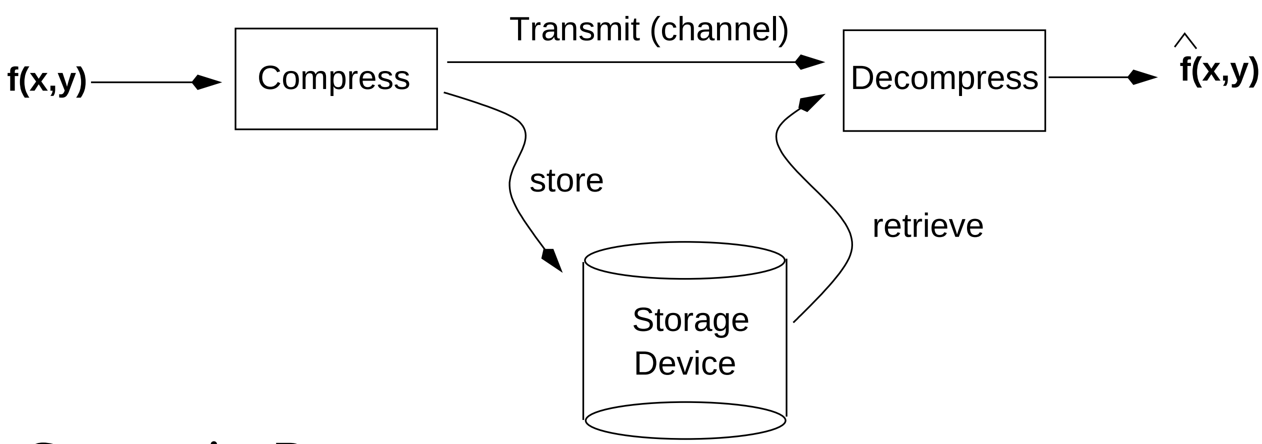
Image Compression

• Introduction

- The goal of image compression is the reduction of the amount of data required to represent a digital image.

- The idea is to remove redundant data from the image (i.e., data which do not affect image quality significantly)

- Image compression is very important for image storage and image transmission



• Compression Rates

- Advanced compression techniques can achieve compression ratios in the range **10:1** to **50:1** without visibly affecting image quality.

- Very high compression ratios of up to **2000:1** can be achieved in compressing video signals.

- In order for a compression system to be useful, compression and decompression must be very fast

• Compression Techniques

Lossless:

- Information preserving

- Low compression ratios

Lossy:

- Not information preserving

- High compression ratios

Tradeoff: image quality **vs** compression ratio

• Main Steps

(1) First, we may want to divide the image into fixed size blocks (e.g., as in JPEG). Then, we choose a set of basis functions that has some desired properties.

(2) Transform the image by projecting it into the chosen basis.

(3) Quantize the coefficients.

(4) Coding (further compress the coefficients using lossy or lossless compression techniques).

Wavelet noise removal

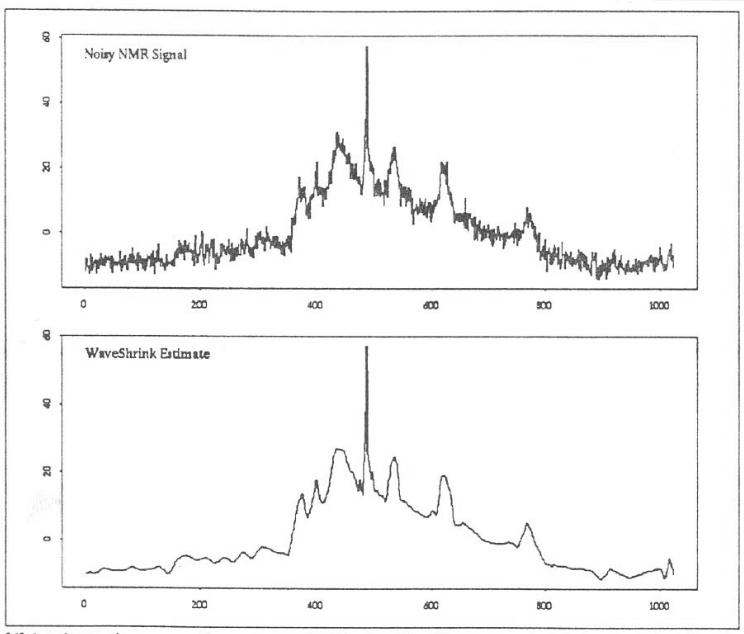
(A. Bruce, D. Donoho, and H. Gao, "Wavelet analysis", *IEEE Spectrum*, pp. 26-35, October 1996)

• Why and how?

- Traditional techniques remove noise by low-pass filtering, thus blurring sharp features in the underlying signal.

- Using wavelets, we set the coefficients below a given threshold to zero, then take the inverse transform to reconstruct the signal minus the noise.

- Wavelet noise removal has been shown to work well for geophysical signals, astronomical data, synthetic aperture radar, acoustic data, infrared images, and biomedical signals.



[4] A noisy nuclear magnetic resonance (NMR) signal [top] is compared with the same signal without the noise [bottom]. Setting to zero the coefficients that do not exceed a certain threshold and then inverting the transform gets rid of the noise in the wavelet domain. An important feature of the noise-removal algorithm is its ability to remove noise while simultaneously preserving nonsmooth features, such as the large spike in the NMR signal. Data are from the laboratory of Adrian Maudsley, University of California, San Francisco.

