9

A Perceptive-oriented Approach to Image Fusion

Boris Escalante-Ramírez[1], Sonia Cruz-Techica[1],
Rodrigo Nava[1] and Gabriel Cristóbal[2]
[1]*Facultad de Ingeniería, Universidad Nacional Autónoma de México,*
[2]*Instituto de Óptica, CSIC*
[1]*Mexico,*
[2]*Spain*

1. Introduction

At present time image fusion is widely recognized as an important tool and has attracted a great deal of attention from the research community with the purpose of searching general formal solutions to a number of problems in different applications such as medical imaging, optical microscopy, remote sensing, computer vision and robotics.

Image fusion consists of combining information from two or more images from the same sensor or from multiple sensors in order to improve the decision making process.

Fused images from multiple sensors, often called multi-modal image fusion system, include at least, two image modalities ranging from visible to infrared spectrum and they provide several advantages over data images from a single sensor (Kor & Tiwary, 2004). An example of this can be found in medical imaging where it is common to merge functional activity as in single photon emission computed tomography (SPECT), positron emission tomography (PET) or magnetic resonance spectroscopy (MRS) with anatomical structures such as magnetic resonance image (MRI), computed tomography (CT) and ultrasound, which helps improve diagnostic performance and surgical planning (Guihong et al., 2001, Hajnal et al., 2001).

An interesting example of single sensor fusion can be found in remote sensing, where pansharpening is an important task that combines panchromatic and multispectral optical data in order to obtain new multispectral bands that preserve their original spectral information with improved spatial resolution.

Depending on the merging stage, common image fusion schemes can be classified into three categories: pixel, feature and decision levels (Pohl & van Genderen, 1998). Many fusion schemes usually employ pixel level fusion techniques but since features, that are sensitive to human visual system (HVS), are bigger than a pixel and they exist in different scales, it is necessary to apply multiresolution analysis which improves the reconstruction of relevant image features (Nava et al., 2008). Moreover, the image representation model used to build the fusion algorithm must be able to characterize perceptive-relevant image primitives.

In the literature several methods of pixel level fusion have been reported using a transformation to perform data fusion, some of these transformations are: intensity-hue-saturation transform (IHS), principal component analysis (PCA) (Qiu et al., 2005), the

discrete wavelet transform (DWT) (Aguilar et al., 2007, Chipman et al., 1995, Li et al., 1994), dual-tree complex wavelet transform (DTCWT) (Kingsbury, 2001, Hill & Canagarajah, 2002), the contourlet transform (CW) (Yang et al., 2007), the curvelet transform (CUW) (Mahyari & Yazdi, 2009), and the Hermite transform (HT) (Escalante-Ramírez & López-Caloca, 2006, Escalante-Ramírez, 2008). In essence, all these transformations can discriminate between salient information and constant or non-textured background.

Of all these methods, the wavelet transform has been the most used technique for the fusion process. However, this technique presents certain problems in the analysis of signals of two or more dimensions, examples of these are the points of discontinuity that cannot always be detected, and its limitation to capture directional information. The contourlet and the curvelet transforms have shown better results than the wavelet transform due to their multi-directional analysis, but they require an extensive orientation search at each level of the decomposition. In contrast, the Hermite transform provides significant advantages to the process of image fusion. First, this image representation model includes some of the more important properties of the human visual system such as the local orientation analysis and the Gaussian derivative model of primary vision (Young, 1986), it also allows multiresolution analysis, so it is possible to describe the salient structures of an image at different spatial scales, and finally, it is steerable, which allows efficiently representing oriented patterns with a small number of coefficients. The latter has the additional advantage of reducing noise without introducing artifacts.

Hereinafter, we assume the input images have negligible registration problems, thus the images can be considered registered. The proposed scheme fuses images at the pixel level using a multiresolution directional-oriented Hermite transform of the source images by means of a decision map. This map is based on a linear dependence test of the Hermite coefficients within a fixed windows size; if the coefficients are linearly dependent, this indicates the existence of a relevant pattern that must be present in the final image.

The proposed algorithm has been tested on both multi-focus and multi-modal image sets producing results that ovecome results achieved with other methods such as wavelets (Li et al., 1994), curvelets (Donoho & Ying, 2007), and contourlets (Yang et al., 2008, Do, 2005). In addition to this, we used other decision rules proving that our scheme best characterized important structures of the images at the same time that the noise was reduced.

2. The Hermite transform as an image representation model

The Hermite transform (HT) (Martens 1990a, Martens 1990b) is a special case of polynomial transform, which is used to locally decompose signals and can be regarded as an image description model. The analysis stage involves two steps. First, the input image $L(x,y)$ is windowed with a local function $\omega(x,y)$ at several equidistant positions in order to achieve a complete description of the image. In the second step the local information of each analysis window is expanded in terms of a family of orthogonal polynomials. The polynomials $Gm,n\text{-}m(x,y)$ used to approximate the windowed information are determined entirely by the window function in such a way that the orthogonality condition is satisfied:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \omega^2(x,y) G_{m,n-m}(x,y) G_{l,k-l}(x,y)\,dxdy = \delta_{nk}\delta_{ml} \tag{1}$$

for $n, k=0,1,\ldots,\infty$; $m=0,\ldots,n$ y $l=0,\ldots,k$; where δ_{nk} denotes the Kronecker function.

The polynomial transform is called Hermite transform if the windows used are Gaussian functions. The Gaussian window is isotropic (rotationally invariant), separable in Cartesian coordinates and their derivatives mimic some processes at the retinal or visual cortex of the human visual system (Martens, 1990b, Young, 1986). This window function is defined as follows

$$\omega(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{2}$$

In a Gaussian window function, the associated orthogonal polynomials are the Hermite polynomials, which are defined as

$$G_{n-m,m}(x,y) = \frac{1}{\sqrt{2^n(n-m)!m!}} H_{n-m}\left(\frac{x}{\sigma}\right) H_m\left(\frac{x}{\sigma}\right) \tag{3}$$

where $H_n(x)$ denotes the nth Hermite polynomial.
The original signal $L(x,y)$, where (x,y) are the pixel coordinates, is multiplied by the window function $\omega(x-p,y-q)$ at the positions (p,q) that conform the sampling lattice S. By replicating the window function over the sampling lattice, we can define the periodic weighting function as

$$W(x,y) = \sum_{(p,q)\in S} \omega(x-p,y-q) \tag{4}$$

This weighting function must be a number other than zero for all coordinates (x,y). Therefore,

$$L(x,y) = \frac{1}{W(i,j)} \sum_{(p,q)\in S} L(x,y)\omega(x-p,y-q) \tag{5}$$

In every window function, the signal content is described as the weighted sum of polynomials $G_{m,n-m}(x,y)$ of m degree in x and $n-m$ in y. In a discrete implementation, the Gaussian window function may be approximated by the binomial window function and in this case, its orthogonal polynomials $G_{m,n-m}(x,y)$ are known as Krawtchouck's polynomials.
In either case, the polynomial coefficients $L_{m,n-m}(p,q)$ are calculated convolving the original image $L(x,y)$ with the analysis filters $D_{m,n-m}(x,y) = G_{m,n-m}(-x,-y)\omega^2(-x,-y)$, followed by subsampling at position (p,q) of the sampling lattice S. That is,

$$L_{m,n-m}(p,q) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} L(x,y) D_{m,n-m}(x-p,y-q)\,dxdy \tag{6}$$

The recovery process of the original image consists of interpolating the transform coefficients with the proper synthesis filters. This process is called inverse transformed polynomial and is defined by

$$\hat{L}(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\sum_{(p,q)\in S} L_{m,n-m}(p,q) P_{m,n-m}(x-p,y-q) \tag{7}$$

The synthesis filters $P_{m,n-m}(x,y)$ of order m and n-m, are defined by

$$P_{m,n-m}(x,y) = \frac{G_{m,n-m}(x,y)\omega(x,y)}{W(x,y)} \tag{8}$$

for m=0,…,n , and n=0,…,∞

2.1 The steered Hermite transform

The Hermite transform has the advantage of high-energy compaction by adaptively steering the HT (Martens, 1997, Van Dijk, 1997, Silván-Cárdenas & Escalante-Ramírez, 2006). Steerable filters are a class of filters that are rotated copies of each filter, constructed as a linear combination of a set of basis filters. The steering property of the Hermite filters explains itself because they are products of polynomials with a radially symmetric window function. The N +1 Hermite filters of Nth-order form a steerable basis for each individual Nth-order filter. Because of this property, the Hermite filters at each position in the image adapt to the local orientation content.

