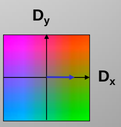
1

Motion Estimation

2 **Motion Estimation**

- Is there any motion?
- How fast?
- Into which direction?

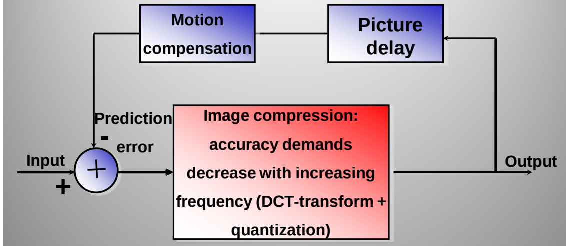


ME

3 **Application dependency of ME**

- Scan rate conversion (true-motion vectors)
  - De-interlacing
  - Picture rate conversion
- Video compression (low average error)
  - MPEG
  - H.264
- True-motion vectors are usually more consistent than coding vectors. Consistency has some, but no dominant relevance for coding efficiency

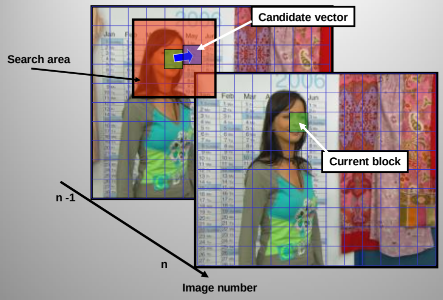
4 **Motion estimation and coding**



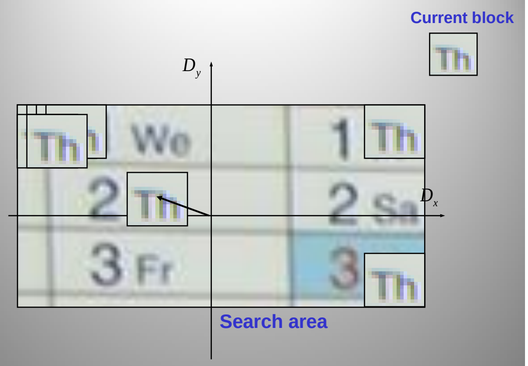
5

Block-matching ME methods: Full-search

6 **Block-matching; find corresponding block in image n-1**



7 **Finding block similarity**



8 **Formal definitions**

Luminance value in previous picture, shifted over candidate vector C: $F(\vec{x}-\vec{C},n-1)$

A block matcher minimises an error function, e.g. a SAD, varying C:

$$\varepsilon(\vec{C},\vec{X},n)=\sum_{\vec{x}\in B(\vec{X})}\left|F(\vec{x},n)-F(\vec{x}-\vec{C},n-1)\right|$$

And the resulting candidate vector for which the error is minimal is assumed to be the displacement vector:

$$\vec{D}(\vec{x},n)$$

9 **Alternative match criteria**

Complexity

- Correlation (NCCF) of pixels in the two blocks
- Mean Square Error (MSE) between pixels in the blocks
- Mean Absolute Difference (MAD) between pixels in the blocks
- Number of significantly different pixels (NSD) in the two blocks

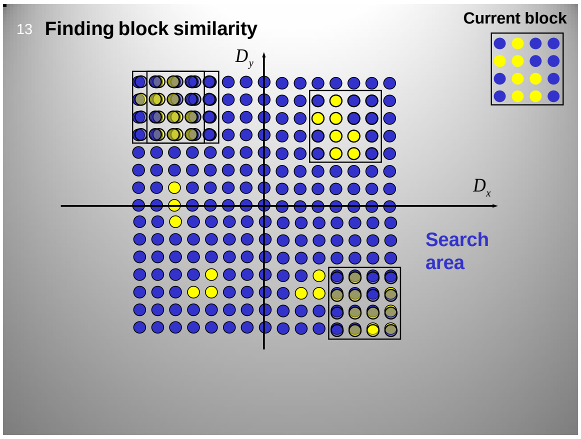
10 **Comparison of match criteria**

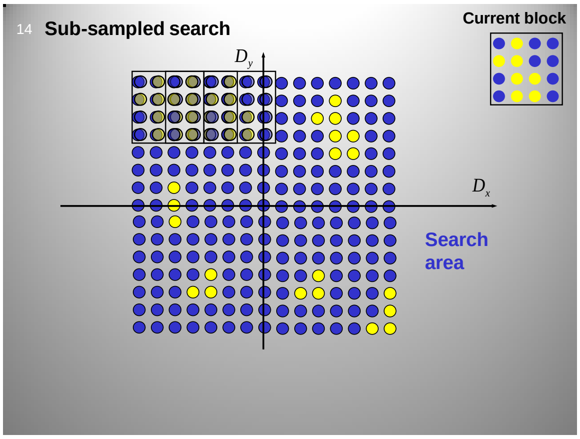
11 **Operations count of full search block matching**

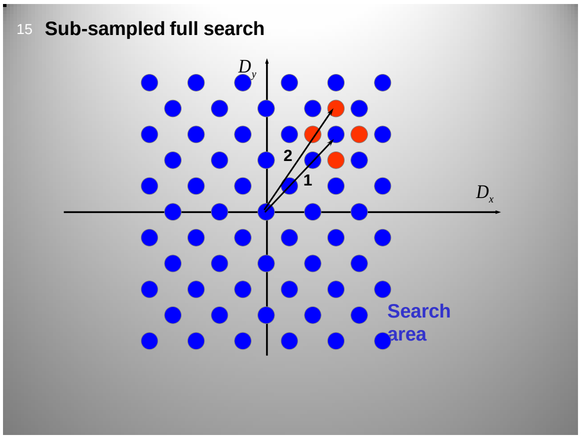
- CCIR signal
  - 720x288x50 (pixels/s)
- Search window for realistic velocities
  - 64x48 (HxV in pixels) = 3000 possible vectors, assuming integer vector accuracy
- Matching error (SAD) calculation only:
  - approximately: $1\text{x}10^{11}$ (ops/s)

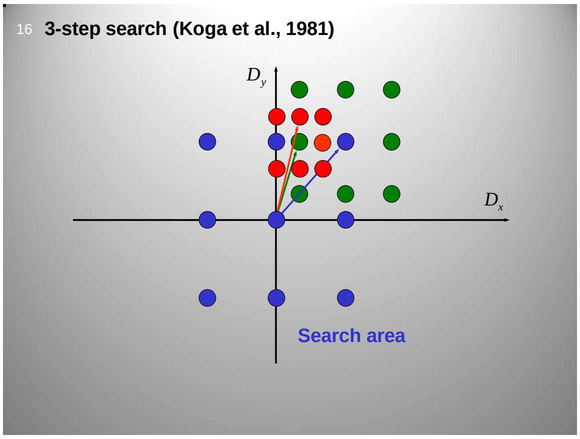
12

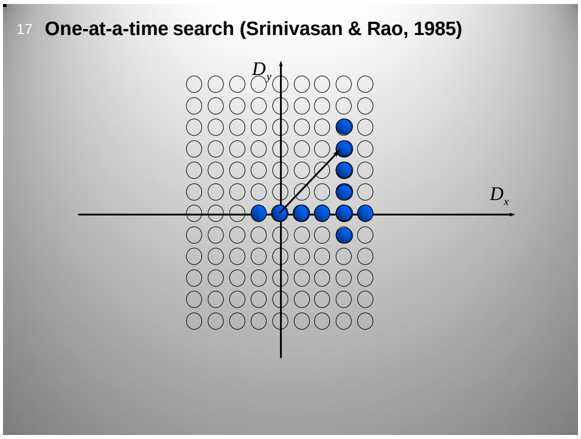
Block-matching efficient search techniques