Content-based image retrieval using wavelets

(C. Jacobs, A. Finkelstein, and D. Salesin, "Fast multiresolution image quering", *Proceedings of SIGGRAPH*, pp. 277-286, 1995)

• Query images

- The query is an approximation of the image to be retrieved, expressed in two possible forms:

(1) A low-resolution image from a scanner or video camera.

(2) A rough sketch of the image painted by the user.

- The query image is typically very different from the "target" image.

(a) 1|2 (b) 1|1 (c)

• Requirements

- An effective "image query metric" is required to accomodate image distortions.

- Retrieval should be fast enough to handle tens of thousands of images at interactive rates.

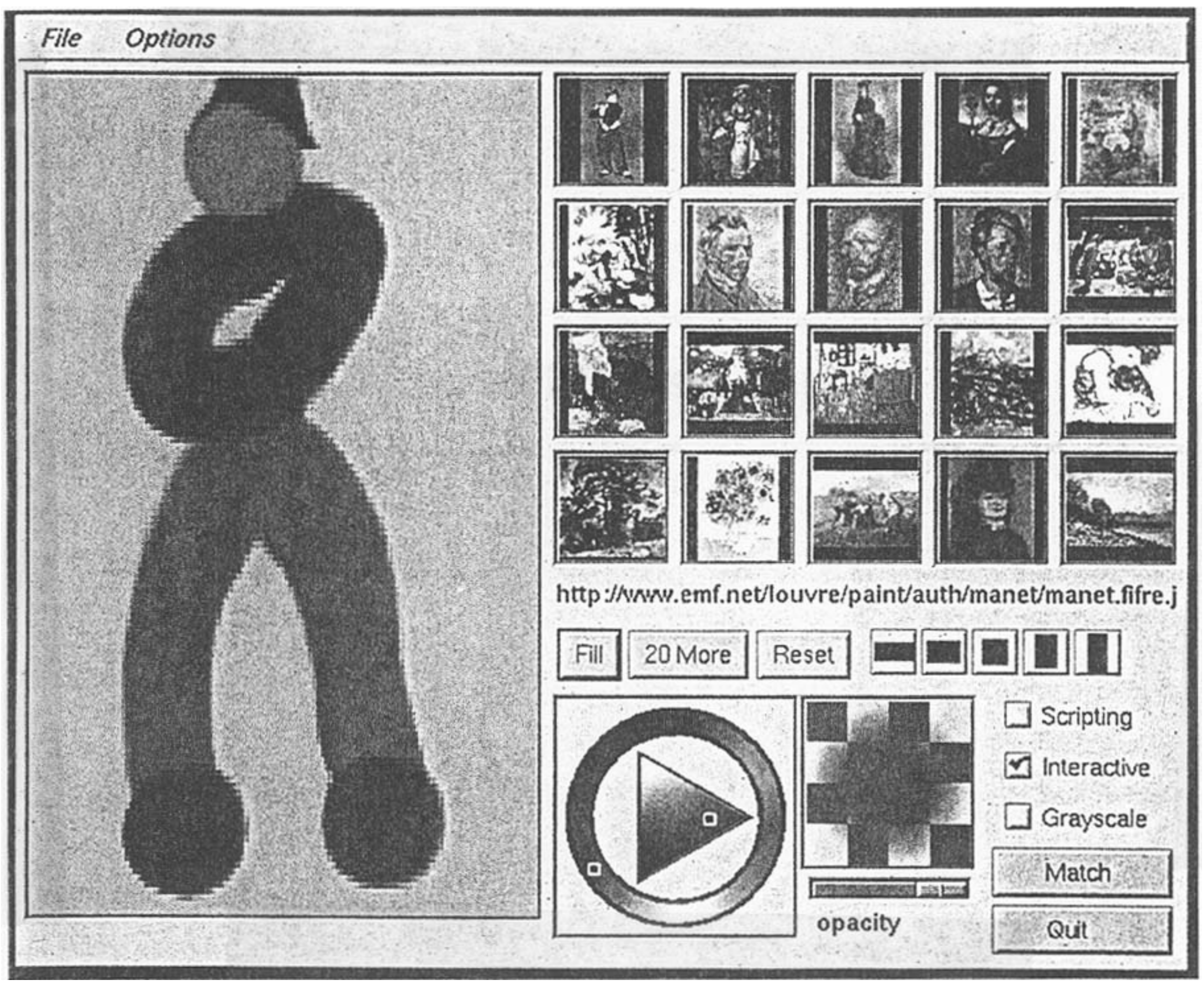
• Overview of proposed method

- The image query metric is based on truncated, quantized versions of the wavelet coefficients (*signature*).

- A novel database organization is used for computing this metric fast.

- System retrieves top 20 matches.

- The system processes a 128 x 128 image query on a database of 20,000 images in under 0.5 seconds.



• **Some advantages**

- The use of wavelets allows a query to be specified at any resolution (e.g., different from that of the target).

- The signature can be extracted from a wavelet-compressed version of the image directly.

- Simple to implement and use algorithm.

• **Common metrics**

- Metrics based on the L_1 and L_2 norms cannot handle inexact matching and are time consuming.

$$||Q - T||_1 = \sum_{i,j} |Q[i, j] - T[i, j]|$$

$$||Q - T||_2 = (\sum_{i,j}(Q[i, j] - T[i, j])^2)^{1/2}$$

- Experiments performed using these metrics have shown that the target image is in the highest 1% of the retrieved images only 3% of the time.

• **Components of the metric**

Color space

- YIQ seems to be the most appropriate for their data.