Thus, for orientation analysis, it is convenient to work with a rotated version of the HT. The polynomial coefficients can be computed through a convolution of the image with the filter functions $D_m(x)D_{n\_m}(y)$. They are separable in spatial and polar domains, and their Fourier transform can be expressed as $\omega_x=\omega cos\theta$ and $\omega_y=\omega sin\theta$, in polar coordinates, then

$$d_m(\omega_x)d_{n-m}(\omega_y) = g_{m,n-m}(\theta)d_n(\omega) \tag{9}$$

where $d_n(\omega)$ is the Fourier transform for each filter function expressed in radial frequency, given by

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}}(-j\omega\sigma)\exp\left(-\frac{\omega\sigma^2}{4}\right) \tag{10}$$

and the orientation selectivity for the filter is expressed by

$$g_{m,n-m}(\theta) = \sqrt{\binom{n}{m}}\cos^m\theta sen^{n-m}\theta \tag{11}$$

In terms of orientation frequency functions, this property of the Hermite filters can be expressed by

$$g_{m,n-m}(\theta-\theta_0) = \sum_{k=0}^{n} c^n_{m,k}(\theta_0)\, g_{n-k,k}(\theta) \tag{12}$$

where $c^n_{m,k}(\theta_0)$ is the steering coefficient.

The Hermite filter rotation at each position over the image is an adaptation to local orientation content. Fig. 1 shows the HT and the steered HT over an image. For the directional Hermite decomposition, first, a HT was applied and then the coefficients were rotated toward the estimated local orientation, according to a criterion of maximum oriented energy at each window position. This implies that these filters can indicate the direction of one-dimensional pattern independently of its internal structure.

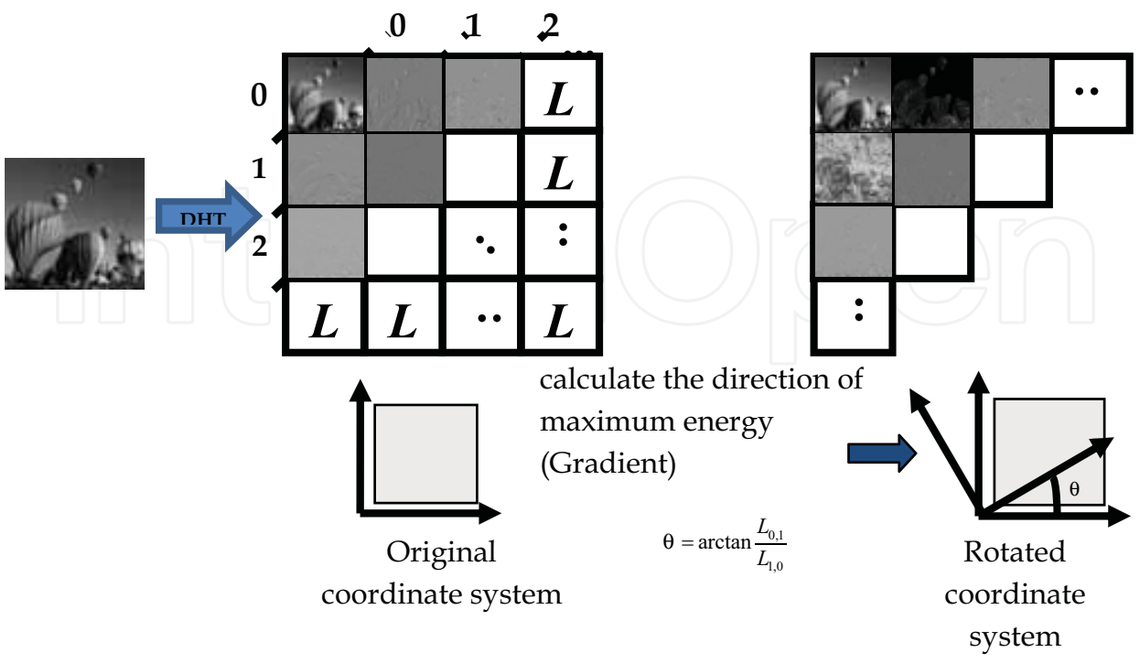

Fig. 1. The discrete Hermite transform (DHT) and the steered Hermite transform over an image

The two-dimensional Hermite coefficients are projected onto one-dimensional coefficients on an axis that makes an angle θ with the x axis, this angle can be approximated as $\theta=L_{01}/L_{10}$, where L_{01} and L_{10} are a good approach to optimal edge detectors in the horizontal and vertical directions respectively.

2.2 The multiresolution directional oriented HT

A multiresolution decomposition using the HT can be obtained through a pyramid scheme (Escalante-Ramírez & Silván Cárdenas 2005). In a pyramidal decomposition, the image is decomposed into a number of band-pass or low-pass subimages, which are then subsampled in proportion to their spatial resolution. In each layer the zero order coefficients are transformed to obtain -in a lower layer- a scaled version of the above. Once the coefficients of Hermite decomposition of each level are obtained, the coefficients can be projected to one dimension by its local orientation of maximum energy. In this way we obtain the multiresolution directional-oriented Hermite transform, which provides information about the location and orientation of the structure of the image at different scales.

3. Image fusion with the Hermite transform

Our approach aims at analyzing images by means of the HT, which allows us to identify perceptually relevant patterns to be included in the fusion process while discriminating spurious artifacts. As we have mentioned, the steered HT allows us to focus energy in a small number of coefficients, and thus the information contained in the first-order rotated

coefficient may be sufficient to describe the edge information of the image in a particular spatial locality. If we extend this strategy to more than one level of resolution, then it is possible to obtain a better description of the image. However, the success of any fusion scheme depends not only on the image analysis model but also on the fusion rule, therefore, instead of choosing for the usual selection operators based on the maximum pixel value, which often introduce noise and irrelevant details in the fused image, we seek a rule to consider the existence of a pattern in a region defined by a fixed-size window.

The general framework for the proposed algorithm includes the following stages. First a multiresolution HT of the input images is applied. Then, for each level of decomposition, the orientation of maximum energy is detected to rotate the coefficients, so the first order rotated coefficient has most edge information. Afterwards, taking this rotated coefficient of each image we apply a linear dependence test. The result of this test is then used as a decision map to select the coefficients of the fused image in the multiresolution HT domain of the input images. If the original images are noisy, the decision map is applied on the multiresolution directional-oriented HT. The approximation coefficients in the case of HT are the zero order coefficients. In most multifocal and multimodal applications the approximation coefficients of the input images are averaged to generate the zero order coefficient of the fused image, but it always depends on the application context. Finally the fused image is obtained by applying the inverse multiresolution HT. Fig. 2 shows a simplified representation of this method.

