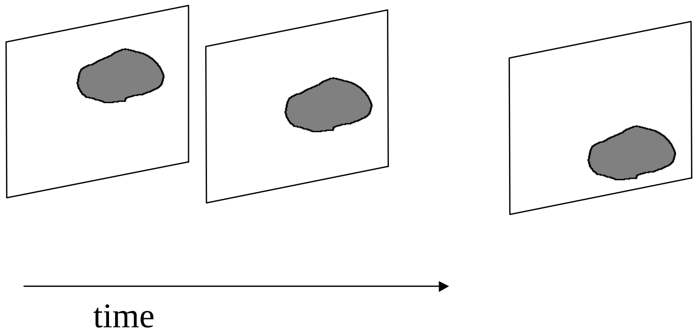
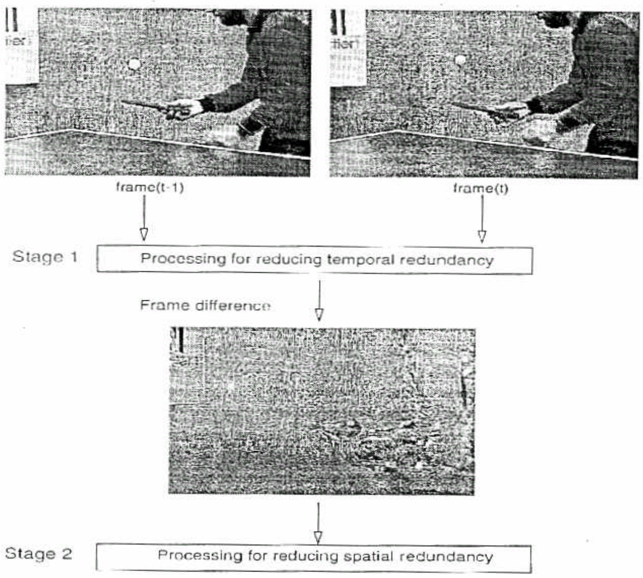
Fundamentals of Lossy Video Compression

Image Sequence Model



Two-stage video coding process



@NTUEE
DSP/IC Lab

Stage 1 : Reducing Temporal Redundancy

- Segment a frame into macroblocks.
- Output energy is increased with the degree of temporal redundancy.
- Interframe coder .

@NTUEE
DSP/IC Lab

Stage 2 : Reducing Spatial Redundancy

- Processing the difference frame (spatially correlated) from stage 1.
- Using DCT coding.
- Intraframe coder.
- Hybrid coding method.

Motion Compensation

- The process of compensating for the displacement of moving objects from one frame to another.
- It is preceded by motion estimation.
- Similar with DPCM.

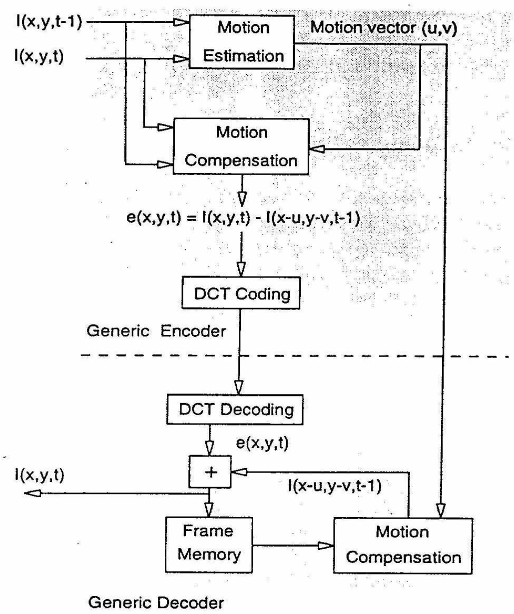


Figure 4.3 Generic hybrid video coder and decoder.

$$k = \sqrt{u^2 + v^2}$$

@NTUEE DSP/IC Lab

Performance Evaluation

Questions before using interframe coding:

- How does intraframe coding compare to interframe coding?
- How do hybrid coding schemes compare to intraframe or interframe coding alone?
- Does the added complexity of motion compensation warrant its use

- 2-D correlated (intraframe)
- Differencing (interframe)
- Motion-compensation (interframe)
- Hybrid (interframe /intraframe)

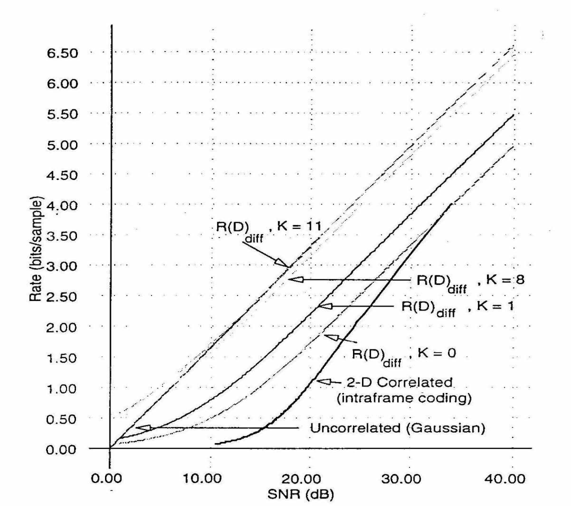
5

@NTUEE DSP/IC Lab

Frame differencing with no motion compensation

- K is the displacement distance
- K>8 offers no improvement, means no correlation
- All cases show that FD is good enough. Intraframe coding does the job best .

$$k = \sqrt{u^2 + v^2}$$



9

@NTUEE DSP/IC Lab

Motion-compensated prediction

- There are 40% improvement even at high motion-estimation error of +/- 1 pixel.
- Different motion precision achieve different performance.

$$R = \frac{1}{2}\log_2 \frac{s_e^2}{D}$$

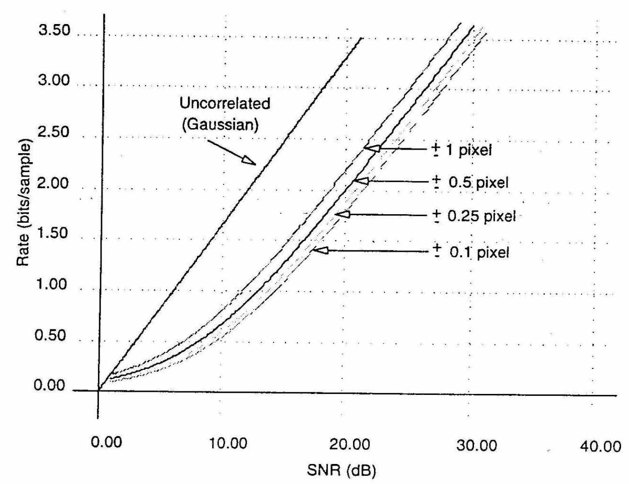


Figure 4.6 Rate-distortion performance for motion-compensated prediction with no intraframe coding.