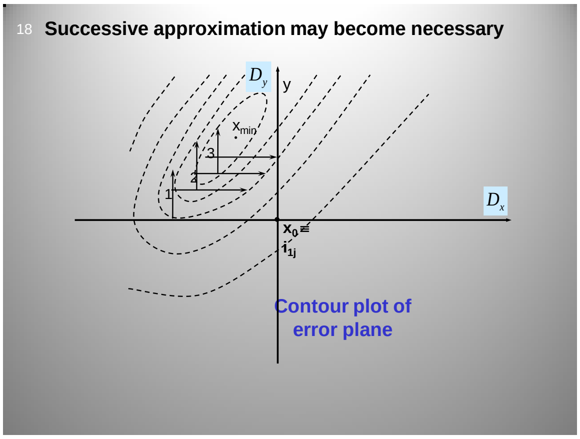

13
Finding block similarity
Current block
D_y
D_x
Search area


14
Sub-sampled search
Current block
D_y
D_x
Search area

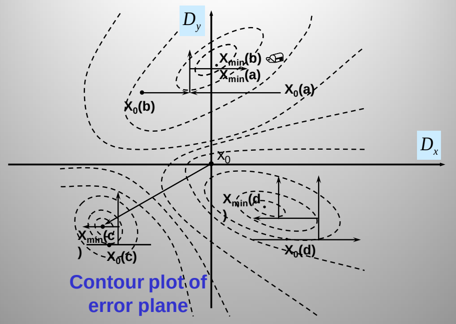

15
Sub-sampled full search
D_y
D_x
2
1
Search area


16
3-step search (Koga et al., 1981)
D_y
D_x
Search area

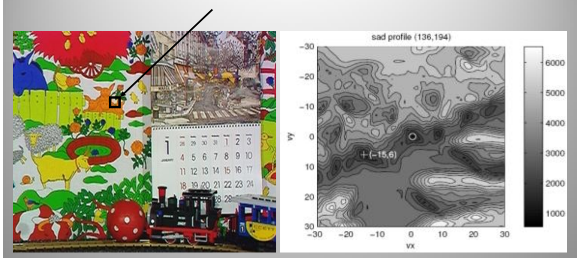

17
One-at-a-time search (Srinivasan & Rao, 1985)
D_y
D_x


18
Successive approximation may become necessary
D_y
y
x_{min}
3
2
1
D_x
$x_0 \equiv$
i_{1j}
Contour plot of error plane

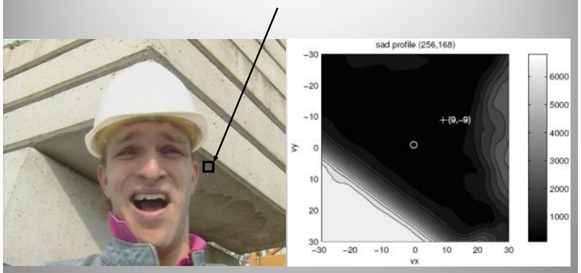
19 **Prevention of trap in local minimum**



20 **Reality is even more complicated…**



21 **And sometimes there is no unique solution…**



22 **Comparison of search techniques**

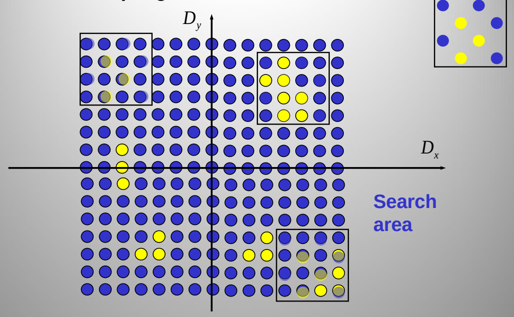
23

Pixel sub-sampling in match function

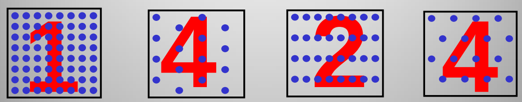
24 **Intermediate conclusion**

- Efficient search techniques can highly reduce the operations count of a block matching motion estimator, but increase the risk of getting trapped in a local minimum of the error function
- Methods to prevent the disadvantages of efficient search, increase complexity again.

25 **Pixel sub-sampling of match error criterion**



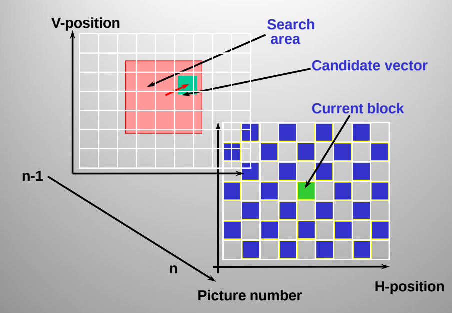
26 **Pixel sub-sampling in match error criterion**



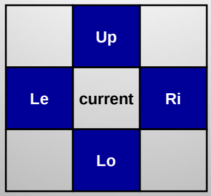
27

Block sub-sampling

28 **Block sub-sampling**



29 **Interpolate missing motion vectors**

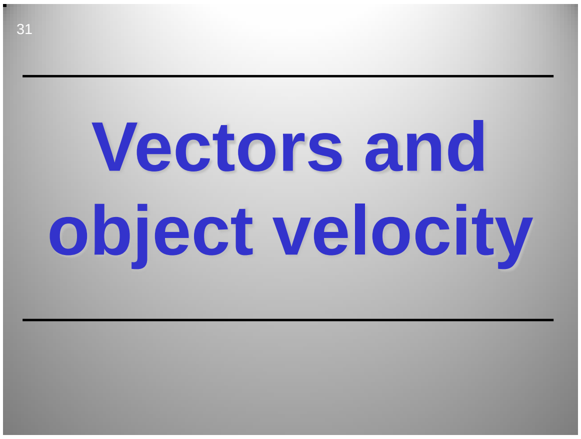


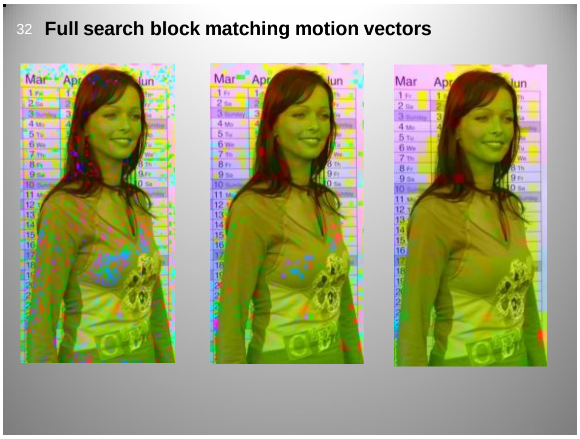
1: Current D_x = median{Le_x, (Up_x+Lo_x)/2, Ri_x}
Current D_y = median{Le_y, (Up_y+Lo_y)/2, Ri_y}