Wavelet type

- Haar wavelets are the fastest to compute and simplest to implement.

Truncation

- Keep only the coefficients with largest magnitude.

- This accelerates the search for a query and reduces storage requirements.

- The 60 largest coefficients in each channel worked best for painted queries.

- The 40 largest coefficients in each channel worked best for scanned queries.

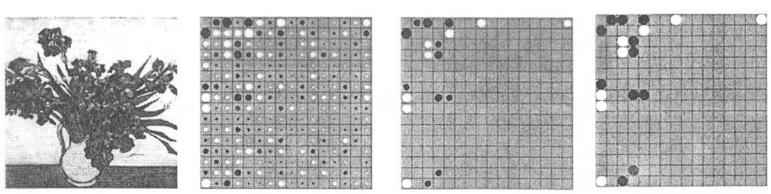
Quantization

- Quantize each of the retained coefficients into three levels: +1, 0 and -1

- Large positive coefficients are quantized to +1 and large negative coefficients are quantized to -1

- The mere presence or absence of these coefficients appears to have more discriminatory power that their precise magnitudes.

- Comparisons can be done much faster and efficiently now.



(a) (b) (c) (d)

FIGURE 5.1 Preprocessing steps: van Gogh's painting "Irises" (a) is first decomposed into wavelet coefficients (b). Next, all but the *m* largest-magnitude coefficients are truncated (c). Finally, the remaining coefficients are quantized (d). In the diagrams above, wavelet coefficients are represented by black and white dots. A dot's color (black or white) gives the sign (positive or negative) of the coefficient it represents. A dot's radius gives the magnitude of the coefficient.

• Wavelet-based metric

- Suppose Q and T represent a single channel of the wavelet decomposition of the query and target images.

- Let $Q[0, 0]$ and $T[0, 0]$ be the scaling function coefficients corresponding to the average intensity of that channel.

- Let $\hat{Q}[i, j]$ and $\hat{T}[i, j]$ represent the truncated, quantized coefficients of Q and T.

$$||Q - T|| = w_{0,0}|Q[0, 0] - T[0, 0]| + \sum_{i,j} w_{i,j}|\hat{Q}[i, j] - \hat{T}[i, j]|$$

• Simplifying the metric

- The above metric is equivalent to

$$||Q - T|| = w_{0,0}|Q[0, 0] - T[0, 0]| + \sum_{i,j} w_{i,j}(\hat{Q}[i, j] \neq \hat{T}[i, j])$$

where $(\hat{Q}[i, j] \neq \hat{T}[i, j])$is 1 if it is true and 0 otherwise.

- Group terms together into "buckets" so that only a small number of weights $w_{i,j}$ needs to be determined.

- Consider only the terms for which $\hat{Q}[i, j] \neq 0$

(1) allows for a query without much detail to match a very detailed target image.

(2) does not allow a detailed query to match a target that does not contain the same detail.

$$||Q - T|| = w_0|Q[0, 0] - T[0, 0]| + \sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)}(\hat{Q}[i, j] \neq \hat{T}[i, j])$$

• Fast computation of the metric

- It is quicker to count the number of matching coefficients than the number of mismatching coefficients (i.e., the majority of database images will not match the query image)

(this has to do with the data structure used to speed-up search - see later)

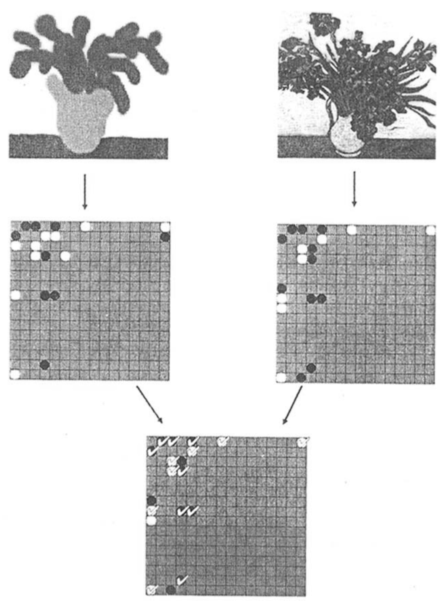
$$(\hat{Q}[i, j] \neq \hat{T}[i, j]) = 1 - (\hat{Q}[i, j] = \hat{T}[i, j])$$

$$\|Q - T\| = w_0|Q[0,0] - T[0,0]| + \sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)} -$$

$$\sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)}(\hat{Q}[i, j] = \hat{T}[i, j])$$

- The term $\sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)}$ does not depend on the target image, we can ignore it in ranking the target images:

$$\|Q - T\| = w_0|Q[0,0] - T[0,0]| - \sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)}(\hat{Q}[i, j] = \hat{T}[i, j])$$



• Algorithm

Preprocessing

(1) Perform a 2D Haar wavelet decomposition of every image in the database.

(2) Store the overall average color and the indices and signs of the m wavelet coefficients of largest magnitude.

(3) Organize the indices for all the images into a single data structure to optimize searching.

Quering

(1) Perform the same wavelet decomposition on the query image.

(2) Throw away all but the average color and the largest m coefficients.

(3) Compute the score of each target image using the above equation.

• Preprocessing (details)

- To optimize the search process, the m coefficients from every image are organized into a set of six arrays (*search arrays*).

- There is an array for every combination of sign (+ or -) and color channel (Y, I, and Q):

$$D_+^Y, D_+^I, D_+^Q, D_-^Y, D_-^I, D_-^Q$$

- The element $D_+^c[i, j]$, for example, contains a list of all images T having a large positive wavelet coefficient $T[i, j]$ in color channer c.