Motion-compensated prediction v.s. intraframe coding

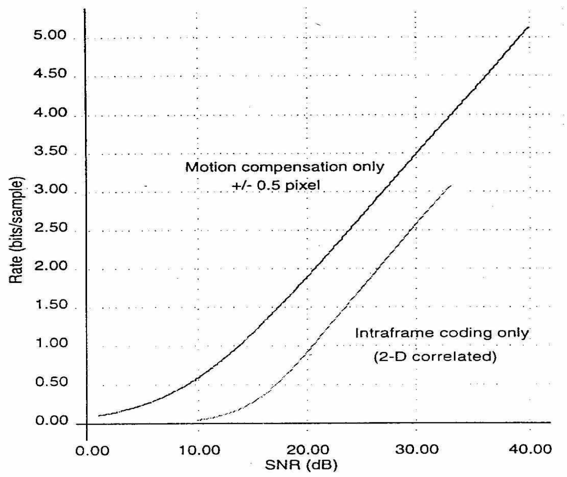


Figure 4.7 Rate-distortion performance for motion-compensated prediction versus intraframe coding.

@NTUEE
DSP/IC Lab

Hybrid Coding: Motion-compensated video followed by intraframe coding

- Hybrid coding can get better results.

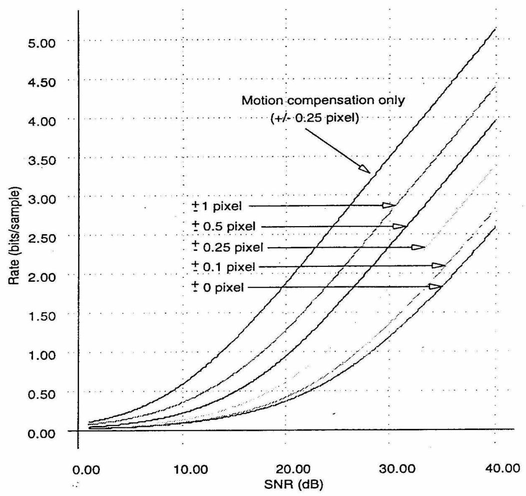
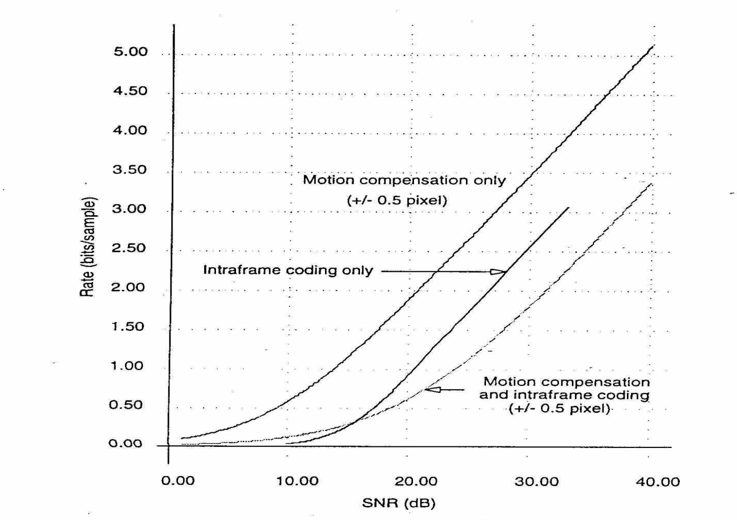


Figure 4.8 Rate-distortion functions for hybrid video coding.

JEE
IC Lab

Rate-distortion functions for Hybrid, interframe, and intraframe video coding



13

@NTUEE
DSP/IC Lab

R(D) Summary

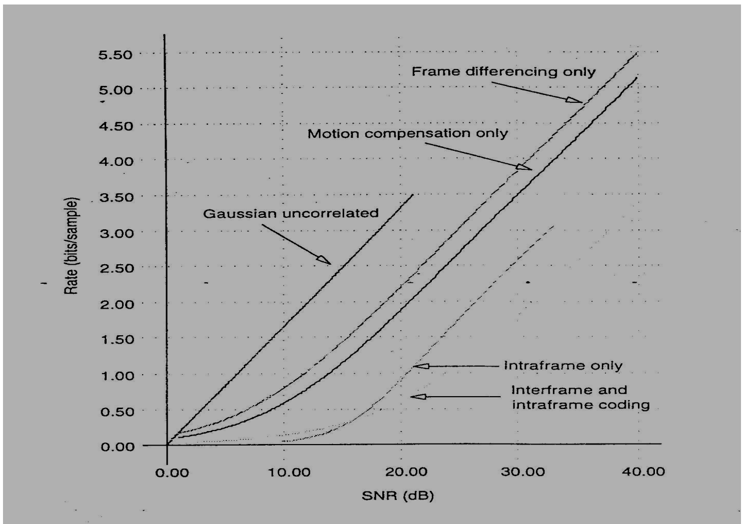


Figure 4.10 Rate-distortion functions for various video coding schemes.

@NTUEE
DSP/IC Lab

Motion-compensated prediction

- Hybrid coding method is quite effective.
- In interframe coding , the motion-estimation problem takes an important role.