3.1 The fusion rule

The linear dependence test evaluates the pixels inside a window of w_s x w_s, if those pixels are linearly independent, then there is no relevant feature in the window. However, if the pixels are linearly dependent, it indicates the existence of a relevant pattern. The fusion rule selects the coefficient with the highest dependency value. A higher value will represent a stronger pattern. A simple and rigorous test for determining the linear dependence or independence of vectors is the Wronskian determinant. The dependency of the window centered at a pixel *(i,j)* is described in

$$D_A(i,j) = \sum_{m=i-w_s}^{i+w_s} \sum_{n=j-w_s}^{j+w_s} L_A^2(m,n) - L_A(m,n) \tag{13}$$

where $L_A(m, n)$ is the first order steered Hermite coefficient of the source image A with spatial position *(m,n)*. The fusion rule is expressed in (14). The coefficient of the fused HT is selected as the one with largest value of the dependency measure.

$$L_F(i,j) = \begin{cases} L_A(i,j) & si \;\; D_A(i,j) \geq D_B(i,j) \\ L_B(i,j) & si \;\; D_A(i,j) < D_B(i,j) \end{cases} \tag{14}$$

We apply this rule to all detail coefficients and in the most of the cases average the zero order Hermite coefficients as (15).

$$L_{00_F}(i,j) = \frac{1}{2}\left[L_{00_A}(i,j) + L_{00_B}(i,j) \right] \tag{15}$$

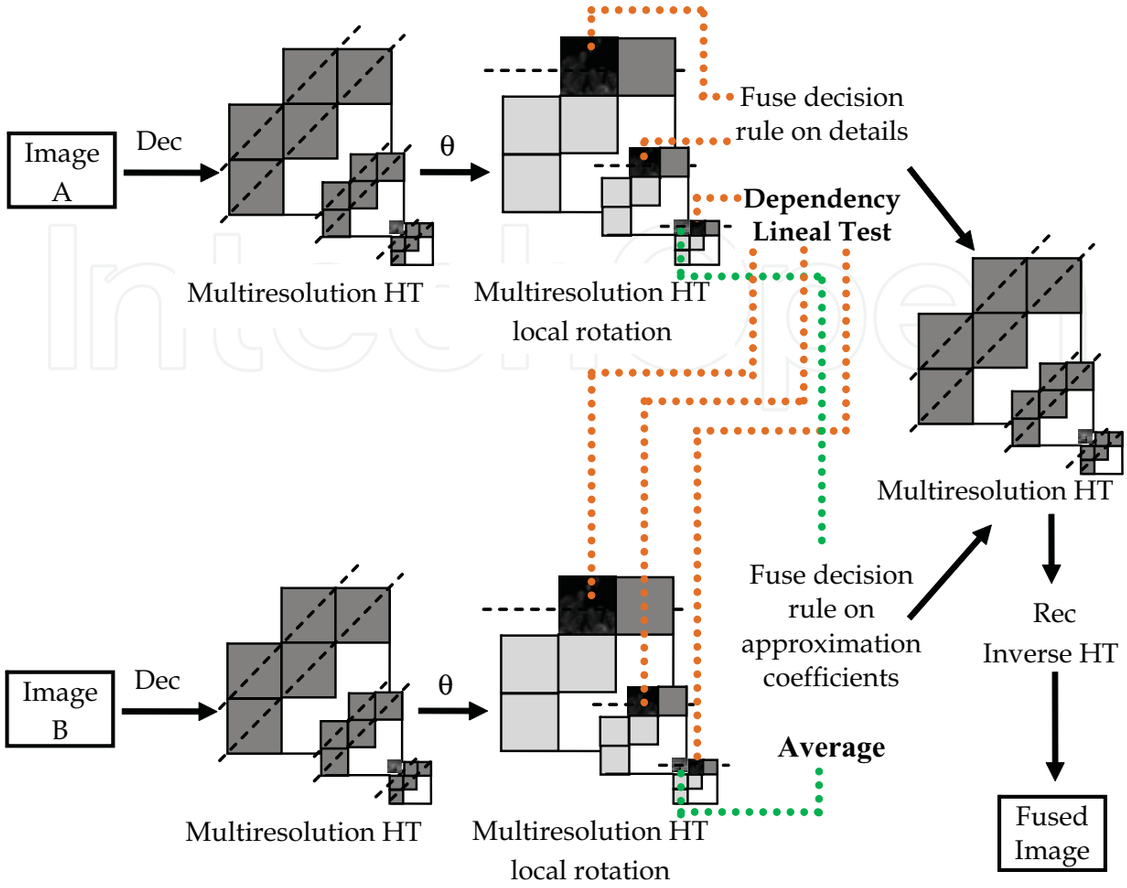

Fig. 2. Fusion scheme with the multiresolution directional-oriented Hermite transform

4. Image fusion results

The proposed algorithm was tested on several sets of multi-focus and multi-modal images, with and without noise degradation. Fig. 3 shows one of the multi-focus image sets used and the results of image fusion achieved with the proposed method using different decision rules. In these experiments, we used a Gaussian window with spread $\sigma=\sqrt{2}$, a subsampling factor T=2 between each pyramidal level and four decomposition levels. The window size for linear dependence test, maximum with verification of consistency and saliency and match measurement (Burt & Kolczynski, 1993), was 3 x 3.

Fig. 4 shows other multi-focus image sets that uses synthetic images. The results of image fusion were achieved with different fusion methods using linear dependence as decision rule. In these experiments, we used a Gaussian window with spread $\sigma=\sqrt{2}$, a subsampling factor T=2 between each pyramidal level and three decomposition levels; the wavelet transform used was db4 and in the case of the contourlet transform, the McClellan transform of 9-7 filters were used as directional filters and the wavelet db4 was used as pyramidal filters. The window size for the fusion rule was 3 x 3. The results were zoomed with the purpose to better observe the different methods performance.

On the other side, Figs. 5, 6 and 7 show the application in medical images comparing with other fusion methods, all of them using the linear dependence test with a window size of 3 x

3. All the transforms have two decomposition levels; the wavelet transform used was db4 and in the case of the contourlet transform, the McClellan transform of 9-7 filters were used as directional filters and the wavelet db4 was used as pyramidal filters.
In Fig. 7, Gaussian noise with σ=0.001 was introduced to the original images in order to show the efficiency of our method in noisy images.

5. Quality assessment of image fusion algorithms

Digital image processing involves many tasks, such as manipulation, storing, transmission, etc., that may introduce perceivable distortions. Since degradations occur during the processing chain, it is crucial to quantify degradations in order to overcome them. Due to their importance, many articles on the literature are dedicated to develop methods for improving, quantifying or preserving the quality of processed images. For example, Wang and Bovik (Wang, et al, 2004) describe a method based on the hypothesis that the HVS is highly adapted for extracting structural information, and they proposed a measure of structural similarity (SSIM) that compares local patterns of pixel intensities that have been normalized for luminance and contrast. In (Nava, et al, 2010; Gabarda & Cristóbal, 2007) two quality assessment procedures were introduced based on the expected entropy variance of a given image. These methods are useful in scenarios where there is no reference image, therefore they can be used in image fusion applications.
Quality is an image characteristic, it can be defined as ``the degree to which an image satisfies the requirements imposed on it'' (Silverstein & Farrell, 1996) and it is crucial for most image processing applications, because it can be used to compare the performance of the different systems and to select the appropriate processing algorithm for any given application. Image quality (IQ) can be used in general terms as an indicator of the relevance of the information presented by an image. A major part of research activity in the field of IQ is directed towards the development of reliable and widely applicable image quality measure algorithms. Nevertheless, only limited success has been achieved (Nava, at al, 2008).
A common way to measure IQ is based on early visual models but since human beings are the ultimate receivers in most applications, the most reliable way of assessing the quality of an image is by subjective evaluations. There are several different methodologies for subjective testing which are based on the idea how a person perceives the quality of images, and so it is inherently subjective (Wang, et al, 2002).
The subjective quality measure, mean opinion score (MOS), provides a numerical indication of the perceived quality. It has been used for many years, and it is considered the best method for image quality. The MOS metric is generated by averaging the results of a set of standard, subjective tests, where a number of people rate the quality of image series based on the recommendation ITU-T J247 (Sheikh, el al, 2006). MOS is the arithmetic mean of all the individual scores, and can range from 1 (worst) to 5 (best).
Nevertheless, MOS is inconvenient because it demands human observers, it is expensive and usually too slow to apply in real-time scenarios. Moreover, quality perception is strongly influenced by a variety of factors that depend on the observer. For these reasons, it is desirable to have an objective metric capable of predict image quality automatically. The techniques developed to assess image quality must depend on the field of application because it determines the characteristics of the imaging task we would like to evaluate. Practical image quality measures may vary according to the field of application and they should evaluate overall distortions. However, there is no single standard procedure to measure image quality.

