The Hermite transform as an efficient model for local image analysis: An application to medical image fusion

Boris Escalante-Ramírez *

Departamento de Procesamiento de Señales, Facultad de Ingeniería, Universidad Nacional Autónoma de México, Edif. Bernardo Quintana, Circuito exterior, Cd. Universitaria, México, D.F. 04510, Mexico

Abstract

The Hermite transform is introduced as an image representation model that can be used to tackle the problem of fusion in multimodal medical imagery. This model includes some important properties of human visual perception, such as local orientation analysis and the Gaussian derivative model of early vision. Local analysis is achieved by windowing the image with a Gaussian function, then a local expansion into orthogonal polynomials takes place at every window position. Expansion coefficients are called Hermite coefficients and it is shown that they can be directly obtained by convolving the image with Gaussian derivative filters, in agreement with psychophysical insights of human visual perception. A compact representation can be obtained by locally steering the Hermite coefficients towards the direction of local maximum energy. Image fusion is achieved by combining the steered Hermite coefficients of both source images with the method of verification of consistency. Fusion results are compared with a competitive wavelet-based technique, proving that the Hermite transform provides better reconstruction of relevant image structures.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Hermite transform; Gaussian derivatives; Image fusion; Medical images; Human visual perception; Local orientation analysis

1. Introduction

This paper introduces the Hermite transform as an efficient tool for medical image fusion. The Hermite transform is an image representation model that mimics some of the more important properties of human visual perception – local orientation analysis and the Gaussian derivative model of early vision. Although this paper only discusses multimodal image fusion, many applications can be developed with a direct–inverse Hermite transform scheme.

It is generally acknowledged that visual perception models must include two major processing stages: (1) initial measurements and (2) high level interpretation. Fleet and Jepson [1] pointed out that early measurement is a rich encoding of image structure in terms of generic properties from which structures that are more
complex are easily detected and analyzed. Such measurement processes should be image-independent and require no previous or concurrent interpretation. Unfortunately, it is not known what primitives are necessary and sufficient for interpreting – or even identifying – meaningful features. We know, however, that for image processing, linear operators that exhibit special kinds of symmetries related to translation, rotation and magnification are of particular interest. A family of generic neighborhood operators that fulfill these requirements consists of the so-called Gaussian derivatives [2]. These operators have long been used in computer vision for feature extraction [3,4] and are relevant in visual system modeling [5]. Formal integration of these operators is achieved with the Hermite transform, first introduced by Martens [6,7] and recently reformulated as a multiscale image representation model for local orientation analysis [8,9]. This transform can take many alternative forms that correspond to different ways of coding local orientations in the image.

Young showed that the Gaussian derivatives model measures receptive field data more accurately than do the Gabor functions [10]. As with receptive fields, both Gabor functions and Gaussian derivatives are spatially local and consist of alternating excitatory and inhibitory regions within a decaying envelope. Nonetheless, the Gaussian derivative analysis has proven more efficient because Gaussian derivatives comprise an orthogonal basis when they belong to the same point of analysis. Gaussian derivatives can be interpreted as local generic operators in a scale-space representation described by the isotropic diffusion equation [2]. In a related work, Gaussian derivatives are interpreted as the product of Hermite polynomials and a Gaussian window [6], where windowed images are decomposed into a set of Hermite polynomials. Some mathematical models based on these operators at a single spatial scale have been described elsewhere [6,11]. The Hermite transform has been extended to the multiscale case [7–9] and has been successfully used in different applications such as noise reduction [12], coding [13], motion estimation for image sequences [15], and image fusion in remote sensing [15,16].

Local orientation analysis is a topic of major concern in this paper. It is well known that local orientation can be estimated by combining the outputs from polar separable quadrature filters [17]. Freeman and Adelson developed a technique to steer filters by linearly combining basis filters oriented at a number of specific directions [18]. The possibilities are, in fact, infinite since the set of basis functions required to steer a function is not unique [19]. The Gaussian derivative family is perhaps the most common example of such functions.

The first part of this paper introduces the Hermite transform as an image representation model and shows how to undertake local analysis with a steered Hermite transform. The second part shows how to use the Hermite transform to achieve multimodal medical image fusion, namely CT with PET, and PET with MRI. It also shows how the proposed method preserves the spatial resolution and image contents of both images.

The fusion algorithm is based on the detection of relevant image structures (primitives) during the analysis stage. To that end, multiscale Gaussian-derivative filters can be used. Local orientation is estimated in order to rotate the transform at every position of the analysis window. Transform coefficients from different image sources are classified in order to select from each image coefficients that contribute to synthesizing the fused image.