Motion Compensation Algorithms in Video Coding

Motion Estimation

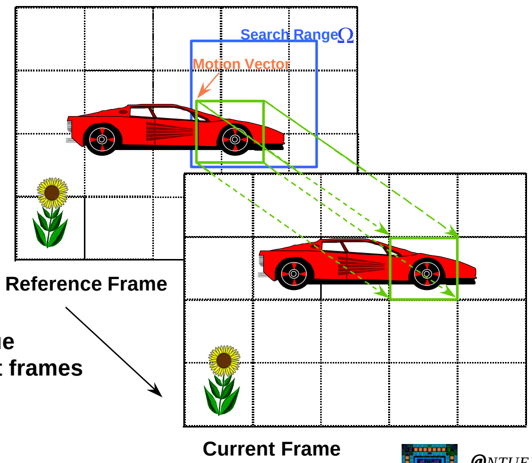
- ❑ ***Objective***
  - ***Predict current frame from neighboring frames***
- ❑ ***Motion Estimation Algorithm***
  - – ***Pixel-based method (Pel-Recursive Algorithm)***
    - • ***Large computation overhead***
  - – ***Block-based method***
    - • ***Regularity and simplicity***
    - • ***Suitable for hardware implementation***
  - – ***Object-based method (content-based)***
- ✡ ***Crucial to make possible a high degree of video compression***

Block-Matching Motion Estimation

$X_t(p,q)$

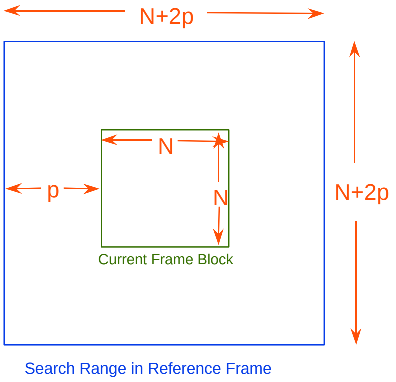
the block at location (p,q) in t-th frame

Motion Vector
$V_t(p.q)$ =(Veci,Vecj)
the location in the search range Ω
that has the maximum correlation value
between blocks in temporally adjacent frames



Factors of Affecting Block-Matching Algorithm(BMA)

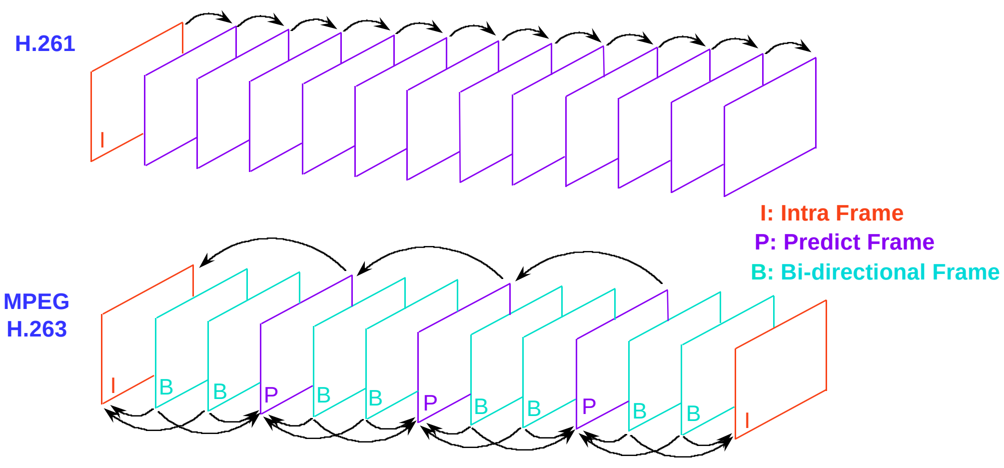
- ☞ ***Search Algorithm***
- ☞ ***Match Criterion***
- ☞ ***Search Range***



19

@NTUEE
DSP/IC Lab

Coding of Moving Pictures



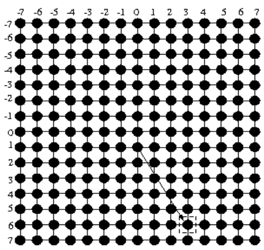
20

@NTUEE
DSP/IC Lab

Brief Review of Previous Algorithms

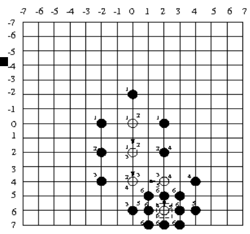
Search Algorithms

- ❒ ***Two-Dimensional Full Search(2DFS)***
  - ***Exhaustive search***
  - ***Extremely Large Computation***
- ❒ ***Fast Search***
  - ***Assumption: Monotonic Distortion Function***
  - ***Reduce computation at the expense of accuracy***
  - ❍ ***2-D Logarithmic Search***
  - ❍ ***Modified 2-D Logarithmic Search***
  - ❍ ***Three-Step Hierarchical Search***
  - ❍ ***Cross Search***
  - ❍ ***One-at-a-time Search***
  - ❍ ***Parallel Hierarchical One-Dimensional Search***
  - ❍ ***One-Dimensional Full Search***

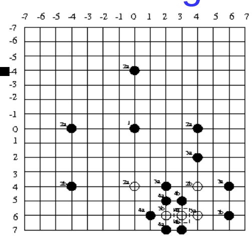


Two-Dimensional Full Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case

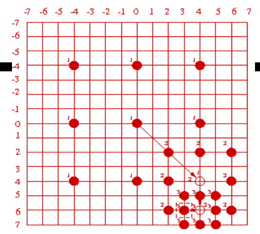
Fast Search Algorithms



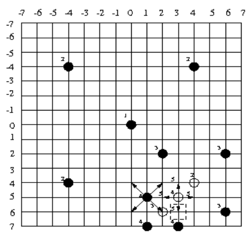
2-D Logarithmic Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case



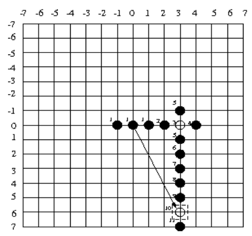
Modified 2-D Logarithmic Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case



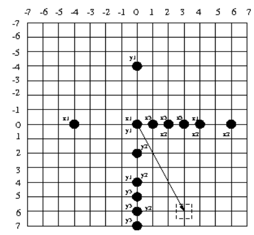
Three-Step Hierarchical Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case



Cross-Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case



One-at-a-time Search(OTS) Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case