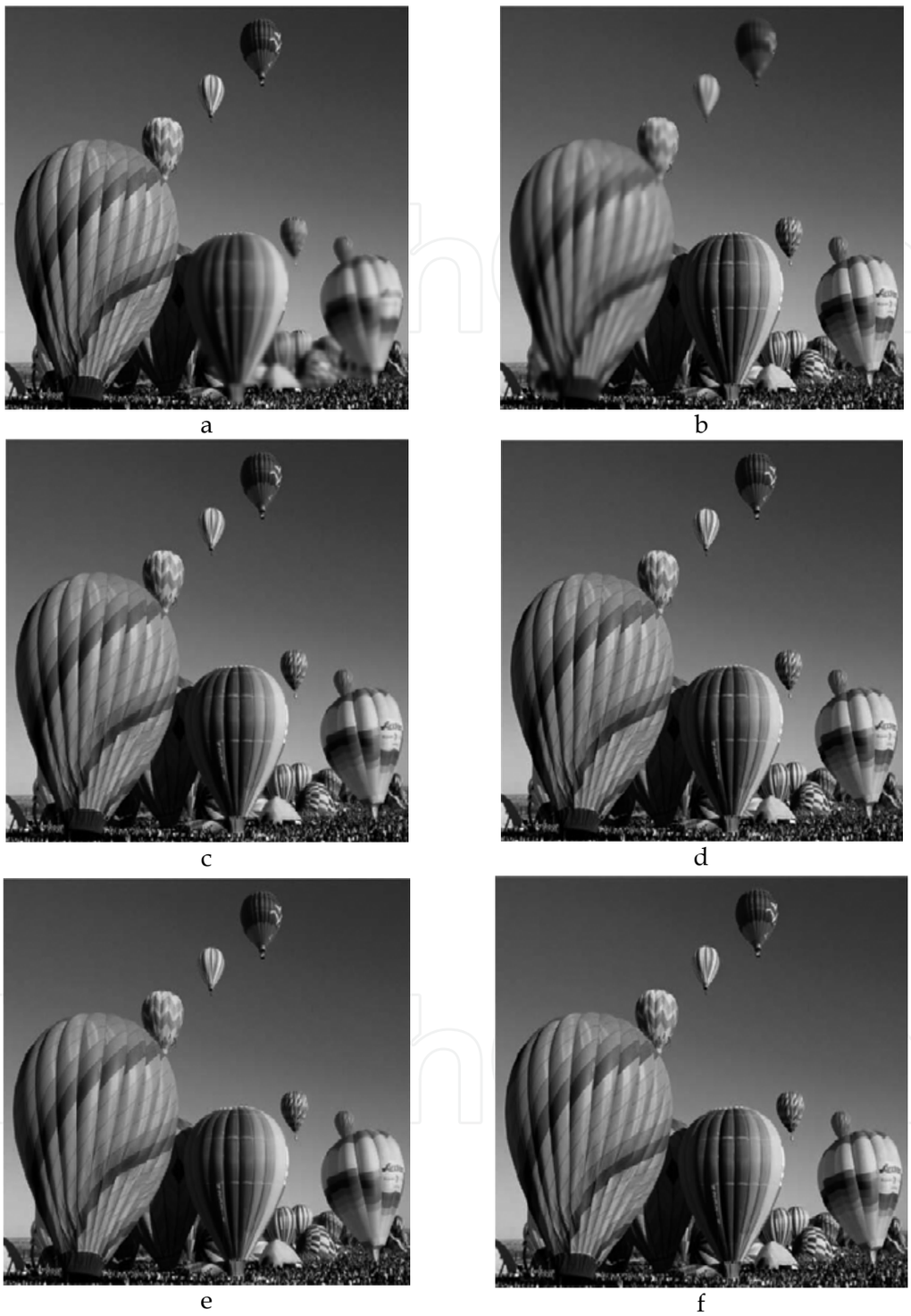

Fig. 3. Results of image fusion in multi-focus images, using multiresolution directional-oriented HT. a) and b) are the source images, c) fused image using absolute maximum selection, d) fused image using maximum with verification of consistency, e) fused image using saliency and match measurement and f) fused image using the linear dependency