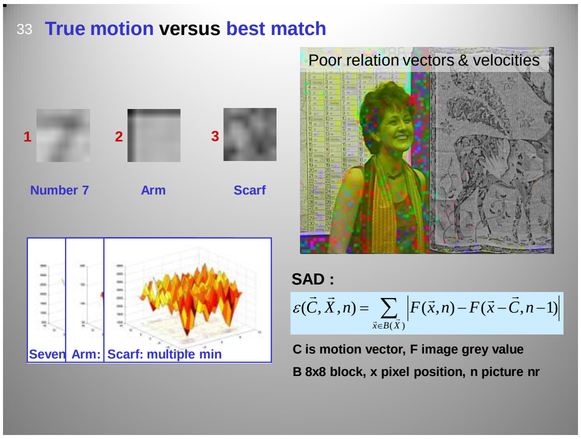
2: Use the vector-median to prevent new vectors

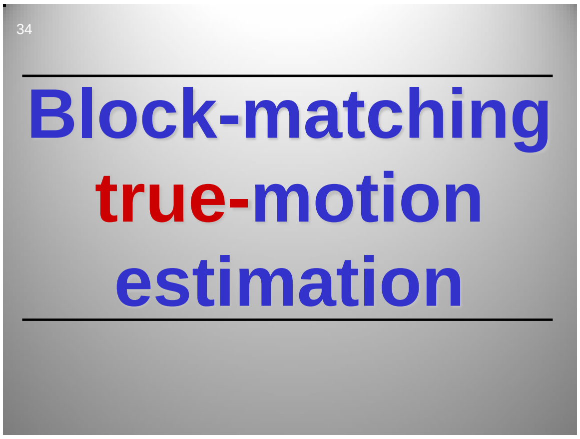
30 **Summary cost reduction block matchers**

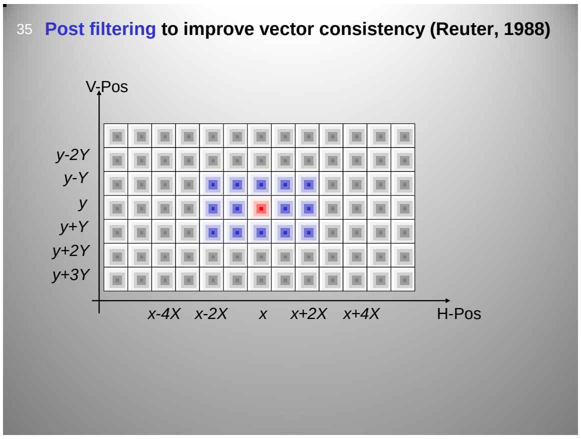
- Simple match criterion
- Efficient search strategy
- Pixel sub-sampling in match criterion
  - a factor of four is usually feasible with little influence on the performance
- Block sub-sampling
  - only valid if motion field is smooth

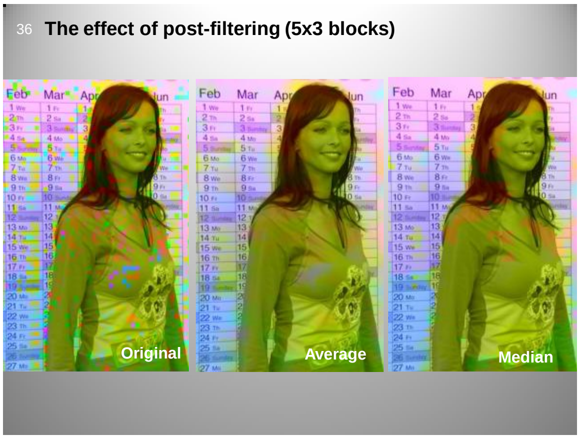

31
Vectors and object velocity

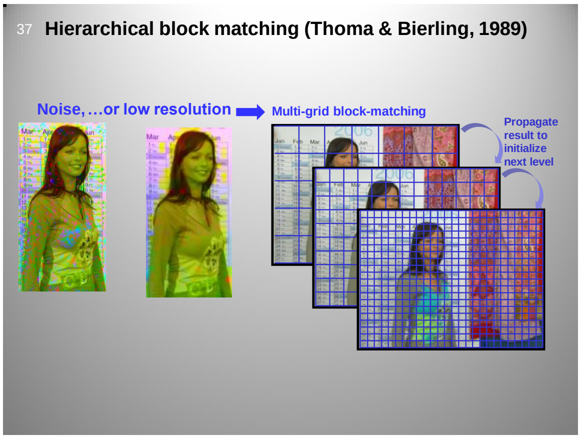

32 Full search block matching motion vectors

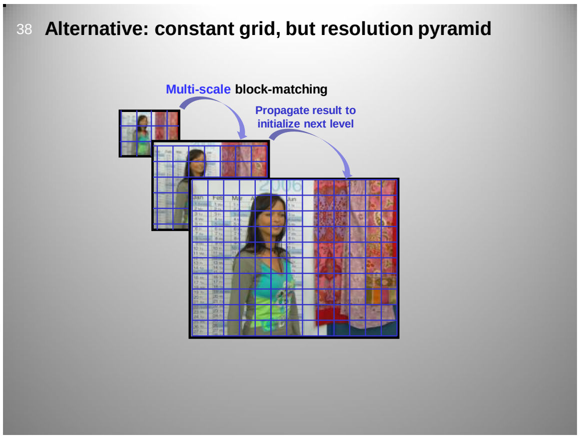

33 True motion versus best match
Poor relation vectors & velocities
1
2
3
Number 7
Arm
Scarf
Seven
Arm:
Scarf: multiple min
SAD :
$\varepsilon(\vec{C},\vec{X},n)=\sum_{\vec{x}\in B(\vec{X})}\left|F(\vec{x},n)-F(\vec{x}-\vec{C},n-1)\right|$
C is motion vector, F image grey value
B 8x8 block, x pixel position, n picture nr