Parallel Hierarchical One-Dimensional Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case

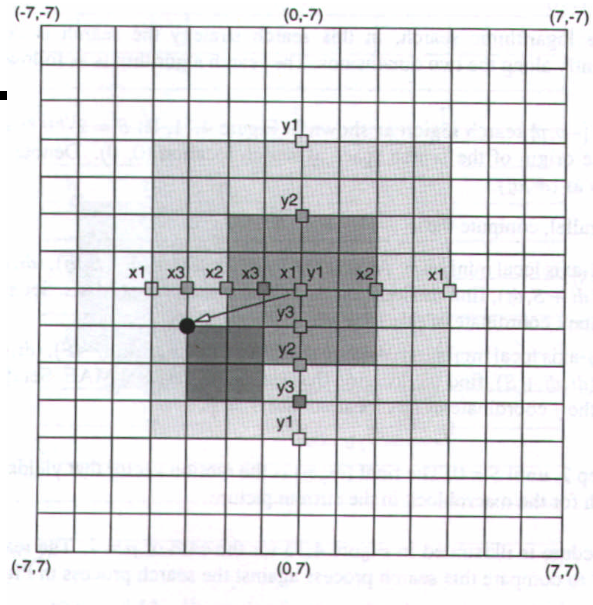


Figure 4.13 Example of PHODS strategy.

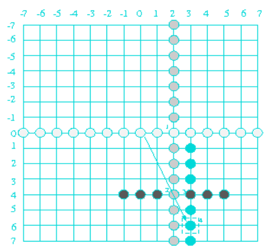
@NTUEE DSP/IC Lab

Fast Search Algorithms

One-Dimensional Full Search

For (i= -P to +P) Veci'= i | min D(i,j)
For (j= -P to +P) Vecj'= j | min D(Veci',j)
For (i= -P/2 to +P/2) Veci = i | min D(i+Veci',Vecj')
For (j= -P/2 to +P/2) Vecj = j | min D(Veci,j+Vecj')

Motion Vector = (Veci,Vecj)



One-Dimensional Full Search Procedure
Search Range -7 to +7
Motion Vector (3,6) in this case

- ***Hardware-Oriented***
- ***Full Data Reuse***
- ***Regular Data Flow***
- ***Fixed-Step Operation***
- ***Good Performance***

Ref: Mei-Juan Chen, Liang-Gee Chen and Tzi-Dar Chiueh, "One-Dimensional Full Search Motion Estimation Algorithm for Video Coding", IEEE Transactions on Circuits and Systems for Video Technology, Vol.4, No.5, October 1994.

@NTUEE DSP/IC Lab

Performance Comparison

	Miss America	Train Calendar	Table Tennis
PSNR (dB)			
2DFS	38.64	19.57	25.48
1DFS	38.39	19.45	24.55
3Step	38.20	19.40	24.06
Entropy (bits/pixel)			
2D FS	3.49	5.93	5.27
1DFS	3.52	5.94	5.39
3Step	3.56	5.94	5.47

30 frames/sec Sequences

Matching Criteria

- ***Cross-Correlation Function(CCF)***
- ***Mean Squared Error(MSE)***
- ***Mean Absolute Error(MAE)***
- ***Pel Difference Classification(PDC)***
- ***Minimized Maximum Error (MiniMax)***

$x_t(k,l)$: luminance for the location (k,l) in $X_t(p,q)$

$x_{t-1}(k+i,l+j)$: luminance for the shifted location by i pels and j lines within the search range

Matching Criteria

Cross-Correlation Function(CCF)

$$CCF(i,j)=\frac{\sum_{k=1}^{N}\sum_{l=1}^{N} x_t(k,l)\cdot x_{t-1}(k+i,l+j)}{\left[\sum_{k=1}^{N}\sum_{l=1}^{N} x_t^2(k,l)\right]^{\frac{1}{2}}\left[\sum_{k=1}^{N}\sum_{l=1}^{N} x_{t-1}^2(k+i,l+j)\right]^{\frac{1}{2}}}$$

(Veci,Vecj)=(i,j) | max *CCF(i,j)*

Mean Squared Error(MSE) --> most widely used in software

$$MSE(i,j)=\frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}[x_t(k,l)-x_{t-1}(k+i,l+j)]^2$$

(Veci,Vecj)=(i,j) | min *MSE(i,j)*

@NTUEE DSP/IC Lab

27

Matching Criteria

Mean Absolute Error(MAE) ➡ ***Most Widely Used in hardware***

$$MAE(i,j)=\frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left|x_t(k,l)-x_{t-1}(k+i,l+j)\right|$$

(Veci,Vecj)=(i,j) | min *MAE(i,j)*

Pel Difference Classification(PDC)

$$T(k,l,i,j)=1,\ \left|x_t(k,l)-x_{t-1}(k+i,l+j)\right|\le Threshold$$
$$0,\ otherwise$$

$$G(i,j)=\sum_{k=1}^{N}\sum_{l=1}^{N}T(k,l,i,j)$$

(Veci,Vecj)=(i,j) | max *G(i,j)*

@NTUEE DSP/IC Lab

28

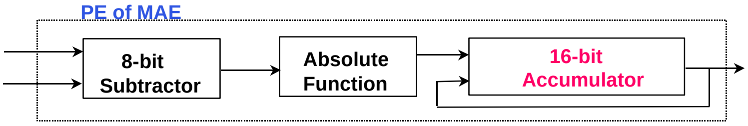
Matching Criteria

Minimized Maximum Error (MiniMax) → Hardware Reduction

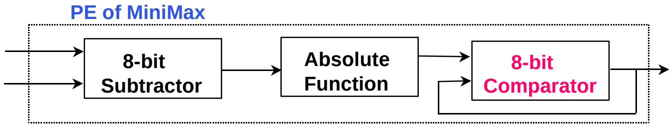
$$G(i,j) = \max \left| x_t(k,l) - x_{t-1}(k+i, l+j) \right|$$

