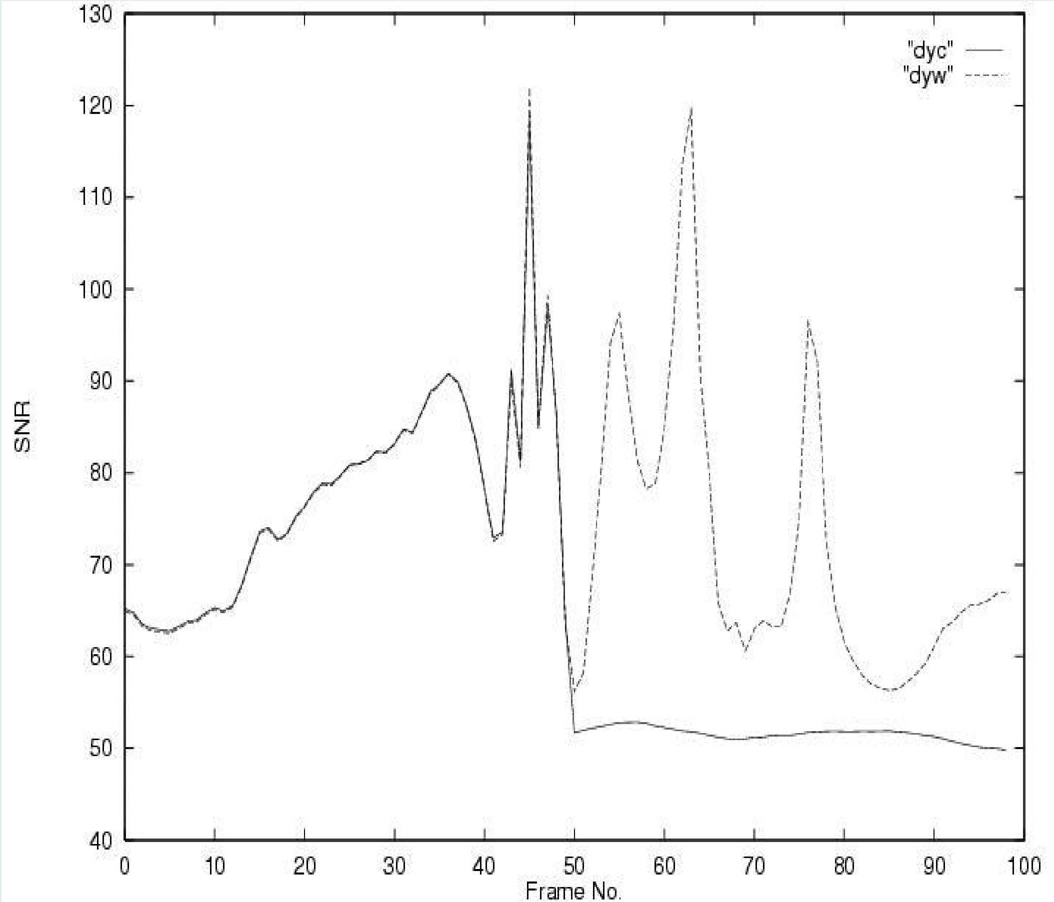
$$Error = \sum_{i,j} \tilde{X} - \hat{X} \quad i, j \in p$$