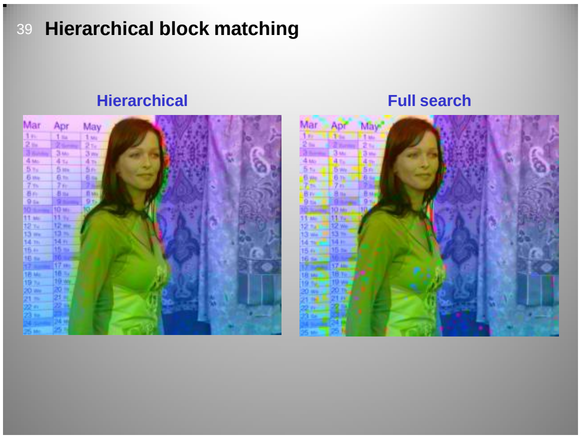

34
Block-matching true-motion estimation

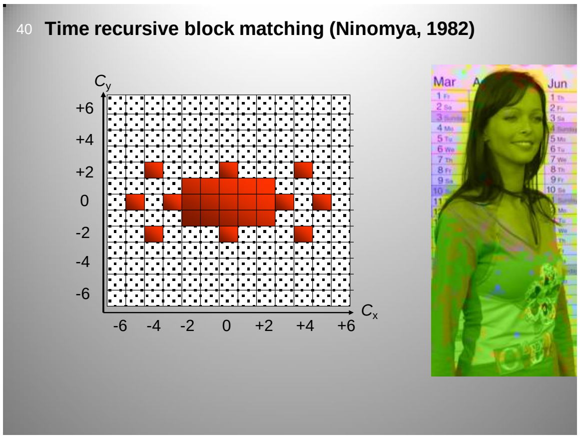

35 Post filtering to improve vector consistency (Reuter, 1988)
V-Pos
y-2Y
y-Y
y
y+Y
y+2Y
y+3Y
x-4X
x-2X
x
x+2X
x+4X
H-Pos

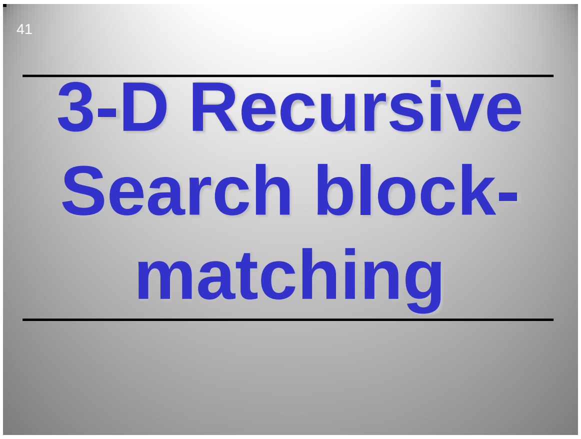

36 The effect of post-filtering (5x3 blocks)
Original
Average
Median

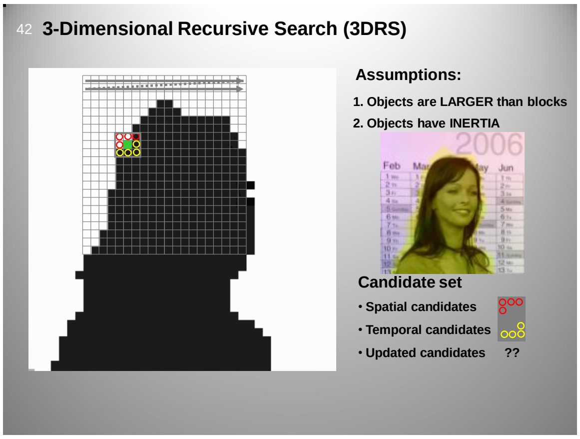

37 Hierarchical block matching (Thoma & Bierling, 1989)
Noise, …or low resolution
Multi-grid block-matching
Propagate result to initialize next level

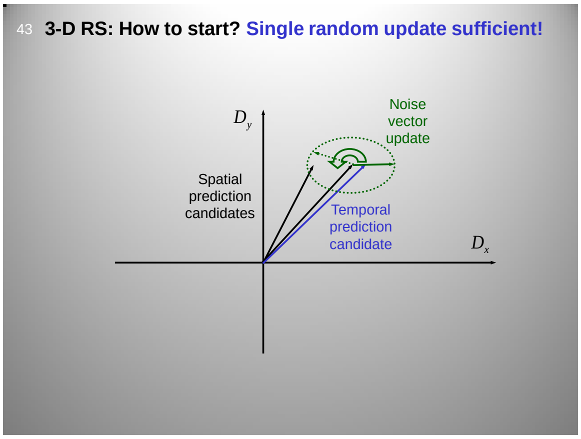

38 Alternative: constant grid, but resolution pyramid
Multi-scale block-matching
Propagate result to initialize next level

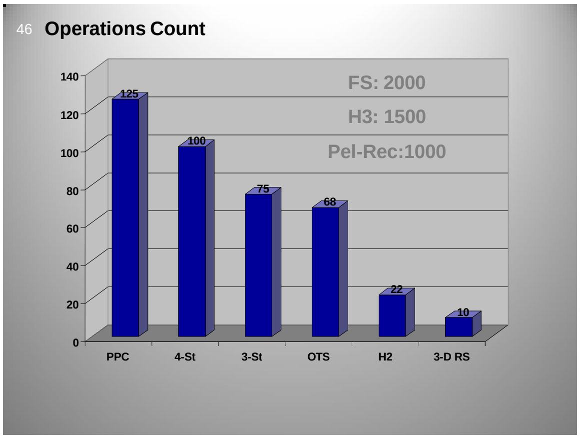

39 Hierarchical block matching
Hierarchical
Full search

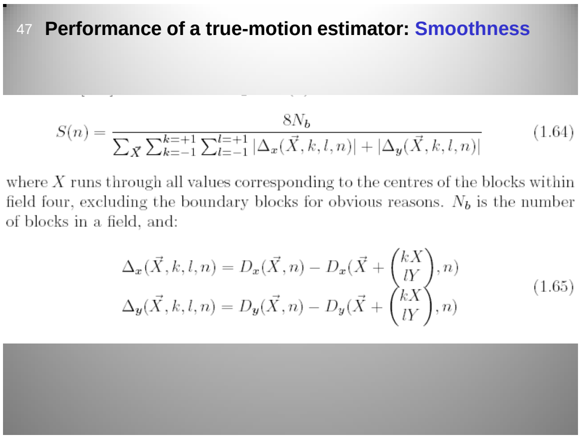

40 Time recursive block matching (Ninomya, 1982)
C_y
+6
+4
+2
0
-2
-4
-6
C_x
-6 -4 -2 0 +2 +4 +6


41
3-D Recursive Search block-matching


42 3-Dimensional Recursive Search (3DRS)
Assumptions:
1. Objects are LARGER than blocks
2. Objects have INERTIA
Candidate set
• Spatial candidates
• Temporal candidates
• Updated candidates ??