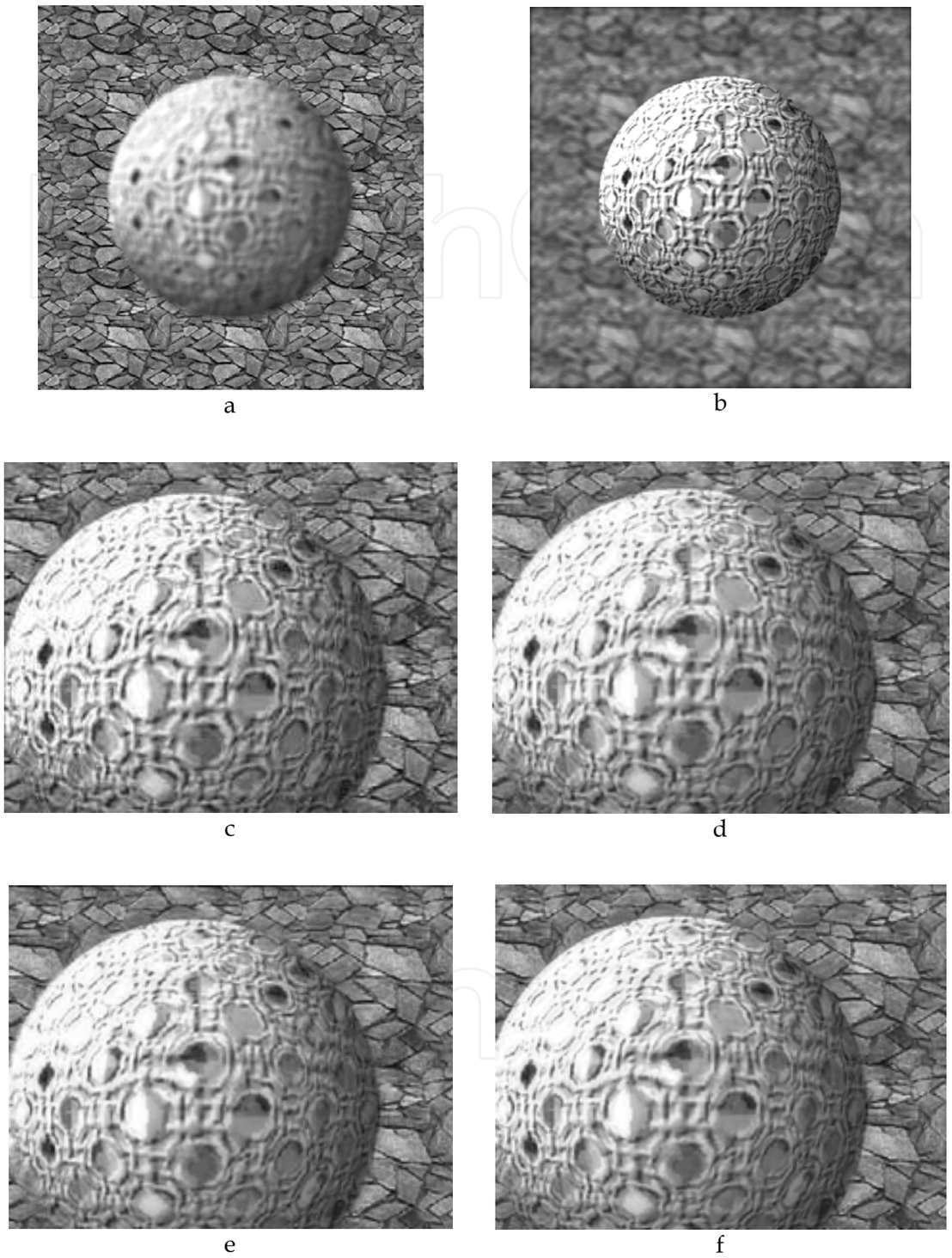

Fig. 4. Results of image fusion in synthetic multi-focus images, using the dependency test rule and different analyze techniques. a) and b) are the source images, c) HT, d) wavelet transform, e) contourlet transform and f) curvelet transform

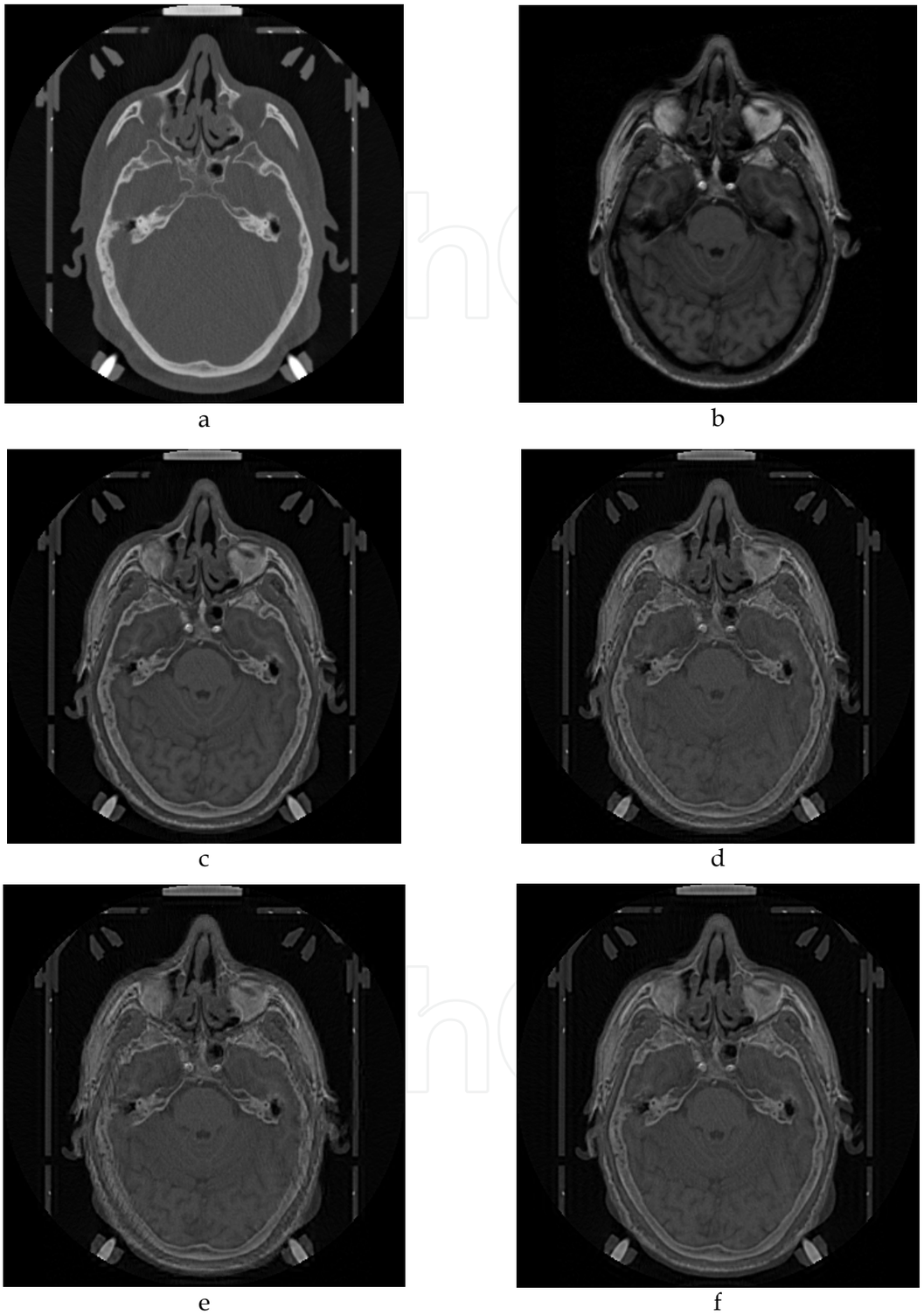

Fig. 5. Results of image fusion in medical images, using the dependency test rule and different analyze techniques. a) CT, b) MR, c) HT, d) wavelet transform, e) contourlet transform and f) curvelet transform