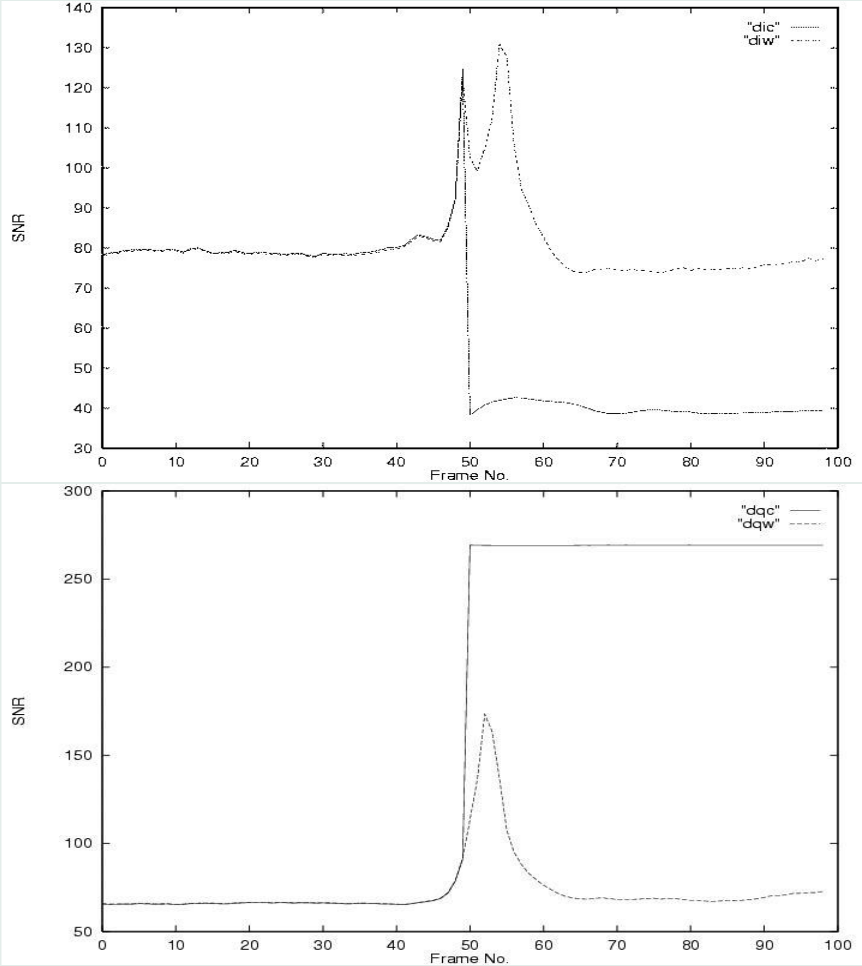
Error Plots for Claire image: Y component
Y axis Error, X axis frame No.



Error Plots for Claire image: Y component, I component and Q component
Y axis Error, X axis frame No.



SNR Plots for Suzie image: Y component Dotted:DWT and Thick:DCT
Y axis SNR, X axis frame No.



SNR Plots for Suzie image: I and Q component Dotted : DWT and thick:DCT
Y axis SNR, X axis frame No.

Suzie video clip. Left col:Original, Middle DCT, right DWT

Multiresolution Motion Estimation

- Estimation motion vectors hierarchically from lower to higher resolution sub-images
- Exploits cross correlation among each layer of wavelet pyramid
- *assumption:* Motion vectors at different levels are highly correlated

- The Sum of Absolute Displacement (SAD)measure is used as

$$SAD(B_i(o,p), B_{i-1}(o+\delta o+\delta p)) = \sum_{m=-o/2}^{o/2} \sum_{n=-p/2}^{p/2} |B_i(o+m, p+n) - B_{i-1}(o+m+\delta o, p+n+\delta p)| \tag{29}$$

- The MRME scheme estimates motion vector as:

$$MV_i(H)(2^{3-i}o, 2^{3-i}p) = 2^{3-i}MV_3(H)(o,p) \quad for\ i = 1, 2$$

- A modified error criterion is defined as:

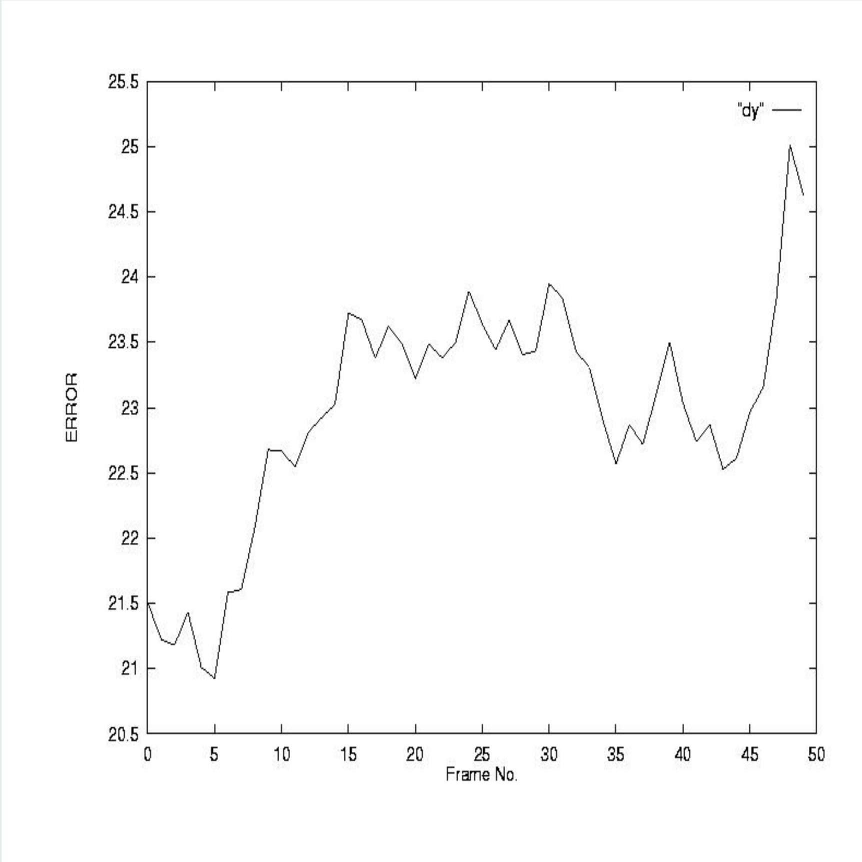
$$\begin{aligned} &MSAD(B_i(o,p), B_{i-1}(o+\delta o, p+\delta p)) \\ &= SAD(B_i(o,p), B_{i-1}(o+\delta o, p+\delta p)) \\ &+SAD(B_i(2*o, 2*p), b_{i-1}(2*(o+\delta o), 2*(p+\delta p))) \\ &+SAD(B_i(4*o, 4*p), b_{i-1}(4*(o+\delta o), 4*(p+\delta p))) \end{aligned}$$

For intra frame:

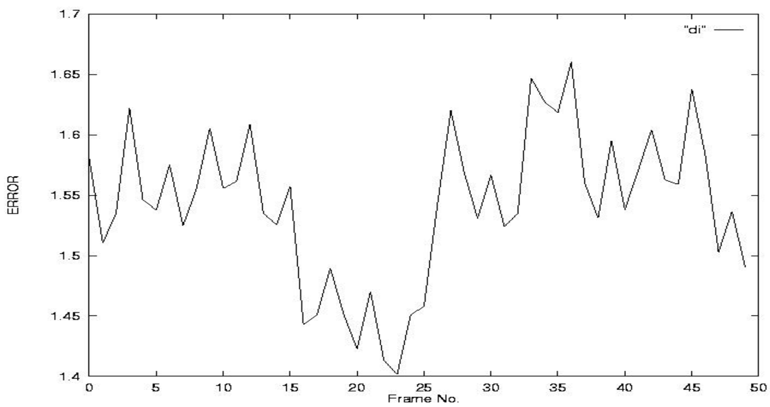
```
Decode and de-quantize coded bit stream ;
Interpolate wavelet coefficients ;
Take 2x IDWT
```

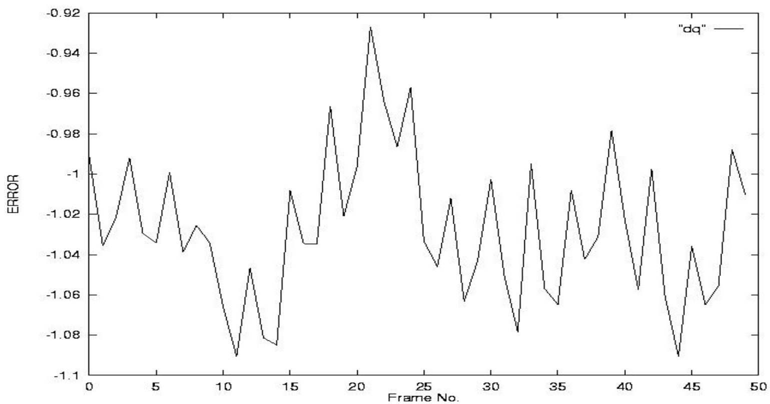
For inter-frame:

```
Decode and de-quantize coded bit stream ;
Generate the motion vectors ;
Interpolate motion vectors ;
Take 2x IDWT ;
```



Error Plots for Y component of Claire image: 0.05 bits/pixel MRME Method
Y axis Error, X axis frame No.