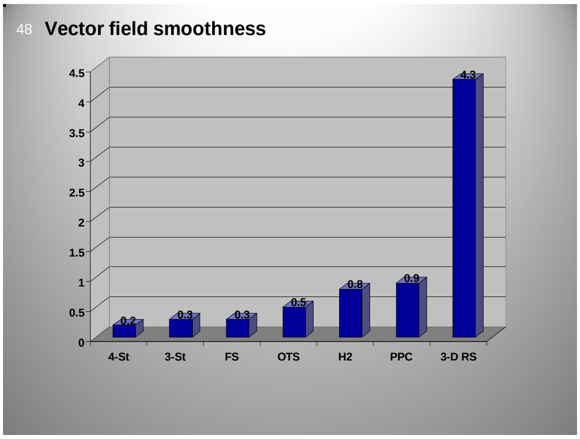




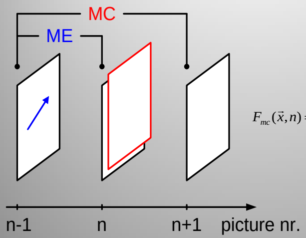
$$S(n) = \frac{8N_b}{\sum_{\vec{X}} \sum_{k=-1}^{k=+1} \sum_{l=-1}^{l=+1} |\Delta_x(\vec{X},k,l,n)| + |\Delta_y(\vec{X},k,l,n)|} \tag{1.64}$$

where X runs through all values corresponding to the centres of the blocks within field four, excluding the boundary blocks for obvious reasons. N_b is the number of blocks in a field, and:

$$\begin{aligned} \Delta_x(\vec{X},k,l,n) &= D_x(\vec{X},n) - D_x(\vec{X} + \begin{pmatrix} kX \\ lY \end{pmatrix}, n) \\ \Delta_y(\vec{X},k,l,n) &= D_y(\vec{X},n) - D_y(\vec{X} + \begin{pmatrix} kX \\ lY \end{pmatrix}, n) \end{aligned} \tag{1.65}$$



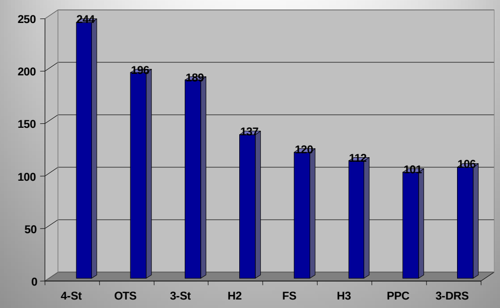
49 **Performance testing of true-motion estimator: M2SE**



$$MMSE(n)=\sum_{\forall\vec{x}}\left(F(\vec{x},n)-F_{mc}(\vec{x},n)\right)^2$$

$$F_{mc}(\vec{x},n)=\frac{1}{2}\left[F(\vec{x}-\vec{D}(\vec{x}),n-1)-F(\vec{x}+\vec{D}(\vec{x}),n+1)\right]$$

50 **M2SE score of ME-methods**



51 **Comparison of vector fields**

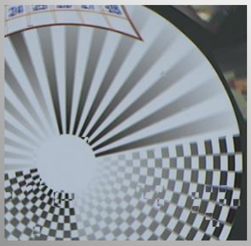
Phase Plane Correlation motion vectors

3-D Recursive Search BM motion vectors

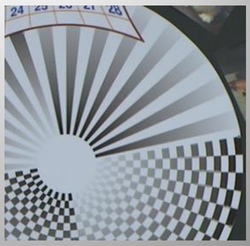
Phase Plane Correlation is a two-step hierarchical motion estimator. The first step is on large blocks (typically 64x64) and operates in the frequency domain, while the second step is on small block (e.g.2x2 or 8x8 (here) in the spatial domain). Algorithm originally proposed by Thomas [6], and applied in professional studio scan converters.

52 **MC up-conversion; Relevance of true-motion vectors**

Interpolated images using full search motion vectors



Interpolated image using 3D-RS motion vectors



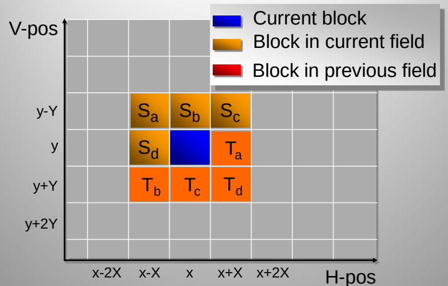
In contrast with coding, for scan rate conversion true-motion is an absolute must. RATHER SMOOTH THAN ACCURATE!!

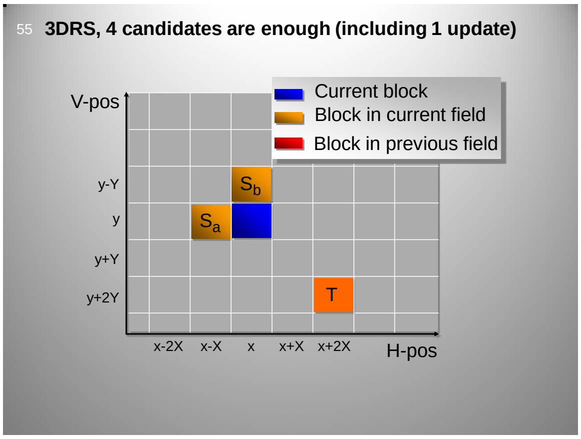
53

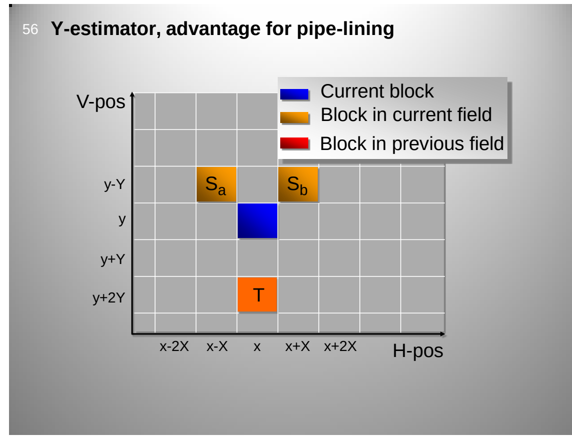
Simplifications

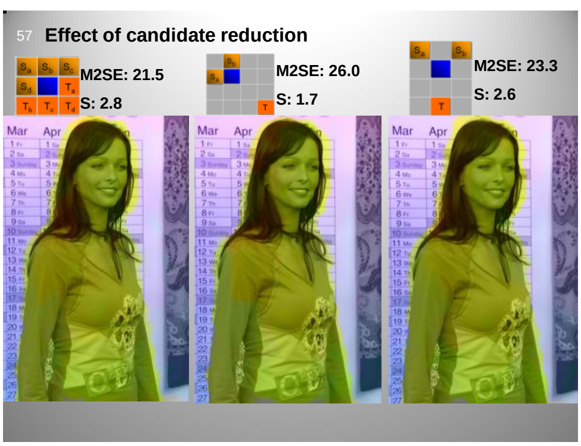
1) Reduced candidate set

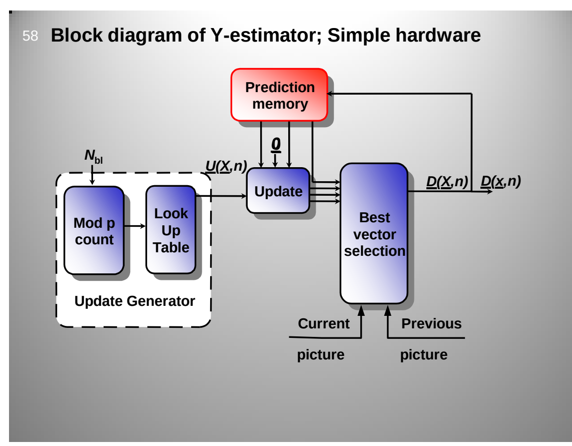
54 **With 8 prediction and 1 update: 9 candidates**