(Veci,Vecj)= (i,j) | min *G(i,j)*

PE of MAE



PE of MiniMax



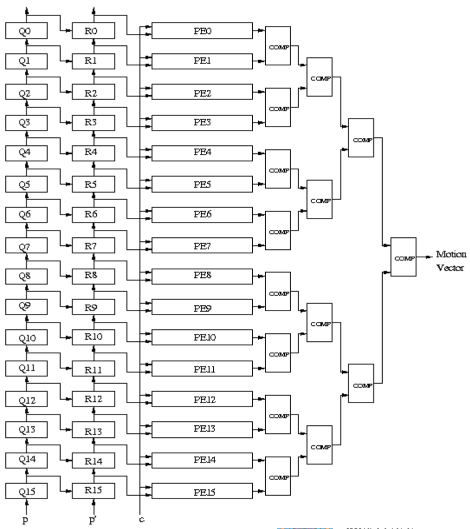
@NTUEE DSP/IC Lab

29

Architecture Design and Chip Implementation

- ***One-Dimensional Full Search***
- ***MiniMax Criterion***

Technology	0.8 um CMOS
Chip Size	6.5 mm x 5.3 mm
Number of Transistors	48170
Chip Clock	43 MHz
Number of Pads	61
Power Dissipation	383 mW



@NTUEE DSP/IC Lab

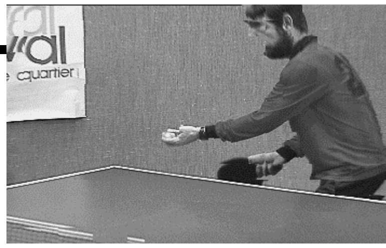
30

Subjective Quality

1DFS'

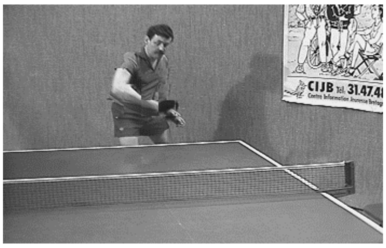
Table Tennis I

Min-max

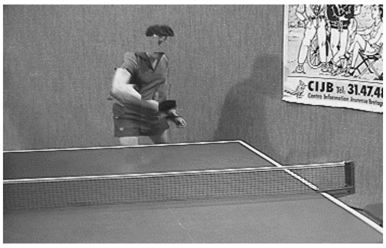


MAE

Table Tennis II



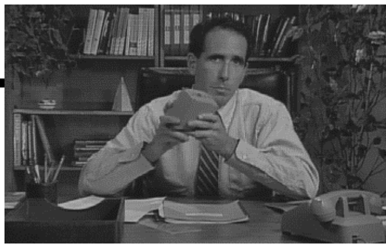
Min-max



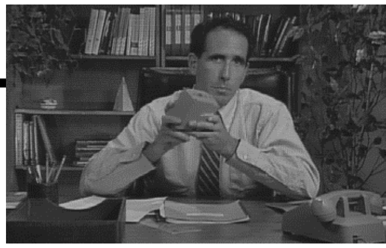
MAE

@NTUEE DSP/IC Lab

Salesman



Min-max



MAE

Train & Calendar

Min-max

MAE

@NTUEE DSP/IC Lab

Other Approaches: Pixel Subsampling for MAE calculations

1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4

Figure 4.15 Pixel decimation for block matching in an 8×8 block.

Pixel Subsampling for MAE calculations

- For each MAE value , we use only 1/4 of the pixels.
- But every pixel in the block will be used.
- It minimizes the possibility of not considering one-pixel-wide horizontal , vertical and diagonal lines .

Projections for MAE calculations

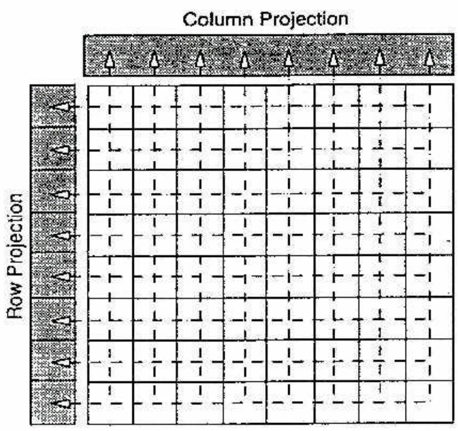


Figure 4.16 Row and column projection of pixels in an 8 × 8 block.

@NTUEE DSP/IC Lab

35

Predictive Motion Estimation Using Boundary/Side Match

- ***Motion***
  - ***Relieve huge computation complexity imposed on full search for large search range motion estimation***
  - ***MPEG and HDTV applications***
  - ***Little attention paid in the past***
- ***Goal***
  - ***Provide solutions for the computation reduction of increasing search area***
  - ***Object-based concept***

@NTUEE DSP/IC Lab

36

Predictive Motion Estimation Algorithms

$$V_t(p,q) = E[V_t(p,q)] + dV_t(p,q) \quad , \quad |dV_t(p,q)| \le w$$

$E[V_t(p,q)]$:Predicted Motion Vector

$dV_t(p,q)$:Refined Displacement

w :Reduced Search Range

(e.g. one-quarter of search range Ω)

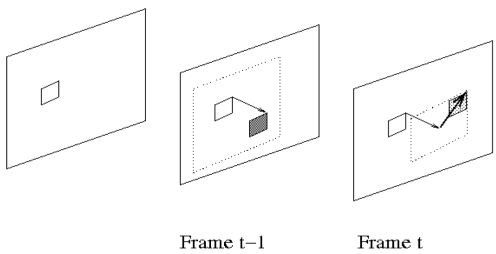
- **_Temporal Prediction_**
  - **_Inter-Frame Prediction_**
- **_Spatial Prediction_**
  - **_Inter-Block Prediction_**
  - **_Median Vector Prediction_**
  - **_Proposed Boundary Match Prediction_**
  - **_Proposed Side Match Prediction_**