Error Plots for Claire image: 0.05 bits/pixel MRME Method
Y axis Error, X axis frame No.

MRME based decompression and zooming.
Top row shows decompressed Claire video after compressing the original video
Bottom row is zoomed version of the same compressed video clip

Video Frame Interpolation

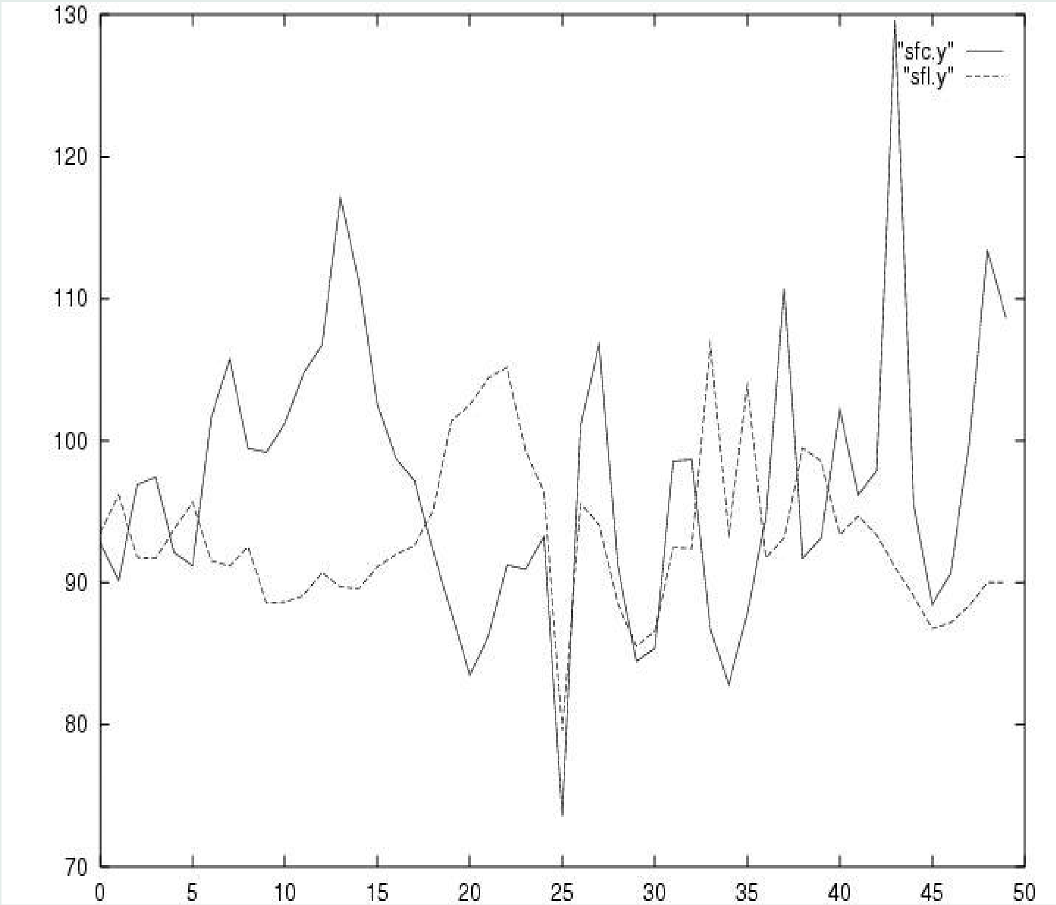
Error is defined as,

$$^{k}\epsilon_p = |\sum_{i=0}^{15}\sum_{j=0}^{15}\{^{k}B_{p+1}^{16}(i,j) -^{k} B_{p-1}^{16}(i,j)\}| \tag{30}$$