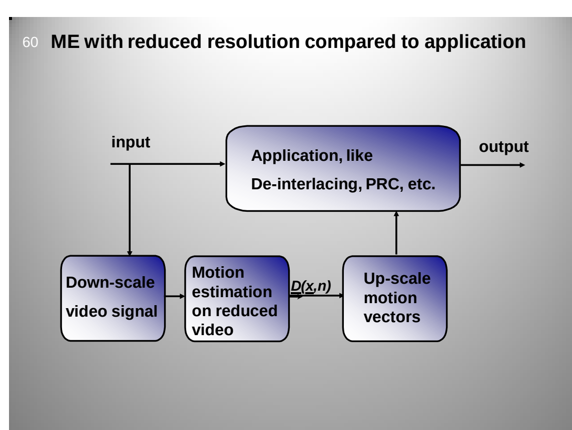

55 3DRS, 4 candidates are enough (including 1 update)
Current block
Block in current field
Block in previous field
V-pos
y-Y
y
y+Y
y+2Y
Sb
Sa
T
x-2X
x-X
x
x+X
x+2X
H-pos


56 Y-estimator, advantage for pipe-lining
Current block
Block in current field
Block in previous field
V-pos
y-Y
y
y+Y
y+2Y
Sa
Sb
T
x-2X
x-X
x
x+X
x+2X
H-pos


57 Effect of candidate reduction
M2SE: 21.5
S: 2.8
M2SE: 26.0
S: 1.7
M2SE: 23.3
S: 2.6


58 Block diagram of Y-estimator; Simple hardware
Prediction memory
0
U(X,n)
Update
Best vector selection
D(X,n)
D(x,n)
Mod p count
Look Up Table
Update Generator
Current picture
Previous picture

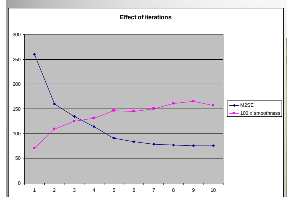

59
Simplifications
1) Reduced resolution for ME


60 ME with reduced resolution compared to application
input
Application, like
De-interlacing, PRC, etc.
output
Down-scale video signal
Motion estimation on reduced video
D(x,n)
Up-scale motion vectors

61

Sophis-tications

62 **Iterating more than once on an image pair**



Once, 1st image 10 times

Remark 1: if estimating in the output domain (100Hz): 2 iterations on video and 4 iterations on film material!

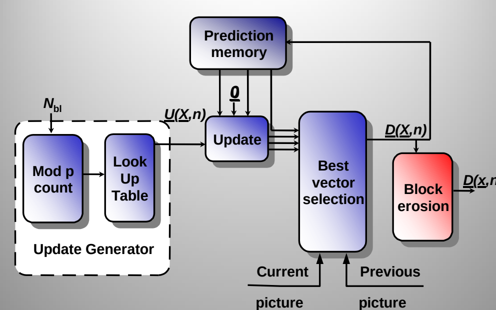
Remark 2: effect mainly shows in 1st image after scene change:

- 1iteration, 10th frame: M2SE: 29, Smoothness: 2.8
- 10 iterations, 10th frame: M2SE: 28, Smoothness: 3.5

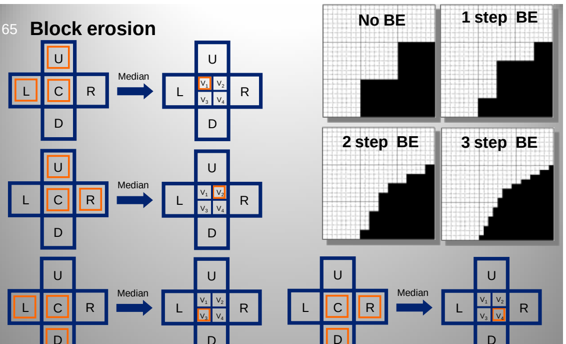
63

Block-erosion

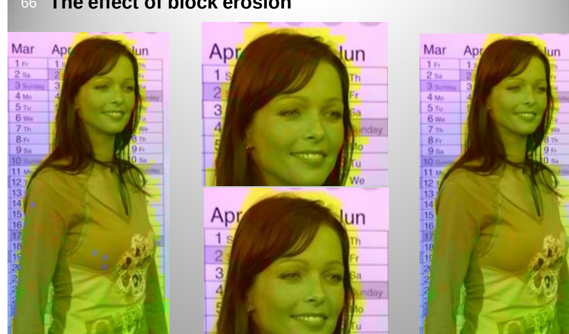
64 **Block diagram of Y-estimator; Simple hardware**



65 **Block erosion**



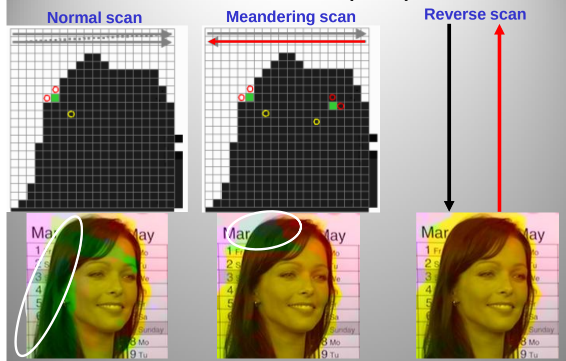
66 **The effect of block erosion**



67

Advanced scanning

68 **3-Dimensional Recursive Search (3DRS)**

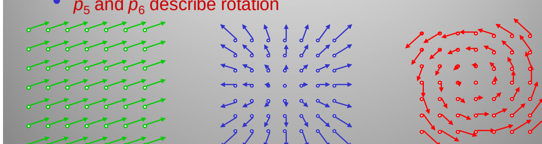


69

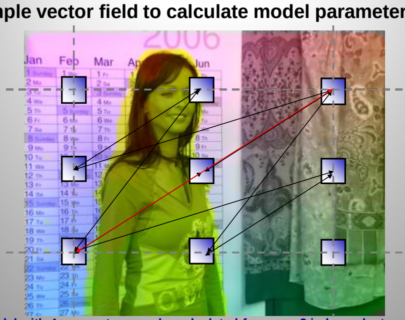
Parametric motion models

70 **Global motion estimation**

- Simple parametric motion model:
  - $D_x(\vec{x},n) = p_1(n) + p_3(n)x + p_5(n)y + ...$
  - $D_y(\vec{x},n) = p_2(n) + p_4(n)y + p_6(n)x + ...$
- p_1 and p_2 describe pan and tilt
- p_3 and p_4 describe zoom
- p_5 and p_6 describe rotation