• Quering (details)

- Compute a score for each target image by looping through each color channel c.

- First compute the difference between the query's average intensity in that channel $Q^c[0, 0]$ and those in the database.

- For each of the m nonzero, truncated wavelet coefficients $Q^c[i, j]$, go through the list corresponding to $D^c_+[i, j]$ or $D^c_-[i, j]$ (i.e., depending on the sign of $Q^c[i, j]$).

- Update the score of each image found in those lists.

- Return the 20 closest matches.

```
func ScoreQuery(Q : array[0..r − 1, 0..r − 1] of color; m : int):
    DecomposeImage(Q)
    Initialize scores[i] ← 0 for all i
    for each color channel c do:
        for each database image T do:
            scores[index(T)] += w^c[0] ∗ |Q^c[0,0] − T^c[0,0]|
        end for
        Q̃ ← TruncateCoefficients(Q, m)
        for each non-zero coefficient Q̃^c[i, j] do
            if Q̃^c[i, j] > 0 then
                list ← D^c_+[i, j]
            else
                list ← D^c_−[i, j]
            end if
            for each element ℓ of list do
                scores[index(ℓ)] −= w^c[bin(i, j)]
            end for
        end for
    end for
    return scores
end func
```

$$||Q - T|| = w_0|Q[0,0] - T[0,0]| - \sum_{i,j:\hat{Q}[i,j]\neq 0} w_{bin(i,j)}(\hat{Q}[i,j] = \hat{T}[i,j])$$

- The function $bin(i, j)$ groups different coefficients into a small number of bins (6 bins per color channel):

$$bin(i, j) = min(max(i, j), 5)$$

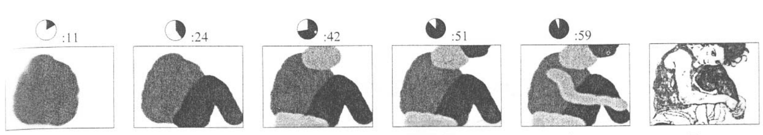
- Each bin is weighted by some constant $w[b]$ (the weights were found experimentally)

	Painted			Scanned		
b	$w^Y[b]$	$w^I[b]$	$w^Q[b]$	$w^Y[b]$	$w^I[b]$	$w^Q[b]$
0	4.04	15.14	22.62	5.00	19.21	34.37
1	0.78	0.92	0.40	0.83	1.26	0.36
2	0.46	0.53	0.63	1.01	0.44	0.45
3	0.42	0.26	0.25	0.52	0.53	0.14
4	0.41	0.14	0.15	0.47	0.28	0.18
5	0.32	0.07	0.38	0.30	0.14	0.27

• **Examples**

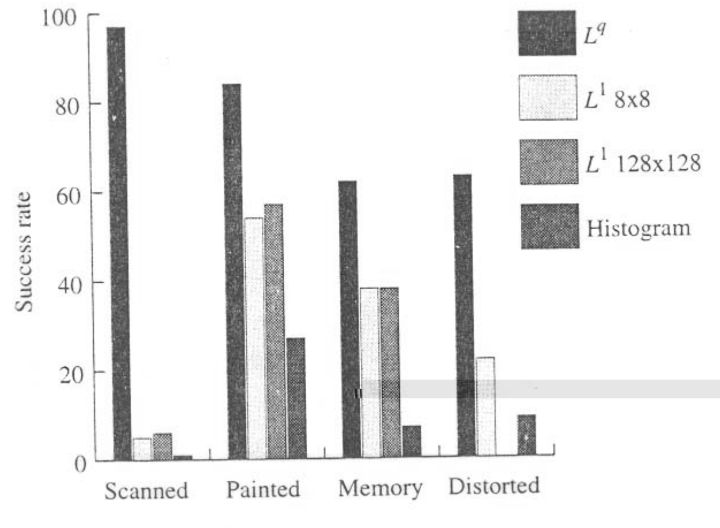
- Query examples using painted/scanned queries (database sizes: 1093 | 20,558)

(a) 1 | 2 (b) 1 | 1 (c)



25 | 945 1 | 171 1 | 29 1 | 16 1 | 1 Target

- Success rate of the proposed metric



-T ime requirements of the proposed metric

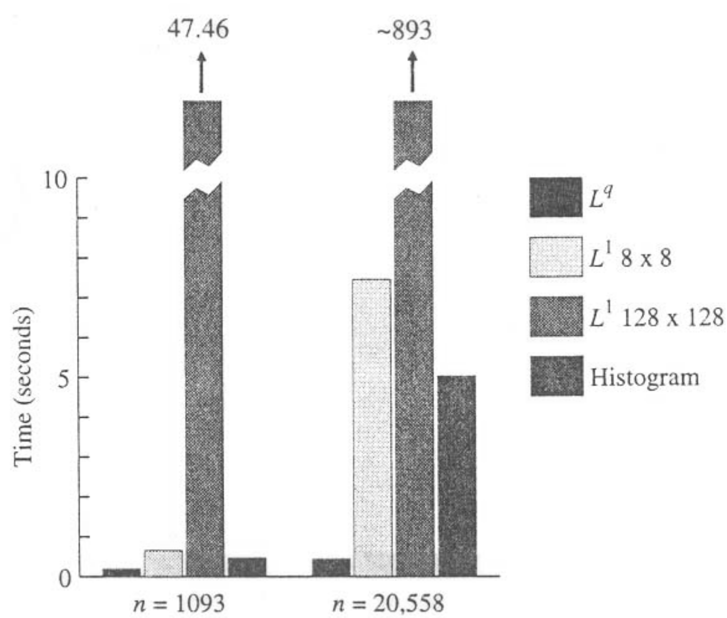


Image Fusion Using the Wavelet Transform

(H. Li, B. Manjunath, and S. Mitra, "Multisensor Image Fusion Using the Wavelet Transform", *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235-245, 1995).

• Image fusion

- The goal of image fusion is to integrate complementary information from multisensor data.

- The new images are more suitable for the purpose of human visual perception and computer processing tasks (e.g., segmentation, feature extraction, object recognition).

- A wavelet-based approach is proposed in this paper.

• Classification of fusion methods

<u>Signal-level fusion:</u> combination of a group of sensors with the objective of producing a signal of better quality and reliability.

<u>Pixel-level fusion:</u> increase the useful information content of an image.

<u>Feature level fusion:</u> enables the detection of useful features with higher confidence.

<u>Symbol-level fusion:</u> information is combined at a higher level of abstraction.

- The wavelet-based approach belongs to the "pixel-level fusion" category.

- **Main idea**

  - Compute the wavelet transform of the input images.

  - Combine the wavelet coefficients (see below how).

  - Take the inverse wavelet transform of the fused wavelet coefficients.

- **Assumptions**

  - The images to be combined have already perfectly registered.

  - Image registration insures that the information from every sensor refers to the same physical structure in the environment.

  - Registration can be done by proper arrangement of the sensors or by finding corresponding features between the images to be combined (see references).

- **Traditional methods**

  <u>Averaging</u>

  Simplest possible fusion method.

  It reduces the contrast of the features.

  <u>Laplacian pyramid</u>

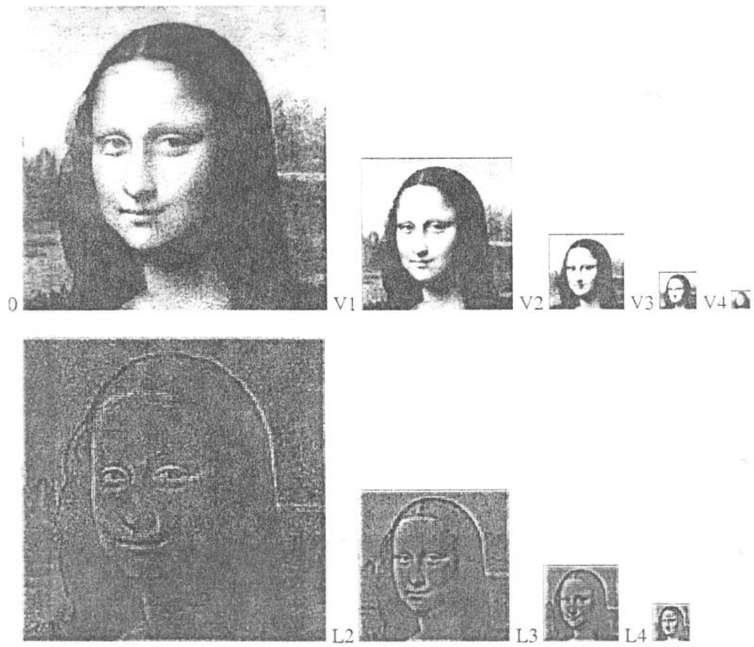
  Convolve images with Laplacians of varying width.

  Features cannot be localized accurately as width increases.

  Laplacian cannot not provide orientation selectivity.

  There is redundancy between different scales.

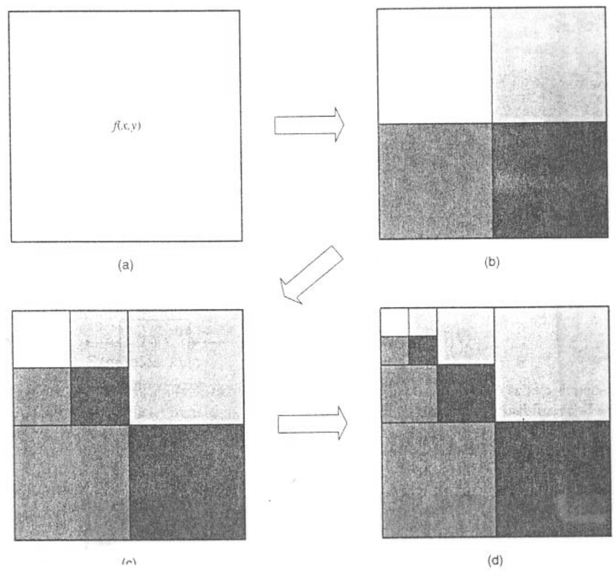
  Much more memory consuming compared to the proposed method.



• **Wavelet decomposition (using filter banks)**

- Image is decomposed into four subimages (subbands) each time:

(1) low-low (image at coarser resolution)
(2) low-high (sensitive to horizontal orientations)
(3) high-low (sensitive to vertical orientations)
(4) high-high (sensitive to diagonal orientations)



- They use the QMF implementation which does not require that the size of the image is a power of 2.

- The size of the wavelet transformed image is the same as the size of the original image.

• The image fusion scheme

(1) Compute the wavelet transform of the input images.

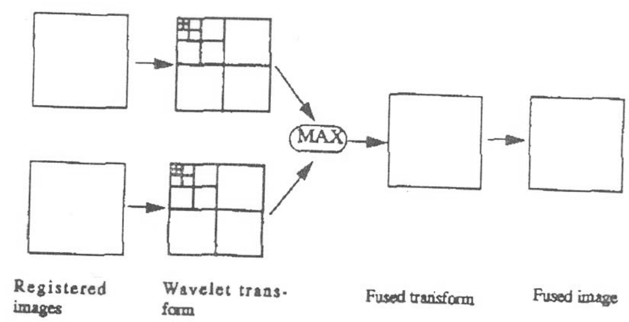
(2) Select the larger (absolute value) of the two wavelet coefficients at each point

(3) Reconstruct the fused image by performing an inverse wavelet transform using the fused coefficients.

- Interpretation of the above idea:

  * Large coefficient values correspond to sharper brightness changes (i.e., salient features such as edges, lines etc.).

  * If the same object appears more distinctly (i.e., has better contrast) in image A than in image B, after fusion the object in image A will be preserved while the object in image B will be ignored.


MAX
Registered images
Wavelet trans-form
Fused transform
Fused image

• Is the inverse wavelet transform stable using this idea?

- The reconstruction base on traditional methods (i.e., Laplacian pyramid) can be unstable in the regions where the two images are different (blocking effect).

- No blocking effect or other artifacts have been observed using the inverse wavelet transform of the fused coefficients.

• Modified algorithm (area-based criterion)

- Coefficient by coefficient selection might not be appropriate since most useful features in an image correspond usually to more than one coefficient.

- Keep the maximum coefficient (absolute) value within a 3x3 or 5x5 window in each image (i.e., assign the max value to the pixel corresponding to the center of the window).

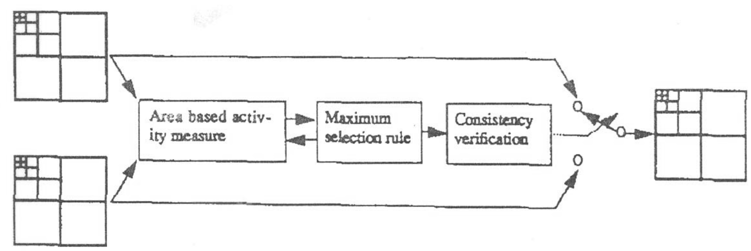
<my figure>

- Create a binary decision map (same size as the wavelet transformed image) to record the selection results based on the above rule (i.e., 1 if the max comes from image A and 0 if it comes from image B).

- Once the binary decision map is computed, apply the following *consistency verification* criterion:

(1) If the center pixel value comes from image A while the majority of the surrounding pixel values come from image B, the center pixel value is switched to that of image B.

(2) If the center pixel value comes from image B while the majority of the surrounding pixel values come from image A, the center pixel value is switched to that of image A.



• Performance measures

- In most cases, the criterion is application dependent.

- Quite commonly, the fusion results are evaluated visually (no quantitative performance measures since it is difficult to define the ideal fused image).

- In this study:

* Create test images for which the desired fusion result is known.

* Define the performance measure as follows (standard deviation of the difference between ideal and fused):

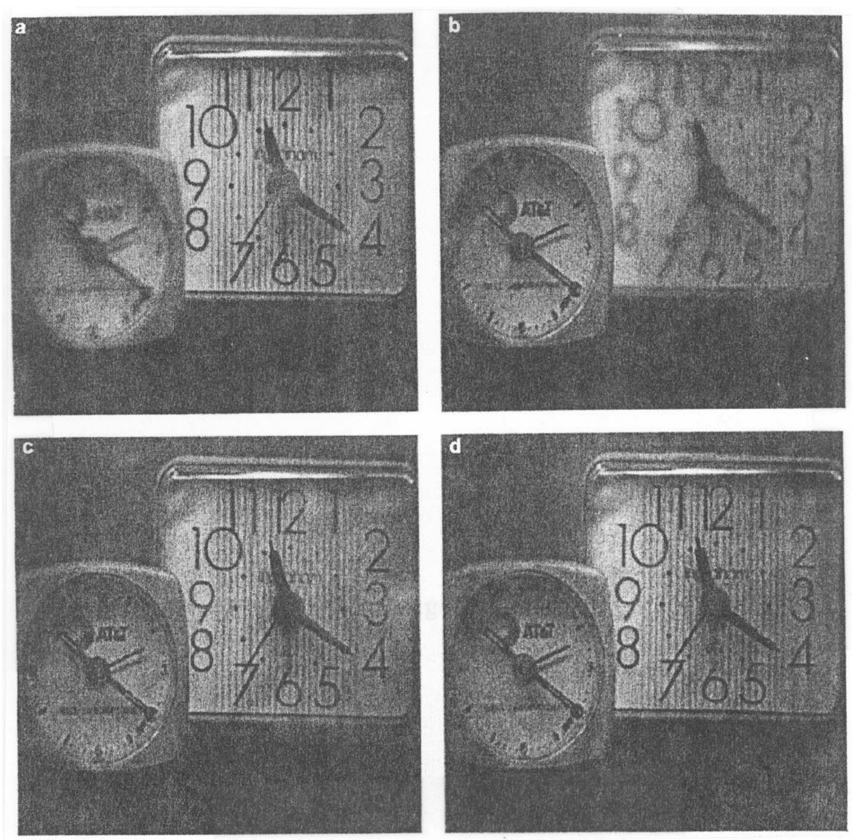
$$\rho = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}[I_{id}(i,j) - I_{fd}(i,j)]^2}{N^2}}$$

• Experimental results

- Seven sets of wavelet filters were chosen (see p. 241)

- Fusion was applied in the following cases:

  (1) multifocus images
  (2) Landsat and Spot images
  (3) Landsat and SAR images
  (4) IR and visible images
  (5) MRI and PET images



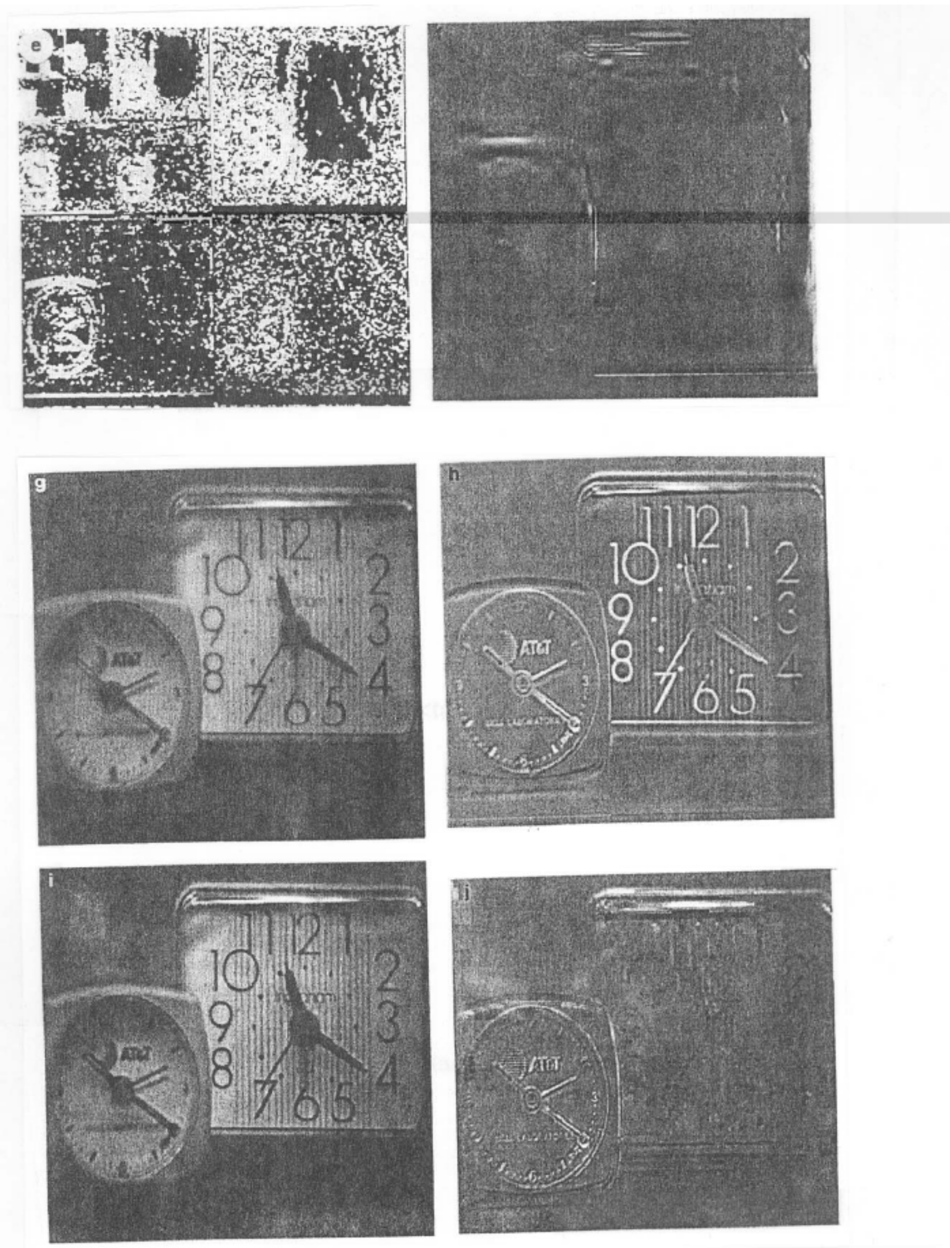


FIG. 5. (a) and (b) A pair of registered images with different focus points. (c) shows a perfectly fused image obtained by manual cut and (d) shows the fusion result using the wavelet transform and the modified feature selection rule. In these images both clocks are in focus. (e) sh binary decision map corresponding to the outcome of the proposed selection rule. (f) shows the normalized difference image between (c) ar (g) shows the fused image obtained by pixel averaging. (h) shows the normalized difference image between (c) and (g). (i) shows the fused obtained by the Laplacian pyramid-based method. (j) shows the normalized difference image between (c) and (i).