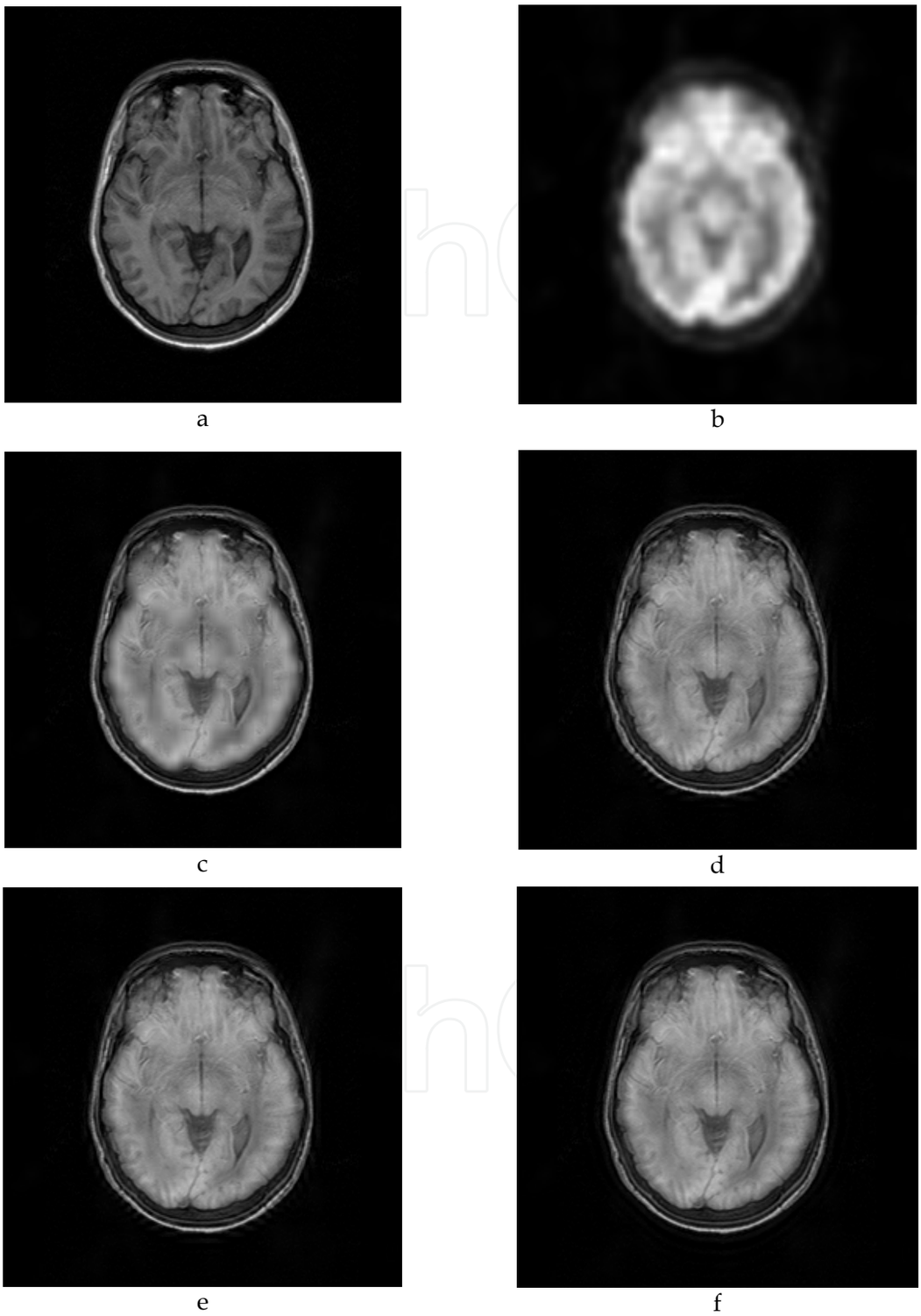

Fig. 6. Results of image fusion in medical images, using the dependency test rule and different analyze techniques. a) RM, b) PET, c) HT, d) wavelet transform, e) contourlet transform and f) curvelet transform

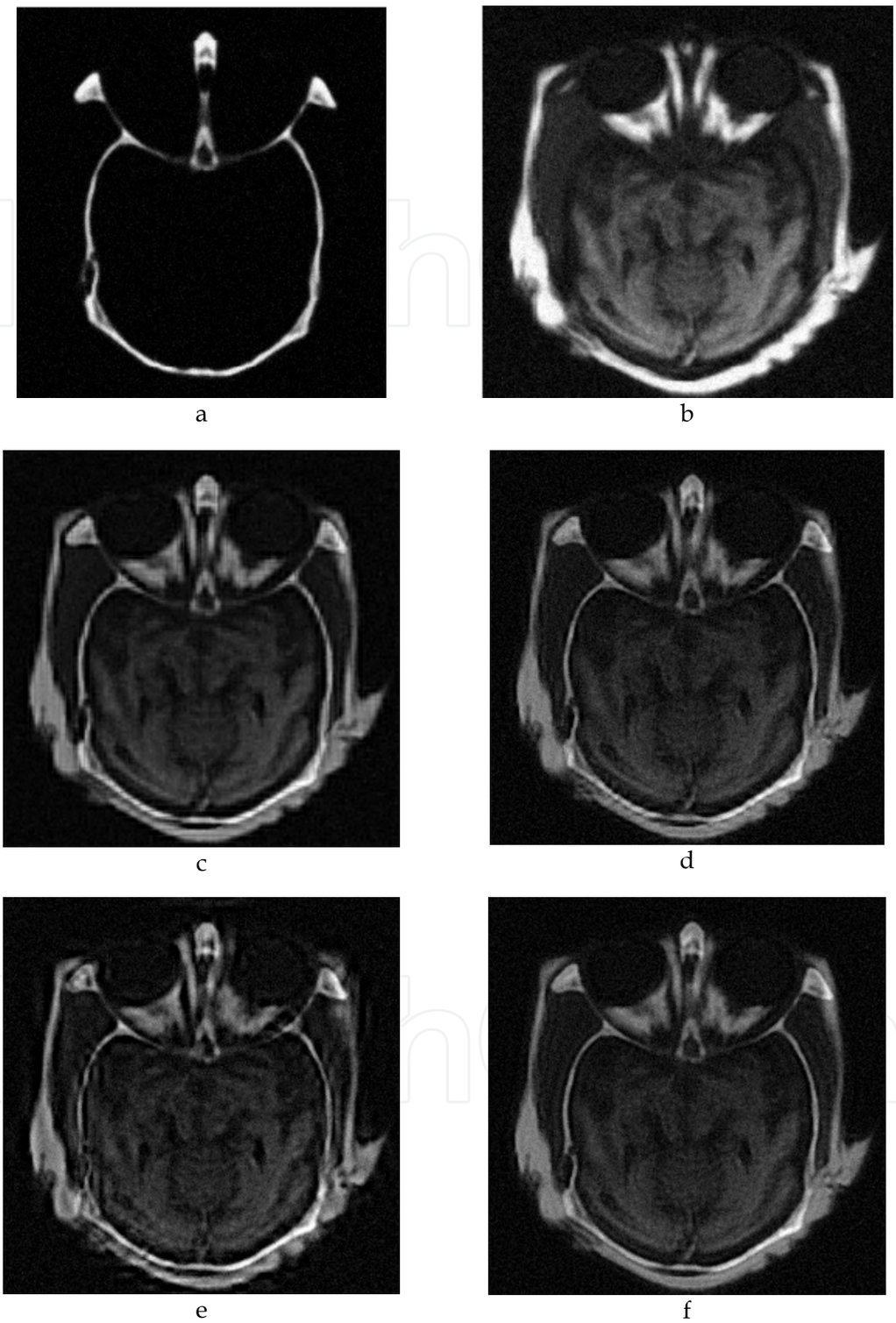

Fig. 7. Results of image fusion in medical images, using the dependency test rule and different analyze techniques. a) CT, b) MR, c) HT, d) wavelet transform, e) contourlet transform and f) curvelet transform. Images provided by Dr. Oliver Rockinger

Objective image quality metrics are based on measuring physical characteristics and they intend to predict perceived quality accurately and automatically. It means, that they should predict image quality that an average human observer will report. One important fact on this issue is the availability of an "original image", which is considered to be distortion-free or perfect quality. Most of the proposed objective quality measures assume that the reference image exists and they attempt to quantify the visibility error between a distorted image and a reference image.

Among the available ways to measure objective image quality, the mean squared error (MSE) and peak signal-to-noise ratio (PSNR) are widely employed because they are easy to calculate and usually they have low computational cost, but such measures are not necessarily consistent with human observer evaluation (Wang & Bovik, 2009). Both MSE and PSNR reflect the global properties of the image quality but they are inefficient in predicting structural degradations. Ponomarenko in (Ponomarenko, et al, 2009) evaluated correspondence of HVS with MSE and PSNR (0.525) where ideal value is 0.99. This shows that the widely used metrics PSNR and MSE have very low correlation with human perception (correlation factors are about 0.5).

In many practical applications, image quality metrics do not always have access to a reference image. However, it is desirable to develop measurement approaches that can evaluate image quality blindly. Blind or non--reference image quality assessment turns out to be a very difficult task, because metrics are not related to the original image (Nava, et al, 2007).

In order to quantitatively compare the different objective quality metrics, we evaluated our fusion results with several methods, including the traditional as well as some of the more recent ones that may correlate better with the human perceptive assessment. Among the first ones, we considered the PSNR and the MSE, and for the second group we used the measure of structural similarity (SSIM), the Mutual information (MI) and the Normalized Mutual Information (NMI) based on Tsallis entropy (Nava et al, 2010). In experiments with no reference image (ground truth) was available, metrics based on mutual information were used.