71 **Sample vector field to calculate model parameters**



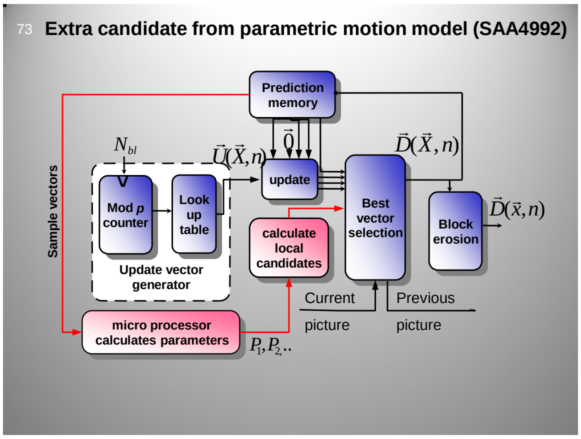
Motion model with 4 parameters can be calculated from any 2 independent sample vectors

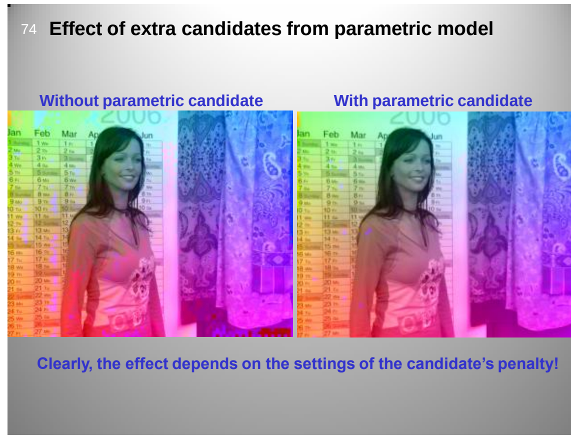
So, in total from these 9 vectors 18 models can be estimated

72 **Derive robust background model from sample vectors**

Take median of all estimated parameters to eliminate outliers:

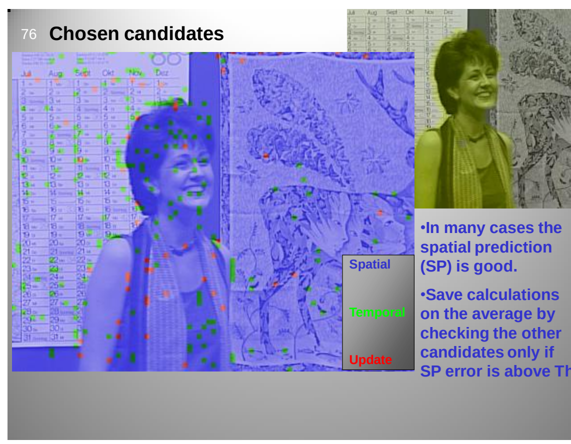
p_1 = median{$p^1{}_1, p^2{}_1, p^3{}_1, ... p^{18}{}_1$}
p_2 = median{$p^1{}_2, p^2{}_2, p^3{}_2, ... p^{18}{}_2$}
p_3 = median{$p^1{}_3, p^2{}_3, p^3{}_3, ... p^{18}{}_3$}
p_4 = median{$p^1{}_4, p^2{}_4, p^3{}_4, ... p^{18}{}_4$}