2. The Hermite transform

2.1. The Hermite transform as an image representation model

The Hermite transform [6,7] is a special case of polynomial transform, which can be regarded as an image description model. Windowing with a local function $\omega(x,y)$ takes place at several positions over the input image and local information at every analysis window is expanded in terms of a family of orthogonal polynomials. The polynomials $G_{m,n}(x,y)$ used to approximate the windowed information are determined by the analysis window function and satisfy the orthogonal condition

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

for $n, k = 0, \cdots, \infty; m = 0, \cdots, n$ and $l = 0, \cdots, k$; where $\delta_{nk}$ denotes the Kronecker function.

Psychophysical insights suggest using a Gaussian window function, which resembles the receptive field profiles of human vision, i.e.,...
\[ \omega(x,y) = \frac{1}{2\pi\sigma^2} \exp\left( -\frac{(x^2 + y^2)}{2\sigma^2} \right) \]  

(2)

The Gaussian window is separable into Cartesian coordinates. It is isotropic – hence rotationally invariant – and its derivatives are good models of some of the more important retinal and cortical cells of the human visual system [5,6].

In a Gaussian window function, the associated orthogonal polynomials are the Hermite polynomials [20]

\[ G_{n,m}(x,y) = \frac{1}{\sqrt{2^n(n-m)!m!}} H_{n-m}(\frac{x}{\sigma}) H_m(\frac{y}{\sigma}) \]  

(3)

where \( H_n(x) \) denotes the \( n \)th 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 \( W(x,y) = \sum_{(p,q)\in S} \omega(x-p,y-q) \). This weighting function must be a number other than zero for all coordinates \((x,y)\). Therefore,

\[ L(x,y) = \frac{1}{W(x,y)} \sum_{(p,q)\in S} L(x,y) \omega(x-p,y-q) \]  

(4)

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. 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 by convolution of the original image \( L(x,y) \) with the function filter \( D_{m,n-m}(x,y) = G_{m,n-m}(-x,-y)\omega^2(-x,-y) \), followed by subsampling at a position \((p,q)\) of the sampling lattice \( S \). That is,

\[ L_{m,n-m}(p,q) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} L(x,y)D_{m,n-m}(x-p,y-q)dx\,dy \]  

(5)

For the Hermite transform, the filter functions \( D_{m,n-m}(x,y) \) can be shown to correspond [20] to Gaussian derivatives of order \( m \) in \( x \) and \( n-m \) in \( y \), in agreement with the Gaussian derivative model of early vision [5,6]. The process of recovering the original image consists of interpolating the transform coefficients with the proper synthesis filters. This process is called the inverse polynomial transform, which is defined by

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

(6)

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

\[ P_{m,n-m}(x,y) = \frac{G_{m,n-m}(x,y)\omega(x,y)}{W(x,y)} \quad \text{for} \quad m = 0, \ldots, n \quad \text{and} \quad n = 0, \ldots, \infty \]

Fig. 1 shows the analysis and synthesis stages of a polynomial transform. Fig. 2 shows a Hermite transform calculated on a medical image.

To define a polynomial transform, we have to select some parameters. First, we have to define the characteristics of the window function. From a perceptual point of view and according to the scale-space theory, the Gaussian window is the best option. Other free parameters are the size of the Gaussian window spread (\( \sigma \)) and the distance between adjacent window positions (subsampling factor) which defines the sampling lattice. The size of the window functions must be related to the spatial scale of the image structures to be analyzed. Fine local changes are better detected with small windows, but representation of low-resolution objects requires...
large windows. Therefore, multiresolution representations can help overcome this compromise. In the case of the Hermite transform, a multiresolution extension has been recently proposed [8,9].

2.2. The steered Hermite transform

High-energy compaction can be attained by adaptively steering the Hermite transform [21], which represents an advantage. 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 $N$th-order form a steerable basis for each individual $N$th-order filter. Because of this property, the Hermite filters at each position in the image adapt to the local orientation content. This adaptability results in significant compaction.