37

@NTUEE
DSP/IC Lab

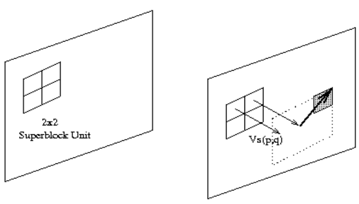
Inter-Frame Prediction

$$E[V_t(p,q)] = V_{t-1}(p,q)$$



Inter-Block Prediction

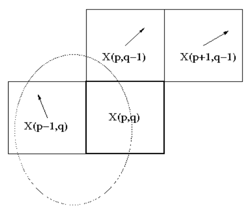
$$E[V_t(p,q)] = V_s(p,q)$$



Median Vector Prediction

$$E[V_t(p,q)] = \arg \underset{v \in V_{cs}}{Median}\, V$$

$$V_{cs} = \{V_t(p,q-1), V_t(p-1,q), V_t(p+1,q-1)\}$$



38

@NTUEE
DSP/IC Lab

Boundary/Side Match

- ***A motion object often covers many small blocks***
- ***Blocks in the same object have similar motion vectors***
- ***Spatial Neighboring Blocks***
  - ***Very likely move in the same direction with similar velocities***
  - ***Highly correlated or dependent***
- ***Utilize the high spatial correlation between the boundary pixels of adjacent blocks to determine a more accurate initial motion estimation center***
- ***Other applications***
  - ***Vector quantization (VQ)***
  - ***Recovery of lost vector or channel error***
  - ***Transform coded image reconstruction***

@NTUEE
DSP/IC Lab

Boundary Match

$$[X(p,q)]_l^c = (x_{l1}, x_{l2}, \cdots, x_{l\,n})^T$$ column vector

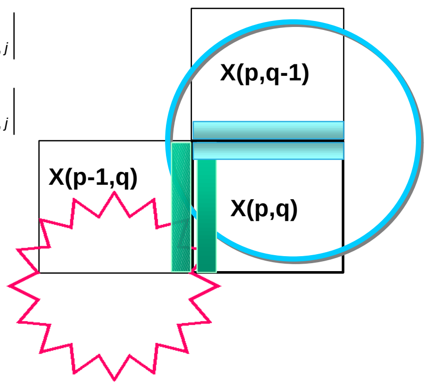
$$[X(p,q)]_l^r = (x_{1l}, x_{2l}, \cdots, x_{n\,l})$$ row vector

$$d_{bm,V_t(p,q-1)} = \sum_{j=1}^{n} \left| [X(p,q)]_{1,j}^r - [X(p,q-1)]_{n,j}^r \right|$$

$$d_{bm,V_t(p-1,q)} = \sum_{j=1}^{n} \left| [X(p,q)]_{1,j}^c - [X(p-1,q)]_{n,j}^c \right|$$

$$E[V_t(p,q)] = \arg \underset{v \in V_{cs}}{Min}\, d_{bm,v}$$

$$V_{cs} = \{V_t(p,q-1), V_t(p-1,q)\}$$



@NTUEE
DSP/IC Lab

Side Match

$$d_{smU} = \sum_{j=1}^{n} \left| [X_{MC}(p,q)]^{r}_{1,j} - [X(p,q-1)]^{r}_{n,j} \right|$$

$$d_{smL} = \sum_{j=1}^{n} \left| [X_{MC}(p,q)]^{c}_{1,j} - [X(p-1,q)]^{c}_{n,j} \right|$$

$$d_{smR} = \sum_{j=1}^{n} \left| [X_{MC}(p,q)]^{c}_{n,j} - [X(p+1,q)]^{c}_{1,j} \right|$$

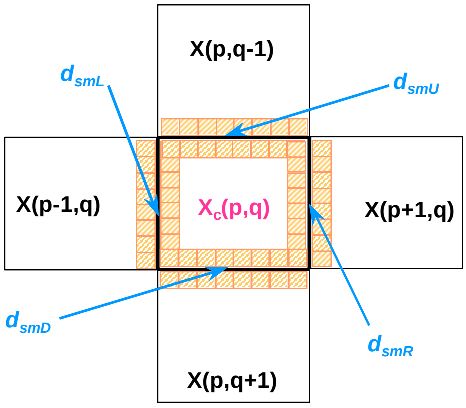
$$d_{smD} = \sum_{j=1}^{n} \left| [X_{MC}(p,q)]^{r}_{n,j} - [X(p,q+1)]^{r}_{1,j} \right|$$

$$d_{sm} = d_{smU} + d_{smL} + d_{smR} + d_{smD}$$

$$E[V_t(p,q)] = \arg \underset{v \in V_{cs}}{Min}\ d_{sm,v}$$

$$V_{cs} = \{V_t(p,q-1), V_t(p-1,q), V_t(p+1,q-1)\}$$

minimum gray-level transition
across four boundaries



Computation Comparison

Algorithms	Search Range	Computation Complexity (SD/block)
Full Search	-p ~ p	$(2p+1)^2$ x m x n
Inter- Frame	-p/2 ~ p/2	$(p+1)^2$ x m x n
Inter- Block	-p ~ p - p/2 ~ p/2	[$1/4(2p+1)^2$ + $3/4(p+1)^2$] x m x n
Median	- p/2 ~ p/2	$(p+1)^2$ x m x n +6
Boundary Match	- p/2 ~ p/2	$(p+1)^2$ x m x n + m + n
Side Match	- p/2 ~ p/2	$(p+1)^2$ x m x n+ 6(m + n)

Objective Performance Measurement

PSNR: Peak-to-Peak Signal-to-Noise Ratio

predicted image compared to the original image

$$PSNR = 10 \times \log_{10} \frac{255^2}{MSE} \quad \textbf{(dB)}$$

Entropy of Prediction Error (bits/pixel)

$$Entropy = -\sum_i p(e_i) \cdot \log_2 p(e_i) = \sum_i p(e_i) \cdot \log_2 \frac{1}{p(e_i)}$$

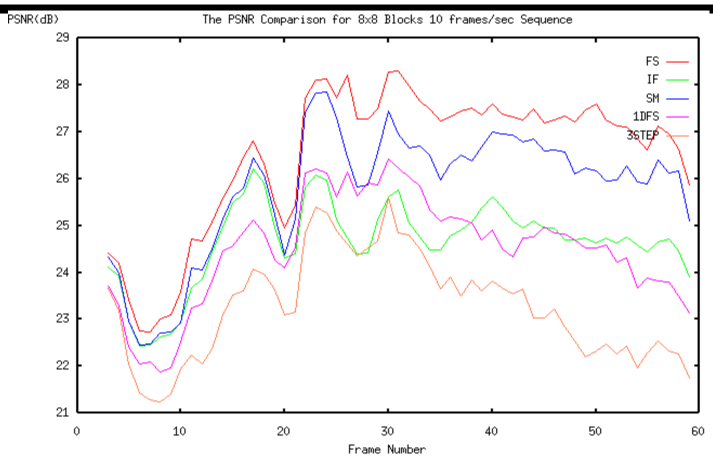
$p(e_i)$ **: probability of occurrence of prediction error pattern**