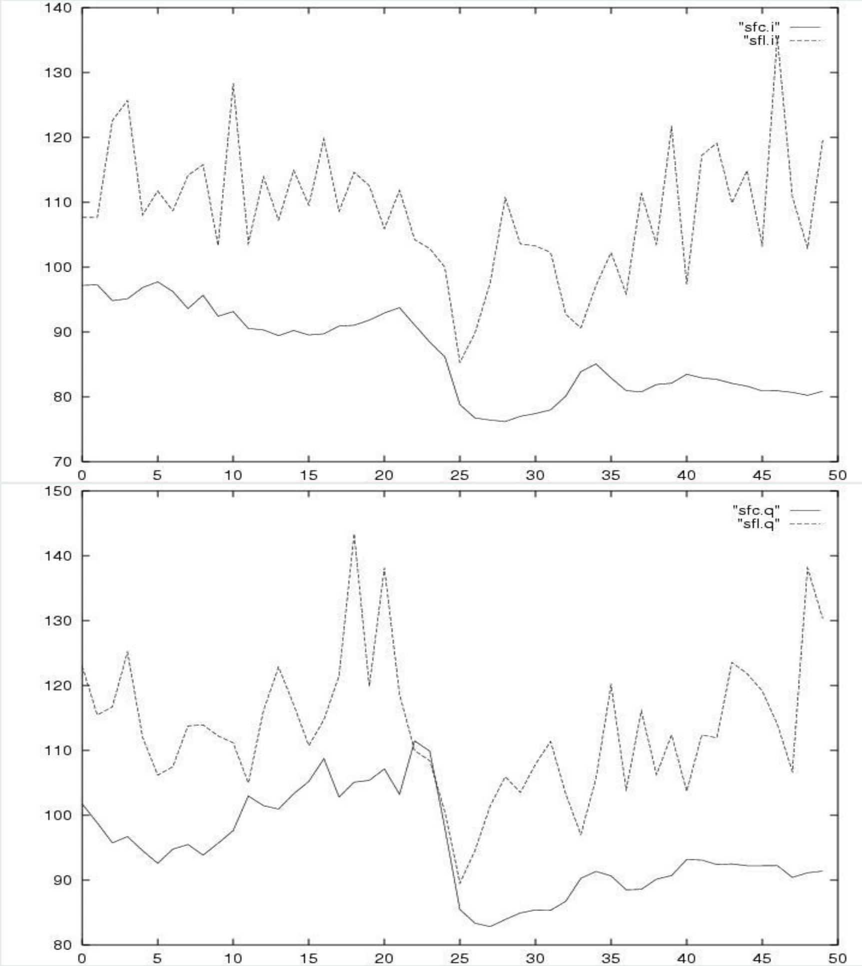
Motion vectors for the p^{th} frame are estimated as $^{k}\hat{w}_x = \lfloor \frac{^{k}w_x}{2} \rfloor$ and $^{k}\hat{w}_y = \lfloor \frac{^{k}w_y}{2} \rfloor$.

The new block location in the p^{th} frame will be,

$$^{k}\hat{B}_p^{16} = (^{k}B_{p-1}^{16}(^{k}\hat{w}_x.x + i,^{k} \hat{w}_y.y + j) +^{k} \epsilon_p \tag{31}$$



SNR Plots for frame interpolated Suzie image: Y component
Dotted:Linear and Thick:Motion vector based
Y axis SNR, X axis frame No.



SNR Plots for frame interpolated Suzie image: I and Q component
Dotted : Linear and thick:Motion vector based
Y axis SNR, X axis frame No.

Frame interpolated Claire video clip. Left col:Original, Middle Linear, right DWT

Observations

- Motion vector interpolation for both DCT and DWT work satisfactorily. Thus the proposed method is independent of compression standards.
- DWT based method gives a better performance than the DCT.
- Quantization will destroy some image details. Thus the zoomed image quality suffers.
- Performance of zooming is dependent on the coder/decoder efficiency.
- Extension to frame interpolation is comparable to linear frame interpolation

Conclusions

- *Color Image Restoration:* Robust, partially blind
- *Image Zooming* A simple algorithm, capable of retaining sharp edges
- *Video Zooming* A novel idea of interpolating motion vectors

Future Directions

- Blind image restoration and different types of blur
- Mathematical justification
- Extend video zoom to to MPEG and transcoder applications
- Downsampling in compressed domain
- A new wavelet basis *tailor made* for zooming applications
-
-
-

THANK YOU