For orientation analysis purposes, it is convenient to work with a rotated version of the Hermite transform. 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 in polar coordinates. If $\omega_x = \omega \cos \theta$ and $\omega_y = \omega \sin \theta$, then

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

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

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

and the orientation selectivity of the filter is expressed by

$$
g_{m,n-m}(\theta) = \sqrt{n \choose m} \cos^m \theta \cdot \sin^{n-m} \theta
$$

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

$$
g_{m,n-m}(\theta - \theta_0) = \sum_{k=0}^{n} c_{m,k}^{(n)}(\theta_0) g_{n-k,k}(\theta)
$$

where $c_{m,k}^{(n)}(\theta_0)$ is the steering coefficient. The Hermite filter rotation at each position over the image is an adaptation to local orientation content. Fig. 3 shows the directional Hermite decomposition over an image. First, a Hermite transform was applied and then the coefficients of this transform were rotated toward the estimated local orientation, according to a criterion of maximum oriented energy at each window position. For local 1D patterns, the steered Hermite transform provides a very efficient representation. This representation consists of a parameter $\theta$ that indicates the orientation of the pattern and a small number of coefficients that represent the...
For a 1D pattern with orientation $\theta$, the following relation holds:

![Diagram showing coefficient orders. Diagonals depict zeroth-order coefficients ($n=0$), first-order coefficients ($n=1$), and so forth. A Gaussian window with spread $\sigma = 1$ and subsampling $T = 2$ was used.]

Fig. 2. (a) Hermite transform calculated on a CT image. (b) Diagram showing coefficient orders. Diagonals depict zeroth-order coefficients ($n=0$), first-order coefficients ($n=1$), and so forth. A Gaussian window with spread $\sigma = 1$ and subsampling $T = 2$ was used.
Fig. 3. Steered Hermite Transform. (a) Original coefficients. (b) Steered coefficients. Interestingly, most coefficient energy is concentrated in the upper row. Coefficient at position 01 now represents the local angle rotation, i.e. \( \tan(\theta) = \frac{L_{0,1}}{L_{1,0}} \), where \( L_{0,1} \) and \( L_{1,0} \) represent the original first-order coefficients.
\[ L_{n-m,m}^0 = \begin{cases} \sum_{k=0}^{n} g_{n-k,k}(\theta)L_{n-k,k}, & m = 0 \\ 0, & m > 0 \end{cases} \] (11)

For such pattern, steering over \( \theta \) results in a compaction of energy into the coefficients \( L_{n,0}^0 \), while all other coefficients are set to zero.

Using Hermite coefficients, the energy content can be expressed as (Parseval Theorem)

\[ E_\infty = \sum_{n=0}^{\infty} \sum_{m=0}^{n} |L_{n-m,m}|^2 \] (12)

Up to order \( N \), the energy \( E_N \) is defined as the addition of all squared coefficients up to order \( N \).

The steered Hermite transform offers a way to describe 1D patterns on the basis of their orientation and profile. We can differentiate 1D energy terms and 2D energy terms. That is, for each local signal

\[ E_{1D}^N(\theta) = \sum_{n=1}^{N} |L_{n,0}^\theta|^2 \] (13)

\[ E_{2D}^N(\theta) = \sum_{n=1}^{N} \sum_{m=1}^{n} |L_{n-m,m}^\theta|^2 \] (14)

### 3. Image fusion with the Hermite transform

Image fusion has become a useful tool for enhancing information from two or more sensors by combining the most relevant features of each image. A wide range of disciplines – including medicine and remote sensing – has taken advantage of fusion techniques. In recent years, these techniques have evolved from simple linear combinations to sophisticated methods based on principal components, color models and signal transformations, among others [22–27]. Further, multiresolution techniques, such as image pyramids and wavelet transforms, have been successfully used [25,28–30]. Several authors have shown that, for image fusion, the wavelet transform approach offers good results [20,21,29,30]. Comparisons of Mallat’s and ‘à trous’ methodologies have been carried out [31].

In the fused image, information must yield improved accuracy (based on redundant information) and improved capacity (based on complementary information). Moreover, from a visual perception point of view, patterns included in the fused image must be perceptually relevant and must not include distracting artifacts. Our approach aims at analyzing images by means of the Hermite transform, which allows us to identify perceptually relevant patterns to be included in the fusion process while discriminating spurious artifacts.

The steered Hermite transform has the advantage of energy compaction. Transform coefficients are selected with an energy compaction criterion from the steered Hermite transform. Therefore, it is possible to reconstruct an image with few coefficients and still preserve details such as edges and textures.

![Fig. 4. Fusion scheme with the Hermite transform.](image-url)
The general framework for fusion by means of the Hermite transform includes five steps. (1) The Hermite transform of the image. (2) Detection of maximum energy orientation with the energy measure $E_N^{D_d}(\theta)$ at each window position. In practice, an estimator of the optimal orientation $\theta$ can be obtained through $\tan(\theta) = L_{0,1}/L_{1,0}$, where $L_{0,1}$ and $L_{1,0}$ are first-order Hermite transform coefficients. (3) Adaptive steering of the transform coefficients, as described in previous sections. (4) Selection of coefficients, based on the method of verification of consistency [28]. This selection rule uses the maximum absolute value in a $5 \times 5$ window over the image (area of activity). The window variance is computed and used as a measurement of the activity associated with the central pixel of the window. Thus, a significant value indicates the presence of a dominant pattern in the local area. A map of binary decision is then created for the registry of results. This binary map is to undergo consistency verification. Lastly, (5) the inverse transform of selected coefficients and their corresponding optimal $\theta$ [13]. Fig. 4 shows a simplified representation of this method.

4. Results

The proposed fusion scheme was tested with CT, MRI, and PET data. Fig. 5 shows the original data used to test the image fusion algorithm.

Fig. 5. Upper left, CT image; upper right, registered MRI image. Lower left, MRI image; lower right, registered PET image.
We fused CT and MRI images, and PET and MRI images. Figs. 6 and 7 show the results of image fusion achieved with the proposed method, and the results obtained using a wavelet transform (Daubechies Spline).

According to Figs. 6 and 7, the image fusion method based on the Hermite transform preserved the spatial resolution and information content of both images. At first glance, the results obtained with the wavelet transform were very similar to those obtained with the Hermite transform. Nonetheless, upon closer inspection, the Hermite transform proved to be more efficient at reconstructing basic patterns, such as edges. Figs. 8 and 9 show details of the results presented in Figs. 6 and 7. Differences are found mostly along the edges of relevant structures. The fact that the Hermite transform performed better at reconstructing edges was not surprising: Gaussian derivatives – the basis functions of this transform – have long been acknowledged to optimally detect, represent and reconstruct perceptually relevant image patterns, such as edges and lines [3–5,10,32].

Fig. 6. Fused CT–MRI images. Left, fusion with the Hermite Transform (using a Gaussian window with spread $\sigma = \sqrt{2}$ and subsampling factor $T = 4$). Right, fusion with a wavelet transform (Daubechies Spline).

Fig. 7. Fused PET–MRI images. Left, fusion with the Hermite Transform (using a Gaussian window with spread $\sigma = \sqrt{2}$ and subsampling factor $T = 4$), Right, fusion with a wavelet transform (Daubechies Spline).
5. Conclusions

The Hermite transform proved to be an efficient image representation model for image fusion in multimodal medical imagery. Other applications such as noise reduction, coding, and motion estimation have been discussed in related works [12–14].

Gaussian derivatives, known to be good operators for the detection of perceptually relevant image primitives, are the basis functions of the Hermite transform. Local orientation analysis is another important feature of the proposed fusion algorithm. All these characteristics are common to some of the more important properties of the early stages of human vision and might help explain why fusion with the Hermite transform achieves better reconstruction of important image features as compared with wavelets, whose orthogonality requirements limit the possibility of having basis functions with a perceptually relevant profile.
The algorithm presented here is formulated in accordance with a single spatial scale scheme, i.e. with a fixed Gaussian window of analysis. Multiresolution, however, is also an important characteristic of human vision that has proved to be an efficient way to construct image-processing solutions. Multiresolution image processing algorithms are straightforward to build from the Hermite transform by means of hierarchical pyramidal structures that replicate, at each resolution level, the analysis-synthesis image processing schemes proposed here. Moreover, a formal approach to the multiresolution Hermite transform for local orientation analysis has been recently developed, clearing the way to propose new multiresolution image processing tasks [8,9].

Acknowledgements

This work was sponsored by UNAM Grant PAPIIT IN105505. The author is indebted to Juan Carlos Muñoz Torres for carrying out the experiments presented in this paper.

References

[28] Li H, Manjunath BS, Mitra SK. Multisensor image fusion using the wavelet transform. Graph Model Im Process 1995;57(3):235–45.


Boris Escalante-Ramírez was born in Mexico City in 1961. He received the Bachelors of Electrical Engineering degree from the Universidad Nacional Autónoma de México in 1985, the Master of Electronic Engineering degree from the Philips International Institute of Technological Studies, Eindhoven, The Netherlands, in 1987, and a PhD degree from the Eindhoven University of Technology in 1992. Since then he has been with the Facultad de Ingeniería, Universidad Nacional Autónoma de México. His research interests are visual information processing, perception models and their applications to image processing including restoration, coding, segmentation, fusion and optic flow.