PSNR Comparison

Algorithms	30 frames/sec	15 frames/sec	10 frames/sec
Full Search	27.33	26.90	26.49
Inter-Frame	25.73	25.23	24.66
Inter-Block	26.09	25.52	25.22
Median	25.57	24.93	24.23
Boundary Match	26.40	26.03	25.68
Side Match	26.52	26.17	25.77
IDFS	26.23	25.09	24.49
3-Step	24.94	23.06	23.26

Block size 8x8 60 frames Table Tennis Sequence

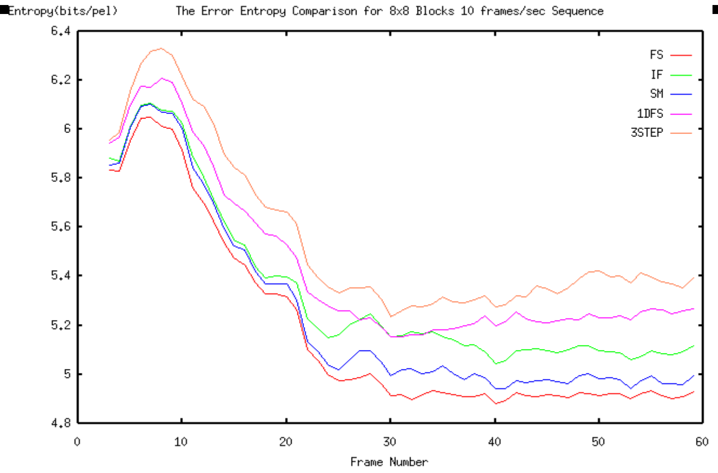
vs Frame Number



@NTUEE
DSP/IC Lab

45

Error Entropy vs Frame Number



@NTUEE
DSP/IC Lab

46

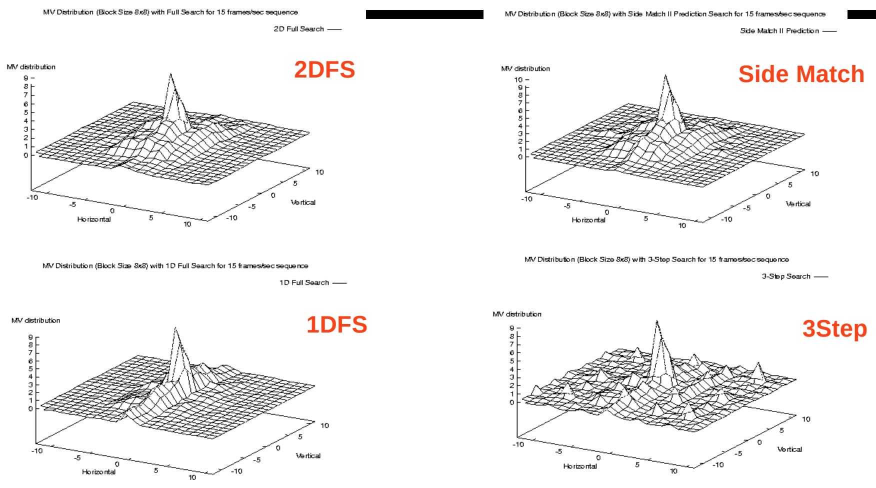
Motion Vector Characteristics

MV Euclidean Distance(ED) $(V_x - V\quad) + (V\ - V\quad)$

Hit Ratio(HR) **Criterion:** $|V\ - V\quad| <= 2$ **and** $|V\ - V\quad| <= 2$

Frame Rate	30 frames/sec		15 frames/sec		10 frames/sec	
Algorithms	ED	HR	ED	HR	ED	HR
Inter-Frame	23.52	72.92%	60.05	67.50%	74.67	62.46%
Inter-Block	13.05	76.88%	54.54	71.37%	72.82	67.55%
Median	20.63	73.33%	57.15	66.98%	82.41	60.64%
Boundary Match	10.08	79.70%	47.88	74.76%	66.91	70.73%
Side Match	9.97	80.93%	46.08	75.72%	65.18	71.86%
IDFS	26.71	61.97%	95.13	41.74%	132.27	31.55%
3-Step	23.14	50.81%	100.92	30.39%	144.47	23.80%

Motion Vector Distribution

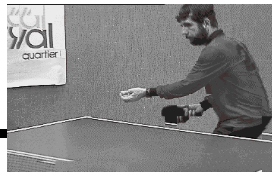


block size 8x8 15 frames/sec Table Tennis Sequence

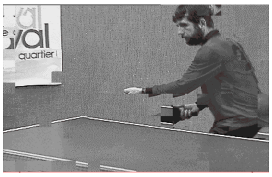
Subjective Quality

16x16 blocks
15 frames/sec Table Tennis Sequence

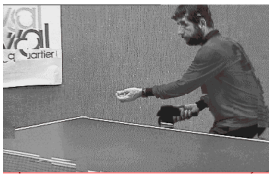
Original



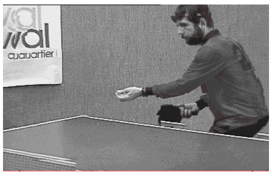
2DFS



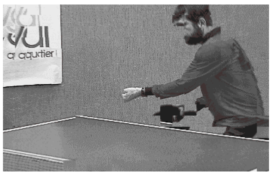
Inter-Frame



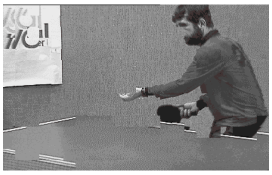
Inter-Block



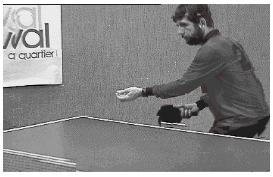
Boundary Match



1DFS

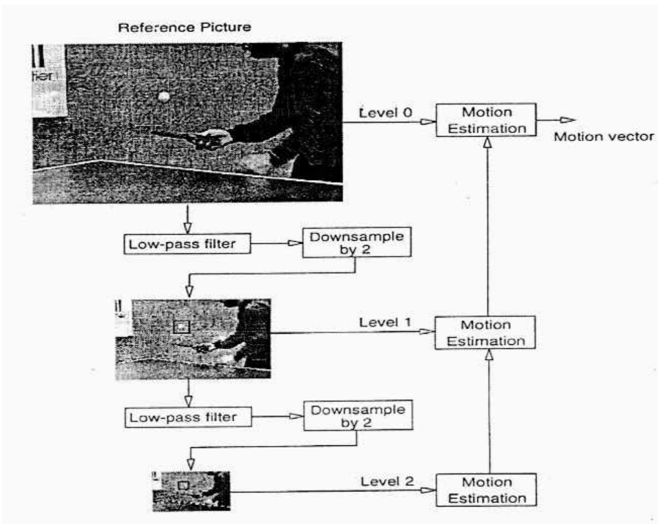


3Step



Side Match

Hierarchical Motion Estimation



Hierarchical Motion Estimation

1. **Level 2**

- Picture size = 180 × 120, and macroblock size = 4 × 4.
- Number of macroblocks = $\frac{Picture\ size}{Macroblock\ size}$ = 1,350. At 30 fps, we have 40,500 macroblocks.
- Search parameter = $\lceil \frac{p}{4} \rceil = 4$.
- Number of search locations = $(2 \times 4 + 1)^2 = 81$.
- Number of operations per search location = macroblock size × 3 = 48.

Complexity for Level 2 = 40,500 × 81 × 48 = 157.46 MOPS.

@NTUEE DSP/IC Lab

51

Hierarchical Motion Estimation

2. **Level 1**

- Picture size = 360 × 240, and macroblock size = 8 × 8.
- Number of macroblocks = 1,350. At 30 fps, we have 40,500 macroblocks.
- Search parameter = 1.
- Number of search locations = 9.
- Number of operations per search location = macroblock size × 3 = 192.

Complexity for Level 1 = 40,500 × 9 × 192 = 69.98 MOPS.

@NTUEE DSP/IC Lab

52

3. **Level 0**

- Picture size = 720 × 480, and macroblock size = 16 × 16.
- Number of macroblocks = 1,350. At 30 fps, we have 40,500 macroblocks.
- Search parameter = 1.
- Number of search locations = 9.
- Number of operations per search location = macroblock size × 3 = 768.

Complexity for Level 0 = 40,500 × 72 × 48 = 279.9 MOPS.

Hierarchical Motion Estimation

- It requires increased storage to keep pictures at different resolutions.
- Motion vector may be inaccurate for regions containing small objects .
- Low-pass filter can reduce noise.

Operations Comparison of Motion-Estimation Methods

Search Method	Operations per Macroblock	Operations for pictures 720 x 40 at 30 fps	
		$p = 15$	$p = 7$
Full-search	$(2p+1)^2 NM3$	29.89 GOPS	6.99 GOPS
Logarithmic	$(8\lceil log_2 p\rceil + 1)NM3$	1.02 GOPS	777.60 MOPS
PHODS	$(4\lceil log_2 p\rceil + 1)NM3$	528.76 MOPS	404.35 MOPS
Hierarchical	$[(2\lceil \frac{p}{4}\rceil + 1)^2 + 180]\ \frac{NM}{16}3$	507.38 MOPS	398.52 MOPS

Table 4.1 Computational complexity and MOPS requirements for various motion-estimation algorithms using the MAE criterion and a [-p, p] search range.

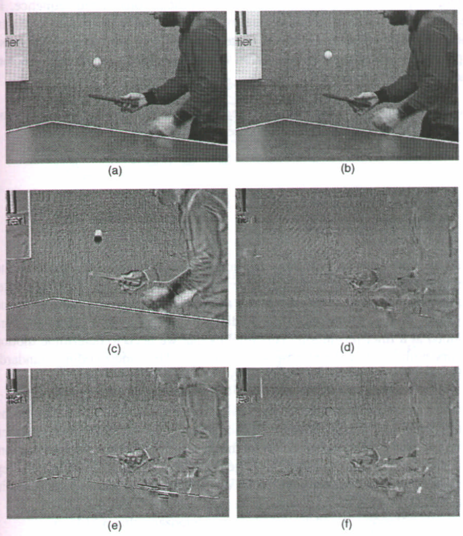


Figure 4.18 Performance comparison of motion-estimation methods. (a) Reference picture; (b) current picture; (c) simple frame differencing; (d) full-search; (e) logarithmic search; (f) three-level hierarchical search.

Sub-pixel-accurate motion estimation

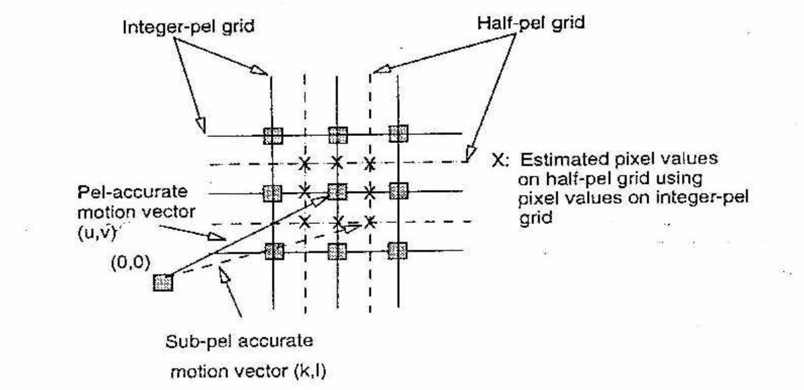


Figure 4.19 Half-pel accurate motion vector estimation.

Bidirectional Temporal Prediction for Progressive Video

- Reference frame must be I or P-frame
- B-picture
- 2 MV