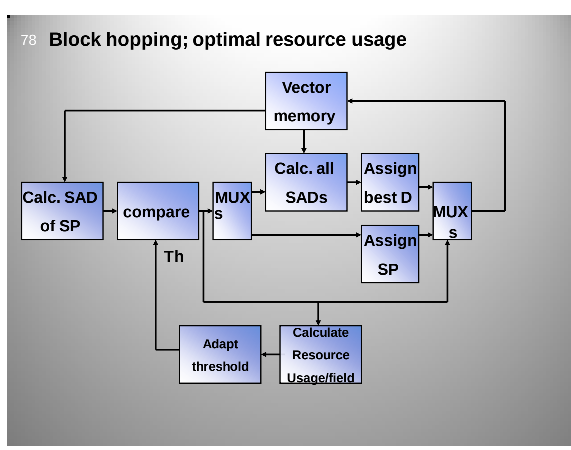


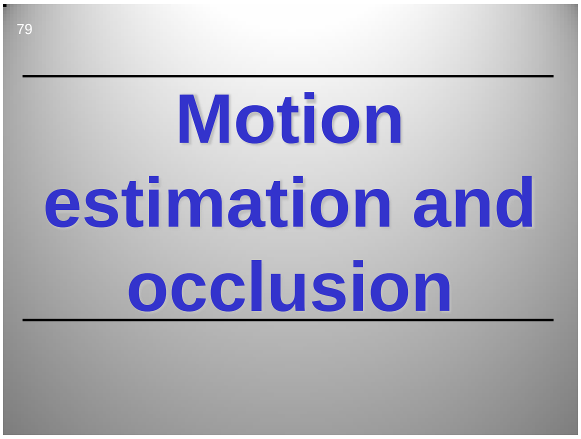


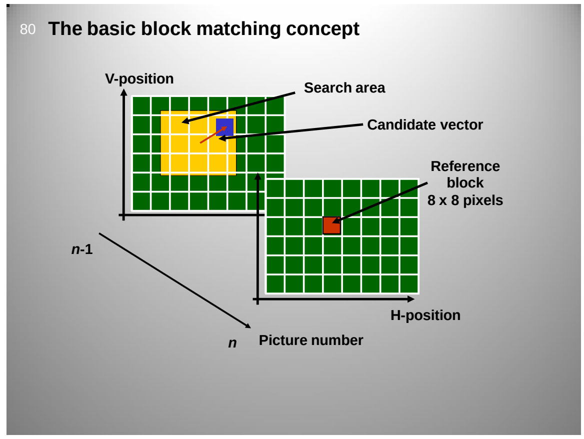
75

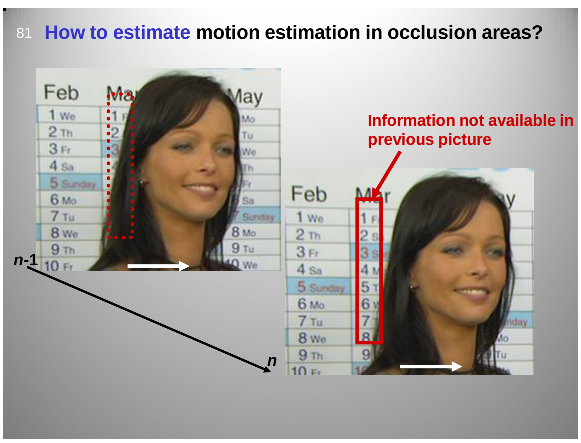
Block-hopping

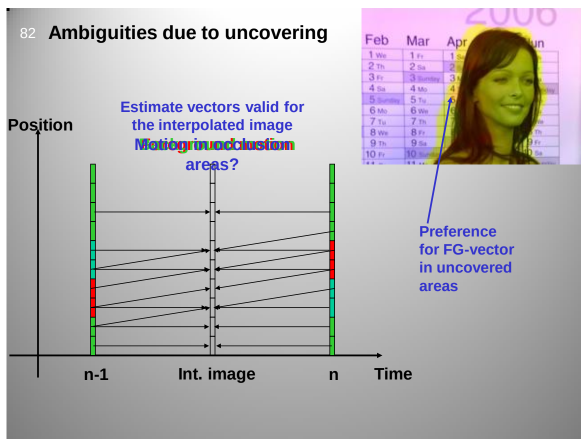


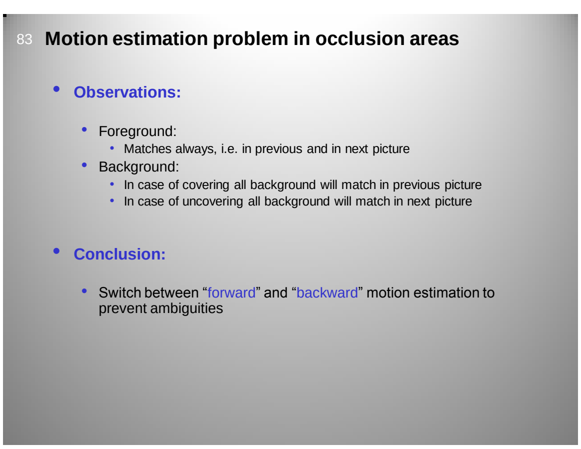


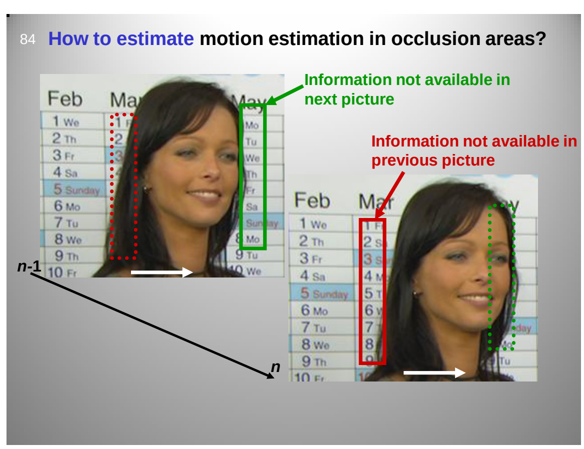

79
Motion estimation and occlusion


80 The basic block matching concept
V-position
Search area
Candidate vector
Reference block 8 x 8 pixels
n-1
H-position
n Picture number

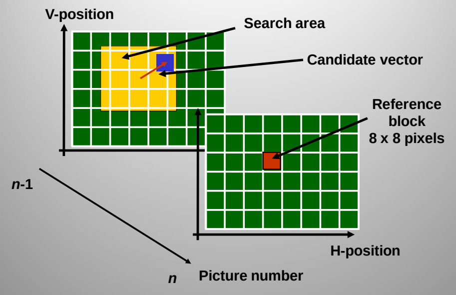

81 How to estimate motion estimation in occlusion areas?
Information not available in previous picture
n-1
n


82 Ambiguities due to uncovering
Position
Estimate vectors valid for the interpolated image
Motion in occlusion areas?
Preference for FG-vector in uncovered areas
n-1
Int. image
n
Time

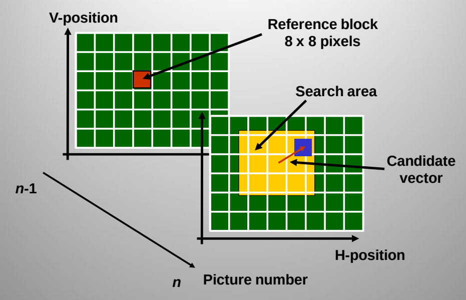

83 Motion estimation problem in occlusion areas
Observations:
Foreground:
Matches always, i.e. in previous and in next picture
Background:
In case of covering all background will match in previous picture
In case of uncovering all background will match in next picture
Conclusion:
Switch between “forward” and “backward” motion estimation to prevent ambiguities


84 How to estimate motion estimation in occlusion areas?
Information not available in next picture
Information not available in previous picture
n-1
n

85 **Solution:** In covering areas “forward” estimation



86 **Solution:** In uncovering areas “backward” estimation



87 **3-frame motion estimation**

Calculate SAD with past AND future frame, choose smallest

88 **Comparison 2 frame and 3 frame motion estimation**

2 frame ME

3 frame ME

89

Global motion estimation

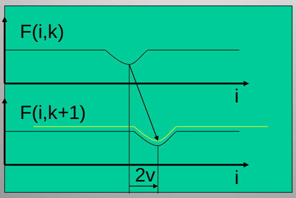
90 **Projection based global motion estimation**

- Algorithm:
  - Accumulate luminance over all lines
  - Accumulate luminance over all collumns
  - Determine global H- and V- motion based on these projections

Demo
Samsung ME

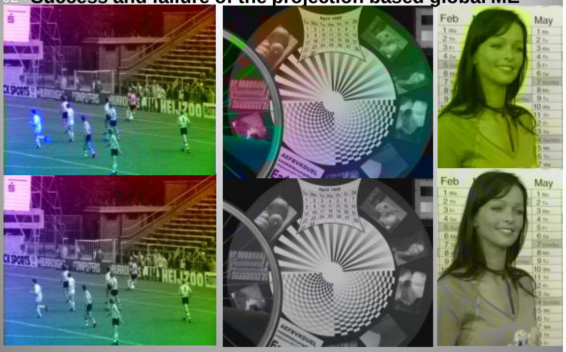
91 **Projection based global motion estimation**

- Global motion:minimum SAD of projection current and previous image
- Ticker tape detection:
  - Minimum of MC SADs only for restricted area



DEMO
Global ME

92 **Success and failure of the projection based global ME**



93 **Conclusions**

- Motion estimators for scan rate conversion differ from estimators for coding, due to additional true-motion constraint
- True motion results from constraints like spatial and temporal consistency
- Such constraints can be forced differently
  - hierarchical approach (e.g. Phase Plane Correlation.)
  - recursive approach (3-D RS)
  - Object based motion estimation

94 **Conclusions**

- Picture rate conversion requires very consistent but not necessarily very accurate motion vectors (integer resolution sufficient), the range should be at least +/-16 pixels
- De-interlacing requires very accurate motion vectors (at least 1/4 pixel) . For larger vectors the accuracy is less important