PSNR is a ratio between the maximum possible power of the reconstructed image and the power of the noise that affects the fidelity of the reconstruction, this is

$$PSNR = 10\log_{10}\frac{255^2(MN)}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[F(i,j)-R(i,j)\right]^2} \tag{16}$$

where *F(i,j)* denotes the intensity of the pixel of the fused image and *R(i,j)* denotes the intensity of the pixel of the original image.

The MSE indicates the error level between the fused image and the ideal image (ground truth), the smaller value of MSE indicates the better performance of the fusion method.

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[F(i,j)-R(i,j)\right]^2}{MN} \tag{17}$$

The SSIM (Wang et al., 2004) compares local patterns of pixel intensities that have been normalized for luminance and contrast and it provides a quality value in the range [0,1].

$$SSIM(R,F) = \frac{\sigma_{RF}}{\sigma_R \sigma_F} \frac{2\mu_R \mu_F}{(\mu_R)^2 + (\mu_F)^2} \frac{2\sigma_R \sigma_F}{\sigma_R^2 \sigma_F^2} \tag{18}$$

Where μ_R is the original image mean and μ_F the fused image mean; σ is the variance and σ_{RF} is the covariance.
MI has also been proposed as a performance measurement of image fusion in the absence of a reference image (Wang et al., 2009). Mutual information is a measurement of the statistical dependency of two random variables and the amount of information that one variable contain about the other. The amount of information that belongs to image A contained in the fused image is determined as follows:

$$MI_{FA}(I_F, I_A) = \sum P_{FA}(I_F, I_A) \log \left[\frac{P_{FA}(I_F, I_A)}{P_F(I_F) P_A(I_A)} \right] \tag{19}$$

where P_F and P_A are the marginal probability densisty functions of images F and A respectively, and P_{FA} is the joint probability density funtion of both images.Then, mutual information is calculated by

$$MI_F^{AB} = MI_{FA}(I_F, I_A) + MI_{FB}(I_F, I_B) \tag{20}$$

Another performance measurement is the Fusion Symmetry (FS) defined in equation (21), it denotes the symmetry of the fusion process in relation to the two input images. The smaller the FS is, the better the fusion process performs.

$$FS = abs\left(\frac{MI_{FA}(I_F, I_A)}{MI_{FA}(I_F, I_A) + MI_{FB}(I_F, I_B)} - 0.5 \right) \tag{21}$$

The NMI (Nava et al 2010) is defined as

$$NMI^q(F,A,B) = \frac{M^q(F,A,B)}{MAX^q(F,A,B)} \tag{22}$$

where

$$M^q(F,A,B) = I^q(F,A) + I^q(F,B) \tag{23}$$

and

$$I^q(F,A) = \frac{1}{1-q}\left(1 - \sum_{f,a} \frac{P(F,A)^q}{(P(F)P(A))^{q-1}} \right) \tag{24}$$

$MAX^q(F,A,B)$ is a normalization factor that represents the total information.
At first glance, the results obtained in Fig. 3 were very similar, thought quantitatively it is possible to verify the performance of the proposed algorithm. Table 1 shows the HT fusion

performance using a ground truth image and different fusion rules, while that Table 2 compares the performance of different fusion methods with the same reference image and the same fusion rule.

Fusion Rule	MSE	PSNR	SSIM	MI
Absolute maximum	4.42934	41.6674	0.997548	5.535170
Maximum with verification of consistency	0.44076	51.6886	0.999641	6.534807
Saliency and match measurement	4.66043	41.4465	0.996923	5.494261
Linear dependency test	0.43574	51.7385	0.999625	6.480738

Table 1. Performance measurement of Fig. 3 using a ground truth image by the multiresolution directional-oriented HT using different fusion rules

Fusion Method	MSE	PSNR	SSIM	MI	NMI
Hermite Transform	0.43574	51.7385	0.999625	6.480738	0.72835
Wavelet Transform	0.76497	49.2944	0.999373	6.112805	0.72406
Contourlet Transform	1.51077	46.3388	0.998464	5.885111	0.72060
Curvelet Transform	0.88777	48.6478	0.999426	6.083156	0.72295

Table 2. Performance measurement of Fig. 3 using a ground truth image applying the fusion rule based on linear dependency with different methods

Tables 3 and 4 correspond to tables 1 and 2 for the case of Fig. 4.

Fusion Rule	MSE	PSNR	SSIM	MI
Absolute maximum	54.248692	30.786911	0.984670	3.309483
Maximum with verification of consistency	35.110012	32.676494	0.989323	3.658905
Saliency and match measurement	38.249722	32.304521	0.989283	3.621530
Linear dependency test	33.820709	32.838977	0.989576	3.659614

Table 3. Performance measurement of Fig. 4 using a ground truth image by the multiresolution directional-oriented HT with different fusion rules

Fusion Method	MSE	PSNR	SSIM	MI	NMI
Hermite Transform	33.820709	32.838977	0.989576	3.659614	0.23967
Wavelet Transform	128.590240	27.038724	0.953244	2.543590	0.24127
Contourlet Transform	156.343357	26.190009	0.945359	2.323243	0.23982
Curvelet Transform	114.982239	27.524496	0.952543	2.588358	0.24024

Table 4. Performance measurement of Fig. 4 using a ground truth image applying the fusion rule based on linear dependency with different methods

From Figs. 5, 6 and 7, we can notice that the image fusion method based on the Hermite transform preserved better the spatial resolution and information content of both images. Moreover our method shows a better performance in noise reduction.

Fusion Method	MI_{FA}	MI_{FB}	MI_{FAB}	FS
Hermite Transform	1.937877	1.298762	3.236638	0.098731
Wavelet Transform	1.821304	1.202295	3.023599	0.102363
Contourlet Transform	1.791008	1.212183	3.003192	0.096368
Curvelet Transform	1.827996	1.268314	3.096310	0.090379

Table 5. Performance measurement of Fig. 5 (CT/RM) applying the fusion rule based on linear dependency with different methods

Fusion Method	MI_{FA}	MI_{FB}	MI_{FAB}	FS
Hermite Transform	1.617056	1.766178	3.383234	0.022038
Wavelet Transform	1.626056	1.743542	3.369598	0.017433
Contourlet Transform	1.617931	1.740387	3.358319	0.018232
Curvelet Transform	1.589712	1.754872	3.344584	0.024691

Table 6. Performance measurement of Fig. 6 (RM/PET) applying the fusion rule based on linear dependency with different methods

6. Conclusions

We have presented a multiresolution image fusion method based on the directional-oriented HT using a linear dependency test as fusion rule. We have experimented with this method for multi-focus and multi-modal images and we have obtained good results, even in the

presence of noise. Both subjective and objective results show that the proposed scheme outperforms other existing methods.
The HT has proved to be an efficient model for the representation of images because derivatives of Gaussian are the basis functions of this transform, which optimally detect, represent and reconstruct perceptually relevant image patterns, such as edges and lines.

7. Acknowledgements

This work was sponsored by UNAM grants IN106608 and IX100610.

8. References

Aguilar-Ponce, R.; Tecpanecatl-Xihuitl, J.L.; Kumar, A.; & Bayoumi, M. (2007). Pixel-level image fusion scheme based on linear algebra, *IEEE International Symposium on Circuits and Systems, 2007. ISCAS 2007* , pp. 2658–2661, May 2007.

Burt, P.J. & Kolczynski, R.J. (1993). Enhanced image capture through fusion, *Proceedings of the Fourth International Conference on Computer Vision, 1993*, pp. 173 –182, 11-14.