TABLE 1

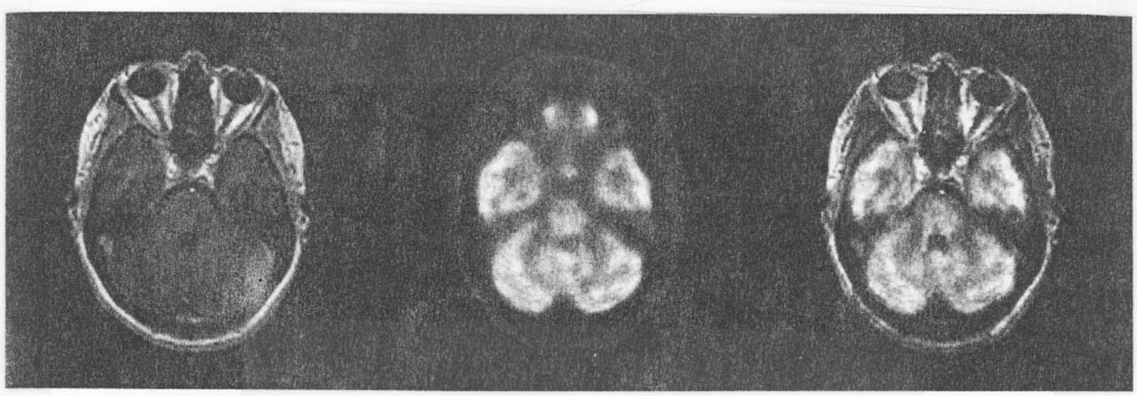
The Standard Deviation ρ of the Difference Images between the Manually Fused Image and the Fused Images Using Various Image Fusion Algorithms

Pixel-by-pixel average	Laplacian pyramid	Wavelet transform		
		Point-based selection rule	Area-based selection rule	Area-based selection rule and consistency verification
5.444	4.668	4.085	3.569	3.279

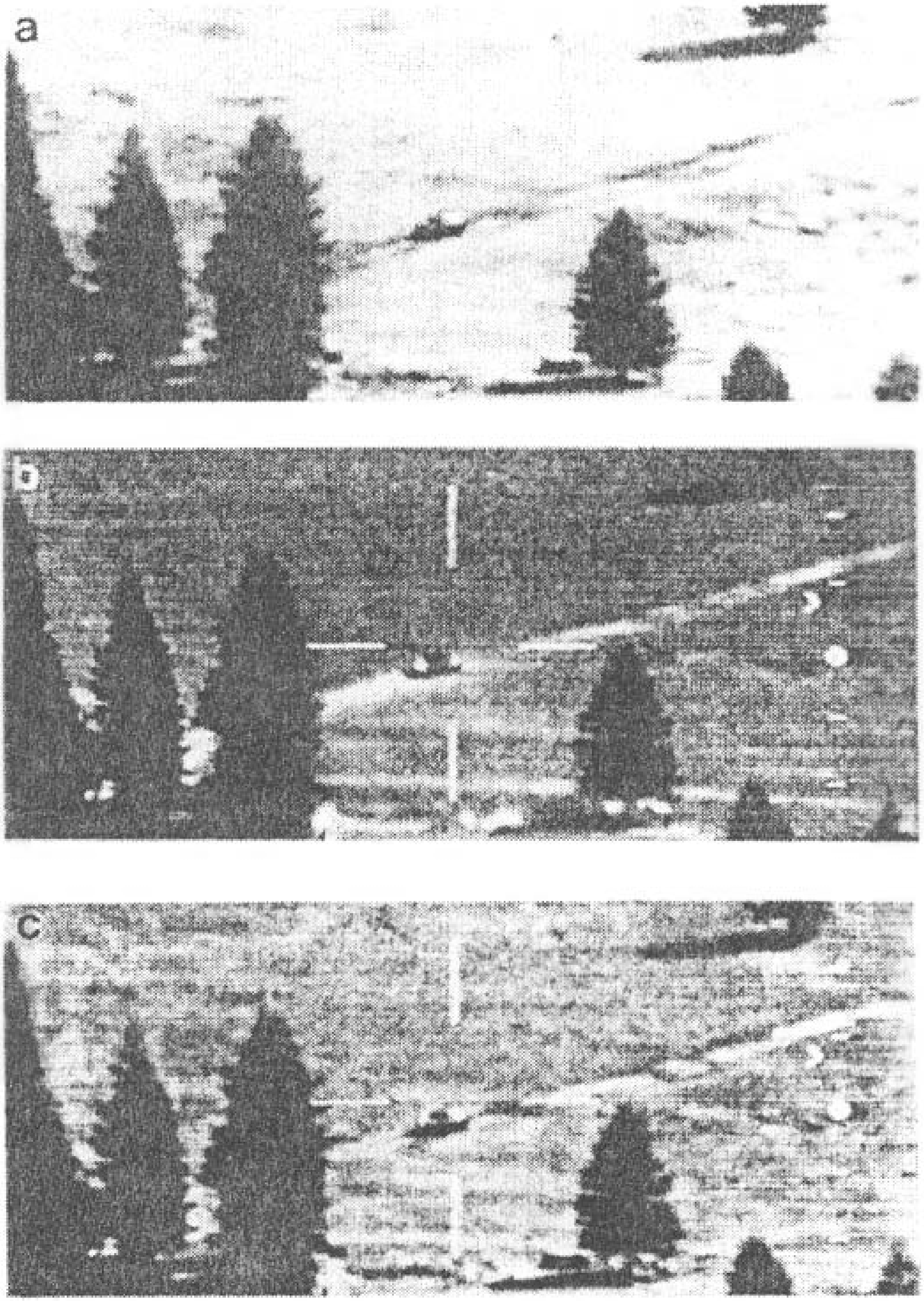
TABLE 2

The Standard Deviation ρ of the Difference Images between the Manually Fused Image and the Fused Images Using Different Sets of Filter Coefficients for the Wavelet Transform

Filter coefficients set						
1	2	3	4	5	6	7
3.388	3.279	3.829	3.559	3.296	5.105	3.490



(MRI provides anatomic information, PET provides functional information)



(fusion of visible with IR could improve target recognition)