Chipman, L.J.; Orr, T.M. & Graham, L.N. (1995). Wavelets and image fusion, *Proceedings of the International Conference on Image Processing*, Vol. 3, pp. 248 –251, Oct. 1995.

Do, M (2005). Contourlet toolbox.
http://www.mathworks.com/matlabcentral/fileexchange/8837, (Last modified 27 Oct 2005).

Donoho, D. & Ying, L. (2007). The Curvelet.org team: Emmanuel Candes, Laurent Demanet. Curvelet.org. http://www.curvelet.org/software.html, (Last modified 24 August 2007).

Escalante-Ramírez, B. & López-Caloca, A.A. (2006). The Hermite transform: an efficient tool for noise reduction and image fusion in remote sensing, In: *Signal and Image Processing for Remote Sensing*, C.H. Chen, (Ed.), 539–557. CRC Press, Boca Raton.

Escalante-Ramírez, B. (2008). The Hermite transform as an efficient model for local image analysis: an application to medical image fusion. *Computers & Electrical Engineering*, Vol. 34, No. 2, 99–110, 2008.

Escalante-Ramírez, B & Silván-Cárdenas, J.L. (2005). Advanced modeling of visual information processing: A multi-resolution directional-oriented image transform based on gaussian derivatives. *Signal Processing: Image Communication*, Vol. 20, No. 9-10, 801 – 812.

Gabarda, S., & Cristóbal, G. (2007). Blind image quality assessment through anisotropy. Journal of the Optical Society of America A, Vol. 24, No. 12, B42--B51.

Guihong, Q.; Dali, Z. & Pingfan, Y. (2001). Medical image fusion by wavelet transform modulus maxima. *Optics Express*, Vol. 9 No. 4, 184–190.

Hajnal, J.; Hill, D.G. & Hawkes, D. (2001). *Medical Image Registration*. CRC Press, Boca Raton.

Hill, P; Canagarajah, N. & Bull, D. (2002). Image fusion using complex wavelets, *Proceedings of the. 13th British Machine Vision Conference*, pp. 487–496, 2002.

Kingsbury, N (2001). Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis*, Vol. 10, No. 3, 234 – 253.

Kor, S. & Tiwary, U. (2004) Feature level fusion of multimodal medical images in lifting wavelet transform domain, *Proceedings of the IEEE International Conference of the Engineering in Medicine and Biology Security,* Vol. 1, pp. 1479–1482, 2004.

Li, H; Manjunath, B.S. & Mitra, S.K. (1994). Multi-sensor image fusion using the wavelet transform, *Proceedings of the IEEE International Conference on Image Processing, ICIP-94,* Vol. 1, pp. 51 –55, Nov. 1994.

Mahyari, A.G. & Yazdi, M. (2009). A novel image fusion method using curvelet transform based on linear dependency test, *Proceedings of the International Conference on Digital Image Processing 2009*, pp. 351–354, March 2009.

Martens J.B. (1990a). The Hermite transform-applications. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 38, No. 9, 1607–1618.

Martens, J.B. (1990b). The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 38, No. 9, 1595–1606.

Martens, J.B. (1997). Local orientation analysis in images by means of the Hermite transform. *IEEE Transactions on Image Processing*, Vol. 6, No. 8, 1103–1116.

Nava, R., Cristóbal, G., & Escalante-Ramírez, B. (2007). Nonreference image fusion evaluation procedure based on mutual information and a generalized entropy measure, *Proceedings of SPIE conference on Bioengineered and Bioinspired Systems III*. Vol. 6592. Maspalomas, Gran Canaria, Spain, 2007.

Nava, R.; Escalante-Ramirez, B. & Cristobal, G (2008). A novel multi-focus image fusion algorithm based on feature extraction and wavelets, *Proceedings of SPIE*, vol. 5616800, 2008, pp. 700028-10.

Nava, R., Escalante-Ramírez, B., & Cristóbal, G. (2010). Blind quality assessment of multi-focus image fusion algorithms, *Proceedings of Optics, Photonics, and Digital Technologies for Multimedia Applications*. 7723, p. 77230F. Brussels, Belgium, Apr. 2010, SPIE.

Pohl, C. & Van Genderen, J.L.(1998). Multisensor image fusion in remote sensing: concepts, methods and applications, *International Journal of Remote Sensing*, Vol. 19, No. 5, 823–854.

Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M., & Battisti, F. (2009). TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics. *Advances of Modern Radioelectronics* Vol. 10, No. 4, 30-45.

Qiu, Y; Wu, J; Huang, H.; Wu, H.; Liu, J. & Tian, J. (2005). Multi-sensor image data fusion based on pixel-level weights of wavelet and the PCA transform, *Proceedings of the IEEE International Conference Mechatronics and Automation*, Vol. 2, 653 – 658, Jul. 2005.

Sheikh, H. R., Sabir, M. F., & Bovik, A. C. (2006). A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms. *IEEE Transactions on Image Processing* , 15 (11), 3440-3451.

Silván Cárdenas, J.L. & Escalante-Ramírez, B. (2006). The multiscale hermite transform for local orientation analysis, *IEEE Transactions on Image Processing,* Vol. 15, No. 5, 1236-1253.

Silverstein, D. A., & Farrell, J. E. (1996). The relationship between image fidelity and image quality, *Proceedings of the International Conference on Image Processing*, 1, pp. 881-884, 1996.

Van Dijk, A.M. & Martens, J.B. (1997). Image representation and compression with steered Hermite transforms. *Signal Processing*, Vol. 56, No. 1, 1 – 16.

Wang, Z., & Bovik, A. C. (2009). Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. *IEEE Signal Processing Magazine* , 26 (1), 98-117.

Wang, Z., Bovik, A. C., & Lu, L. (2002). Why is image quality assessment so difficult?, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. IV-3313-IV-3316), 2002.

Wang, Z.; Bovik, A.C.; Sheikh, H.R. & Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing*, Vol. 13, No. 4, 600 –612.

Wang, Q.; Yu, D, & Shen Y. (2009). An overview of image fusion metrics, *Proceedings of the Instrumentation and Measurement Technology Conference, 2009. I2MTC '09. IEEE*, pp. 918–923, May 2009.

Yang, L.; Guo, B.L. & and Ni, M. (2008). Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. *Neurocomputing*, Vol. 72, No. 1-3, 203 – 211, 2008. Machine Learning for Signal Processing (MLSP 2006) / Life System Modelling, Simulation, and Bio-inspired Computing (LSMS 2007).

Young, R. (1986). The gaussian derivative theory of spatial vision: analysis of cortical cell receptive field line-weighting profiles. Technical Report GMR-4920, General Motors Research, 1986.

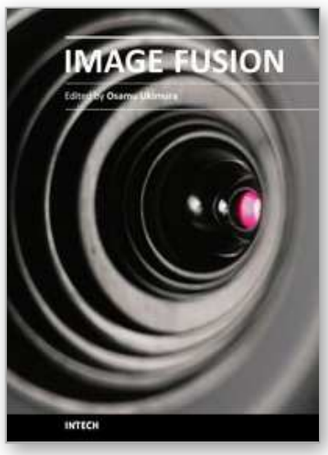

Image Fusion
Edited by Osamu Ukimura

ISBN 978-953-307-679-9
Hard cover, 428 pages
Publisher InTech
Published online 12, January, 2011
Published in print edition January, 2011

Image fusion technology has successfully contributed to various fields such as medical diagnosis and navigation, surveillance systems, remote sensing, digital cameras, military applications, computer vision, etc. Image fusion aims to generate a fused single image which contains more precise reliable visualization of the objects than any source image of them. This book presents various recent advances in research and development in the field of image fusion. It has been created through the diligence and creativity of some of the most accomplished experts in various fields.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen