

Motion Estimation of Magnetic Resonance Cardiac Images Using the Wigner–Ville and Hough Transforms¹

N. Carranza^a, G. Cristóbal^a, P. Bayerl^b, and H. Neumann^b

^a Instituto de Óptica Daza de Valdés (CSIC), Serrano 121, 28006 Madrid, Spain

^b University of Ulm, Dept. of Neural Information Proc., D-89069 Ulm, Germany

e-mail: noemi@optica.csic.es, gabriel@optica.csic.es, Pierre.bayerl@uni-ulm.de, heiko.neumann@uni-ulm.de

Received December 28, 2006

Abstract—Myocardial motion analysis and quantification is of utmost importance for analyzing contractile heart abnormalities and it can be a symptom of a coronary artery disease. A fundamental problem in processing sequences of images is the computation of the optical flow, which is an approximation of the real image motion. This paper presents a new algorithm for optical flow estimation based on a spatiotemporal-frequency (STF) approach. More specifically it relies on the computation of the Wigner–Ville distribution (WVD) and the Hough Transform (HT) of the motion sequences. The latter is a well-known line and shape detection method that is highly robust against incomplete data and noise. The rationale of using the HT in this context is that it provides a value of the displacement field from the STF representation. In addition, a probabilistic approach based on Gaussian mixtures has been implemented in order to improve the accuracy of the motion detection. Experimental results in the case of synthetic sequences are compared with an implementation of the variational technique for local and global motion estimation, where it is shown that the results are accurate and robust to noise degradations. Results obtained with real cardiac magnetic resonance images are presented.

PACS numbers: 87.61.-c

DOI: 10.1134/S0030400X07120077

1. INTRODUCTION

Motion estimation is a key problem for the analysis of image sequences. Visual images and their dynamic evolutions provide enormous amounts of information about our surrounding environment. By observing changes in a scene over time, 3D scene properties and motion parameters can be obtained. We can discover the 3D structure of the scene to make predictions about collisions and to infer material properties of objects, such as their stiffness and transparency; much of this information is revealed by the motion of the different parts of the scene.

Because of the richness of motion as an information source, analysis of visual motion is essential for many practical applications. These range from image-processing problems, such as efficient coding or enhancement of motion pictures, to passive machine vision problems, such as determining the shape of a moving object or recovering the motion of the camera relative to the scene, to active perception applications in which an autonomous agent must explore its environment [1]. It is also useful for performing motion segmentation, for computing the focus of expansion and time-to-collision, for performing motion-compensation image encoding, and for measuring blood flow and heart-wall motion in medical imagery.

The analysis of heart-wall deformation has important clinical implications for the assessment of viability of the heart wall and provides quantitative estimates of the location and extent of ischemic myocardial injury. Quantifying the extent of regional wall motion abnormality can aid in determining the myocardial effects of coronary artery diseases. Among all of the available imaging methods, cardiac magnetic resonance imaging (CMR) is currently recognized as the best imaging method for the dynamic exploration of the cardiac function and it is used not only for the scientific purpose of understanding heart motion, but also for the clinical need to diagnose heart disease. However, the estimation of optical flow is a challenging problem for this particular case of imagery due to a wide range of possible motions and the presence of noise. In addition to this, the nonrigid motion of the heart makes cardiac-motion estimation a complex problem.

In this paper, an algorithm for computing the optical flow applied to CMR based on the computation of the Wigner–Ville distribution and the Hough Transform is presented.

The rationale of this approach is based on the fact that in many situations, motion analysis can be interpreted as result of disparity information. Therefore, correlation-based spectral methods such as the provided by the PWD will constitute a valuable tool. The paper is organized as follows. Section 2 briefly reviews

¹ The text was submitted by the authors in English.

the problem of computing optical flow and explains the frequency-based methods. Section 3 provides an explanation of the new algorithm proposed. Section 4 evaluates the method with synthetic and real images and compares it with a variational approach and Section 5 concludes the paper.

2. OPTICAL FLOW ESTIMATION

From the information available from a sequence of images, it is only possible to derive an estimate of the motion field called "optical flow." Although optical flow is generally not equivalent to the true motion field, it is quite similar in most of the cases.

Numerous theoretical and practical studies have been performed on optical flow estimation from image sequences and on the useful information it contains. Despite this wide variety of approaches, algorithms for computing optical flow are usually divided into the three following categories [1].

Differential techniques: Also known as "gradient" techniques, these methods estimate optical flow vectors from the derivatives of image intensity over space and time. They are typically derived directly by considering the total temporal derivative of an unchanged quantity, such as brightness.

Matching techniques. These methods operate by matching small regions of image intensity or specific "features" from one frame to the next. The matching criterion is usually the least squared or normalized correlation measure.

Frequency-based or filter-based techniques. These methods are based on spatio-temporally oriented filters (i.e., velocity-sensitive), and they are typically motivated and analyzed by considering the motion problem in the Fourier domain. They fall into two categories, energy-based and phase-based.

Each technique has several advantages and disadvantages. Numerical differentiation is sometimes impractical because of small temporal support or poor signal-to-noise ratio. In these cases, it is natural to consider correlation-based techniques. Among the advantages brought by the frequency-based methods, it is found that motion-sensitive mechanisms operating on spatiotemporally oriented energy in Fourier space can estimate motion in images for which matching approaches would fail [2]. One example can be the motion of random dot patterns.

The differential method has a major drawback in the estimation of the first and second derivatives of the pixel intensity, mainly in the case of noisy images. To improve noise robustness, a common strategy is to use regularization methods based on variational integrals. On the contrary, frequency-based methods are in general more robust to noise [3]. Due to the fact that cardiac images are quite noisy, a frequency-based approach has been chosen for the implementation of our algorithm.

2.1. Optical Flow

The initial hypothesis for measuring image motion is that the intensity structures of local time-varying image regions are approximately constant under motion for at least a short duration [4]. That is, changes in the image intensity are due only to translation of the local image intensity and not due to changes in lighting, reflectance, etc. According to this assumption, the total derivative with respect to time of the image intensity function should be zero at each position in the image and at every time.

Let $i(\bar{x}, t)$ denote the intensity function, where $\bar{x} = (x, y)$ represents the pixel position and t is the time. If the intensity remains constant, then

$$i(\bar{x}, t) = i(\bar{x} + \delta\bar{x}, t + \delta t), \quad (1)$$

where $\delta\bar{x}$ is the displacement of the local image region at (\bar{x}, t) after time δt . Expanding the left-hand side of this equation in a Taylor series yields

$$\nabla i \bar{v} + i_t = 0, \quad (2)$$

where $\nabla i = (i_x, i_y)$ is the spatial intensity gradient, i_t is the derivative of the intensity with respect to time, and $\bar{v} = (v_x, v_y)$ is the image velocity.

This equation is called the optical flow constraint equation (*OFCE*). Unfortunately, one scalar equation is not enough for finding both components of the velocity field; it gives only the component in the direction of the gradient, that is, the normal flow. This problem is usually called the aperture problem [5].

2.2. Frequency Domain Description

Many important insights about the estimation of motion may be explained in a straightforward manner by considering the problem in the Fourier domain. This can be easily appreciated by considering the motion of a one-dimensional signal. We can represent such a signal as an intensity image in which the intensity of each pixel corresponds to the value of the signal at a particular location and time. A translating one-dimensional signal has the appearance of a striped pattern, where the stripes are oriented at an angle of $\alpha = \arctan(1/v)$, where v is the velocity of the signal [1]. Clearly, the Fourier decomposition of this signal is a set of sinusoids of this same orientation, and varying wave number (spatial frequency magnitude). Thus, the power spectrum of the Fourier transform will be located along a line that passes through the origin at angle α , as illustrated in Figs. 1a and 1b.

The problem in two dimensions is analogous, although more difficult to visualize. A translating two-dimensional pattern has the appearance of oriented "bundles of fibers" in a 3D space-time (x, y, t) . The Fourier transform spectrum of an image undergoing rigid translation lies in a plane in the spatiotemporal frequency domain.

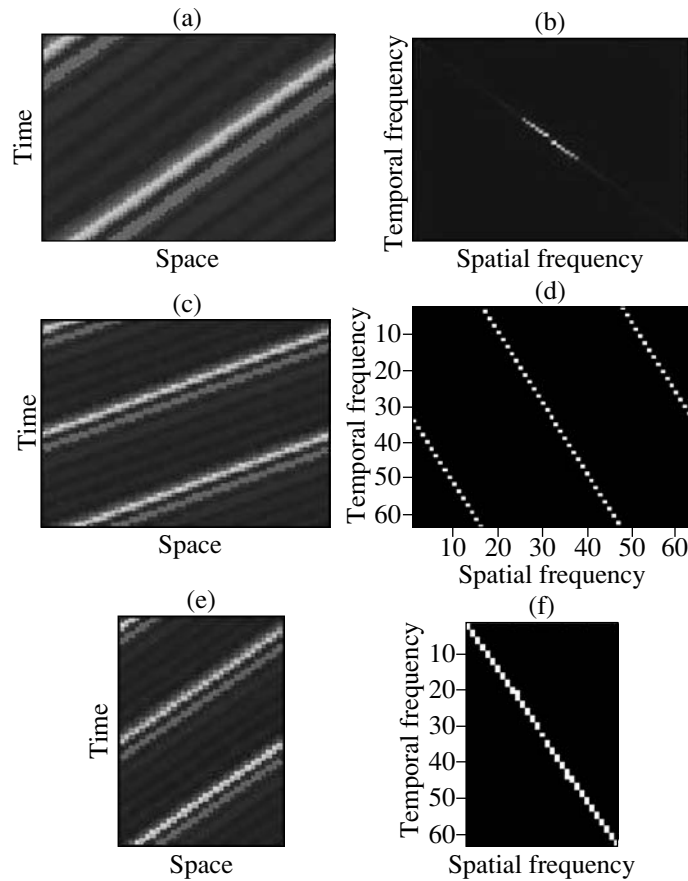


Fig. 1. (a) Sinc function translating ($v = 1$ pixel/s). (b) Fourier transform of the signal (a). (c) Sinc function translating ($v = 0.5$ pixel/s). (d) Spectrum of the signal (c). Note the presence of aliasing. (e) Signal downsampled in space. (f) Spectrum of the signal (e).

For analyzing a video sequence through a three-dimensional Fourier transform, let us assume again that we can represent an image sequence through a function $i_0(x, y)$, such as

$$i(x, y, t) = i_0(x - v_x t, y - v_y t). \quad (3)$$

The main assumption here is that moving objects must displace with a uniform velocity vector (v_x, v_y) and must have constant illumination. Now, by calculating the spatial and temporal Fourier transform of the sequence $i(x, y, t)$, we obtain

$$I(f_x, f_y, f_t) = I_0(f_x, f_y) \delta(v_x f_x + v_y f_y + f_t), \quad (4)$$

where I_0 represents spatial Fourier transform of i_0 and δ is the Dirac delta function.

Thus, $I(f_x, f_y, f_t)$ is nonzero only on a plane, which is called the motion plane [6]. This plane passes through the frequency origin. Its equation is given by

$$v_x f_x + v_y f_y + f_t = 0. \quad (5)$$

Estimating parameters of this plane leads to estimate the velocity vector components (v_x, v_y) of the moving

object. Therefore, Eq. (5) can also be viewed as the Fourier transform of the OFCE

$$\frac{\partial i(x, y, t)}{\partial t} + v_x \frac{\partial i(x, y, t)}{\partial x} + v_y \frac{\partial i(x, y, t)}{\partial y} = 0. \quad (6)$$

2.3. STF Approach

Among the different techniques for computing optical flow using frequency-based methods, the spatiotemporal-frequency approach (STF) has been proposed, which gives a simultaneous representation of a signal in space and spatial frequency variables [7]. One implementation of the STF approach employs the Wigner–Ville distribution (WVD) as the underlying STF image representation.

The major motivation for considering the use of STF image representation approach as a basis for computing optical flow comes from the literature on mammalian vision. In particular, some investigations have demonstrated that many neurons in various cortical areas of the brain behave as spatiotemporal-frequency bandpass filters [8]. In the field of nonstationary signal analysis, the WVD has been used for the representation of speech and image. Jacobson and Wechsler [7, 9] were

the first to suggest the use of the WVD for the optical flow estimation.

The Wigner Distribution was introduced by Wigner as a phase space representation in Quantum Mechanics, and it gives a simultaneous representation of a signal in space and spatial frequency variables [9]. Later, in the area of signal processing, Ville derived the same distribution that Wigner proposed several years before [10]. The WVD can be considered as a particular occurrence of a complex spectrogram in which the shifting window function is the function itself [11].

The WVD of a moving sequence is a 6-D function defined by

$$W_i(x, y, t, w_x, w_y, w_t) = \iiint R_i(x, y, t, \alpha, \beta, \tau) \times \exp[-j(\alpha w_x + \beta w_y + \tau w_t)] d\alpha d\beta d\tau, \quad (7)$$

where

$$R_i(x, y, t, \alpha, \beta, \tau) = i(x + \alpha, y + \beta, t + \tau) i^*(x - \alpha, y - \beta, t - \tau), \quad (8)$$

and where * denotes complex conjugation.

Again, for the case where a time-varying image $i(x, y, t)$ is uniformly translating at some constant velocity, the WVD of this image is

$$W_i(x, y, t, w_x, w_y, w_t) = \delta(v_x w_x + v_y w_y + w_t) W_i(x - v_x t, y - v_y t, w_x, w_y). \quad (9)$$

From (9), the WVD of a linearly translating image with velocity (v_x, v_y) is everywhere zero except in the plane defined by

$$\{(x, y, t, w_x, w_y, w_t): v_x w_x + v_y w_y + w_t = 0\}. \quad (10)$$

Equivalently, for an arbitrary pixel at x, y, t , each local STF spectrum of the WVD is zero everywhere except on the plane defined by (10). For this reason, if a procedure for estimating the velocity associated with a given STF spectrum is found, we will obtain a space and time varying optical flow function. In the next sections we will describe how the Hough Transform can be used for computing the optical flow together with the WVD.

3. IMPLEMENTATION ISSUES

In Section 2 details about the frequency description on motion estimation, specifically on a spatiotemporal approach, have been presented. In this section, this background is used to make a description about the method proposed. As mentioned before, it is necessary to choose a procedure for estimating the slope of the plane found in the spectrum. In our algorithm, the technique proposed is the *HT* due to the ability to discard

cross-terms that can be introduced by the WVD. In the next section, a short explanation of the *HT* is made, followed by the description of the algorithm proposed.

3.1. Hough Transform

The problem of determining the location and orientation of straight lines in images is of great importance in the fields of computer vision and image processing. One approach used for the detection of lines is the Hough Transform (*HT*), a nonlinear filtering technique to estimate the position and direction of certain curves in a discrete image [12]. Another pose estimation method commonly used is the least squares (*LS*) algorithm, which is based on the minimization of an objective function: the sum of squared distances between image features and the model. The *HT* is the classical approach for finding the parameters of lines in a binary image, and maps each image point to all points in the parameter space, which could have possibly produced the image point. Thus, each image point votes for the shape parameters that could have produced it. The points in the parameter space that accumulate the greatest number of votes, which appear as peaks, are the most likely to have produced the shapes in the true image. Therefore, the *HT* reduces the problem of detecting spatially spread patterns in the image space to that of finding localized peaks in a dual parameter space [13]. Its main advantages are an ability to discard features belonging to other objects and robustness against incomplete data and noise [14]. These characteristics are very important in our problem due to the presence of cross terms of the Wigner–Ville distribution; thus, we can use a more efficient method for determining the motion plane.

The parameterization specifies a straight line by the angle θ of its normal and its perpendicular distance ρ from the origin. If we restrict θ to the interval $[0, \pi]$, then the normal parameters for a line are unique. With this restriction, every line in the x - y plane corresponds to a unique point in the θ - ρ plane [15].

3.2. Algorithm Description

As seen in the previous section, a line can be completely characterized in the Hough plane, as well as a line and also a plane. This feature has been used to determine the velocity by means of the *HT* applied to the STF spectrum.

Our first approach was based on the use of the *HT* on the whole spectrum in order to find the plane. In this way, each pixel of the spectrum with a nonzero value was represented in the Hough plane. However, pixels of the WVD that belong to cross terms influence the final result, sometimes producing incorrect solutions. For this reason, we have used another approach based on the *HT* computation for each of the frames of the spectrum in order to detect a line on each of them and in this

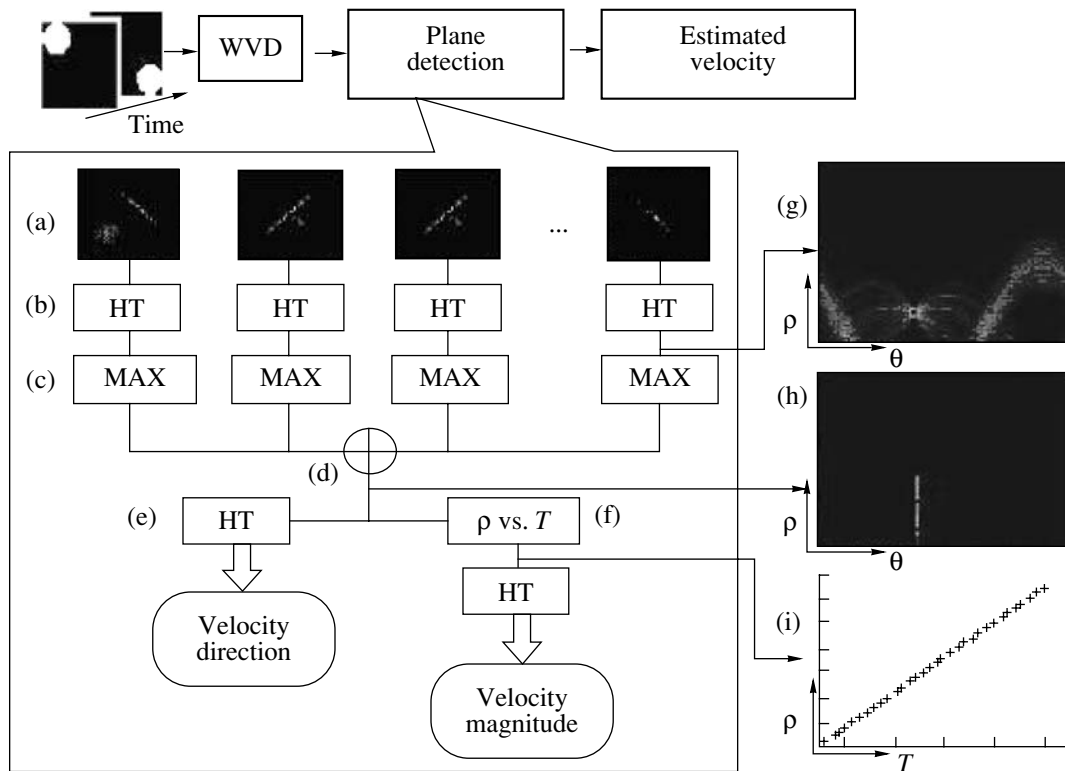


Fig. 2. General block diagram of the method proposed.

way discarding the information from cross-terms pixels. Furthermore, this implementation is computationally less demanding.

In the end, our problem can be reduced to find one straight line in each temporal frame. Furthermore, we already know that the spectrum contains a frequency origin and we can observe that the lines in different temporal frames are parallel. By taking into account these facts and by applying the HT, the plane will be detected.

The main scheme of the implementation of the algorithm for motion estimation is represented in Fig. 2. We will illustrate our algorithm with an example. In Fig. 2, the block diagram of the implementation of the plane detection stage is shown. We started with a simplistic testing sequence, composed of a circular object moving with an oblique velocity along the coordinates X and Y . By performing the WVD, we will find a plane that is represented in each temporal-frequency frame by a line. In Fig. 2a, several frames of the spectrum are shown.

The next step of the algorithm is performing the HT for each of the frames of the spectrum (Fig. 2b). One example of the result obtained after this stage is shown in Fig. 2g. Taking the maximum value of this HT (Fig. 2c), the information of the position for each line is provided. As every straight line found in one frame is

parallel to the others, the maxima found in all of the HT lie all along a line, which means that every line of the spectrum has the same angle θ . Summing up all the HT transform of the frames of the spectrum (Fig. 2d), it is straightforward to find that all these maxima belong to one line (see Fig. 2h).

Actually, the information provided by the angle of one of the peaks would be sufficient to estimate the direction of the velocity. Most of the time, however, ideal conditions are not met (for example, the presence of cross terms induced by the WVD) and, in some cases, as a result of considering only one of the frames, we can end up with an erroneous solution due to noise or other external factors. In order to estimate the direction of the velocity, we propose to use the redundant information of all the frames and the property shown in the Fig. 2h, where all the maxima form a straight line and by applying the HT to the summation of all the peaks (Fig. 2e), erroneous peaks can be easily discarded.

In order to estimate the magnitude of the velocity, the values of the different ρ obtained have been used (i.e., the distance from the lines to the origin), so as to estimate the slope of the plane. As Fig. 2i shows, the values of ρ lie along a line, whose slope can be measured by means of another HT (Fig. 2f).

Table 1. Translations in pixels/frame for several examples

Actual translation		Estimated translation	
V_x	V_y	v_x	v_y
1	0	1.0000	-9.4941e-016
1	-1	0.9717	-0.9748
-0.7	0.5	-0.6773	0.5203
0.7	-1	0.6935	-0.9434
-1.2	-0.7	-1.1973	-0.6790

3.3. Implementation of the Algorithm for Local Motion Estimation

This implementation should be used when the a priori information of assuming only one object in the sequence is unknown. Thus, a small window is assigned to each pixel of the sequence, and the algorithm presented in the previous section is executed for each of the windows. However, using only one fixed size of window can lead to several problems, such as aliasing effects. For this reason, the technique above described has been modified in a hierarchical coarse-to-fine framework.

A hierarchical scheme allows the images to be decomposed in different scales of resolution in the form of Gaussian or Laplacian pyramids. Because of a low-frequency representation at coarser resolutions, the optical flow constraint equation becomes applicable in the case of large image motions. In addition to handling fast motions, hierarchical processing also offers increased computational efficiency [2]. With only one scale of resolution, because of low sampling rates and aliasing effects, the OFCE becomes inappropriate.

In the case of our implementation, the problems which have been solved with the hierarchical implementation are the aliasing effect, the aperture effect and the problem of measuring large or too small image motions. Considering again the example shown in Fig. 1, let us suppose a decrease in the velocity to 0.5 pixels per second (see Fig. 1c). Its spectrum is shown in Fig. 1d), where the aliasing is clearly visible. With a coarse resolution in space, taking only one pixel

out of two, the velocity is increased a factor of two (Fig. 1e) and, consequently, the aliasing is eliminated (Fig. 1f). With this hierarchical implementation, the most accurate velocity is chosen after having obtained the results for all the scales of resolution. The decision about the velocity measurement is made taking into account the accuracy of the HT.

4. RESULTS

4.1. Results with Synthetic Sequences

The new methodology proposed here was applied for evaluation purposes to synthetic images of a moving circular object with constant intensity, which can be described as follows:

$$\begin{cases} x(t) = \rho \cos(\theta) + x_0 + t v_x \\ y(t) = \rho \sin(\theta) + y_0 + t v_y, \end{cases} \quad (11)$$

where ρ represents the radius of the circle with a constant gray level, θ varies from 0 to 2π , $[x_0, y_0]$ is the initial point, t corresponds to the variation of the position with the time and finally, v_x and v_y are the velocity components of the circle.

In order to estimate the global motion, we have obtained a smoothed 3D frequency spectrum by means of a Hanning filter, which was introduced previously [3].

The method was applied to distinct values of ρ , $[x_0, y_0]$, and $[v_x, v_y]$. Some results are shown in Table 1. For these simple sequences, when we consider a moving object with a uniform velocity and we calculate a global motion, an accurate information about the optical flow can be obtained by means of the method based on WVD-HT.

A further step on the analysis has been completed, estimating the motion locally by means of the hierarchical implementation explained above. An example with two hierarchical levels will be discussed later in Fig. 3a, where the image size is $128 \times 128 \times 25$ pixels and window sizes are 25 and 50 pixels. For these sizes of window, we can observe that the optical flow estimated in the regions near to the border of the circle is

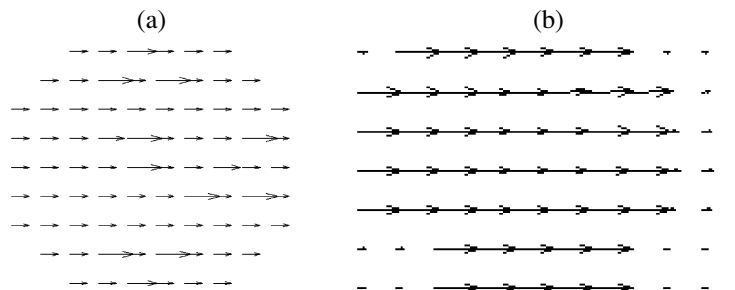


Fig. 3. (a) Optical flow with local estimation using the method based on WVD-HT; (b) optical flow obtained using variational methods.

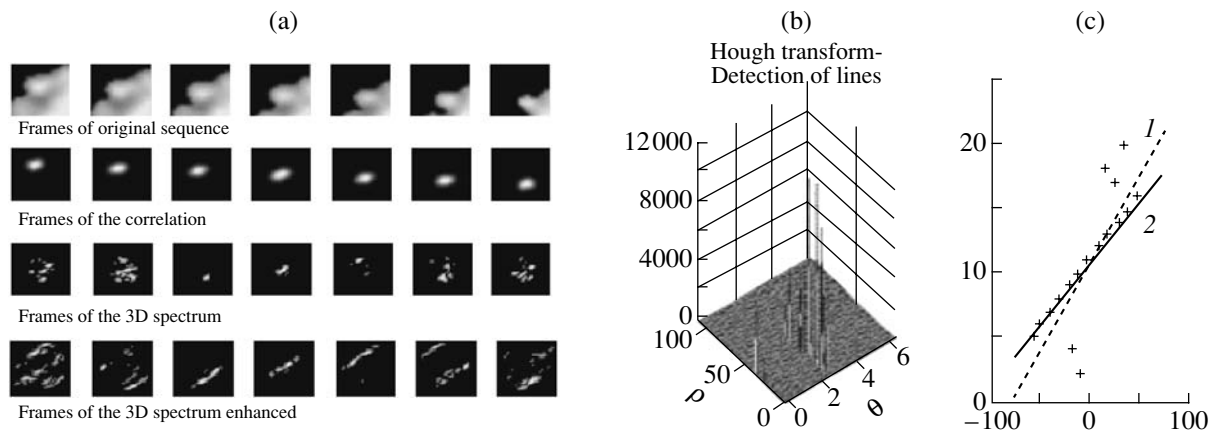


Fig. 4. (a) First row: Frames of a cropped region of a cardiac magnetic resonance; second row: correlation of the sequence; third row: spectrum obtained; fourth row: spectrum after processing the correlation. (b) Summation of all HT maxima. (c) Values of ρ and estimation by *LS* (dashed line) and *HT* (solid line).

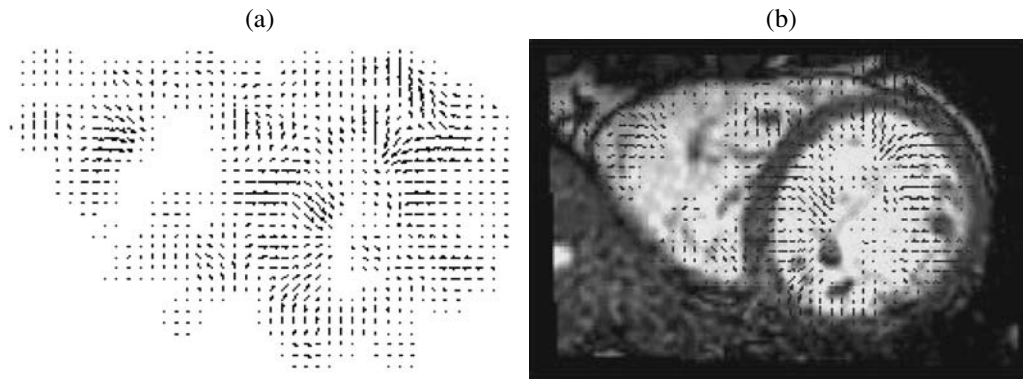


Fig. 5. (a) Optical flow obtained during a period of systole. (b) Same optical flow superimposed over the first frame of the sequence.

very accurate (compare with the results obtained with the variational method in Fig. 3b). Only regions inside the circle provide uncertainty due to the aperture problem and, therefore, values for optical flow are less accurate.

We have applied the variational method to the same simple sequence used before. We can observe the results for an arbitrary frame obtained for $v_x = 1$ and $v_y = 0$ in the Fig. 3b, where it is shown the optical flow obtained after regularization. We can see that the values of optical flow are about 1 pixel per frame, but the optical flow doesn't always fit the right positions. The uncertainty of the optical flow obtained in regions inside the circle, which is due to the aperture problem, has been partially corrected by means of the regularization.

Differential methods provide an optical flow for each pixel of the sequence, so it can be a tough task to perform a direct comparison with the global motion

estimation method based on *WVD-HT*, but we can make a comparison through the local motion estimation. Figures 3a and 3b show that results obtained with our method are comparable with the variational technique and the motion field obtained with our method fits better the right positions.

Finally, to test the robustness against noise of the new algorithm proposed, several experiments have been conducted by adding different types of noise to the sequence. Table 2 shows results of estimated translation for our sequence with added noise, when the actual translation is $V_x = 1$ pixel/frame and $V_y = -1$ pixel/frame. Gaussian and Speckle added noises are characterized by its variance, while salt and pepper noise is characterized by its density.

In the case of the variational method, the added noise of each pixel influences the estimation of the first and second derivatives of the pixel intensity. The opti-

Table 2. Estimated translations provided by WVD-HT method in pixels/frame for actual translation $V_x = 1$ and $V_y = -1$ for different additive noise variances

Noise variance/density	Gaussian		Speckle		Salt and pepper	
	v_x	v_y	v_x	v_y	v_x	v_y
0.05	0.97	-0.97	0.97	-0.97	0.96	-0.98
0.1	0.98	-0.95	0.97	-0.97	0.96	-0.98
0.2	0.95	-0.95	0.98	-0.96	0.98	-0.97
0.4	1	0	0.97	-0.98	0	-1

cal flow obtained obviously gets worse as the noise increases.

4.2. Motion Estimation for CMR Sequences

The method proposed has been applied to real CMR sequences in order to estimate myocardial deformation and its performance has been evaluated for these non-ideal cases. A local analysis of motion has been carried out, assigning a small window to each pixel of the sequence, and executing the algorithm for each of the windows. Thus, an optical flow for all the pixels of the sequence is obtained when the algorithm is executed for each of the pixels of the sequence.

In this case, the analysis is more complex because the ideal conditions are not met. The first condition is that objects must move under a uniform velocity. In a real sequence, a uniform velocity cannot be guaranteed. So the size of the sequence along the temporal axis must be small enough for assuming uniform movements. However, the number of images must be enough to obtain reliable information about the slope of the plane contained on the 3D Wigner–Ville distribution.

On the other hand, the main problem of these sequences is the presence of a deformable cardiac wall instead of a rigid moving object. Because of this reason, some cross terms are found in the spectrum and, in some cases, these terms can mask the plane detection. In order to trying to make uniform these small changes between the images of the sequence, a previous segmentation stage has been applied for extracting the main cardiac walls. Such segmentation turns to be difficult due to the low contrast of the frames and it has been attained by simply subtracting each frame with its corresponding negative frame and then by applying a simple image thresholding. However, a more reliable segmentation procedure should be developed in the future for improving the technique. Nevertheless, the problem of the presence of a deformable object will be the main one.

Figure 4a presents several frames of one of the windows of the original sequence (extracted from a cardiac magnetic resonance), where it is shown how the object doesn't preserve the form. The correlation of this sequence (Fig. 4b), therefore, is variable with time. For this reason, as is seen in Fig. 4c, the spectrum does not

have a clear plane detectable for the HT. This fact is due to the nonuniformity of this correlation. Thus, it is necessary to have a uniform correlation in order to obtain a better plane. A probabilistic approach based on Gaussian mixtures has been implemented. In this way, information about the position and shape of each correlation frame is extracted and processed so as to get a uniform correlation where the main information from the original one is preserved. Thus, the plane is detectable (Fig. 4d) and the accuracy of the motion detection is improved.

Figure 4e shows the results of the HT applied to the spectrum of Fig. 4d. The first figure represents the summation of all the maxima of the HT. As mentioned before, almost all of them belong to the same angle and, therefore, the direction of the movement is easily estimated. In the second figure, the values of ρ , information needed to evaluate the slope of the plane, are represented. Line 1 represents an approximation by least squares and line 2 corresponds to the HT result. As seen, the LS approach gives a worse result because it does not eliminate outliers.

Unfortunately, the ground truth is not available in these real sequences, so it is not possible to achieve a quantitative result. In the example shown in Fig. 4a, a velocity of $v_x = 0.38$ pixels per frame and $v_y = -0.5$ pixels per frame has been obtained (where the window size was $20 \times 20 \times 20$ pixels). Looking at the first and last image, it is possible to make a qualitative evaluation of the result, which confirms its feasibility.

The optical flow obtained for another sequence, taking the frames of a systolic period, is shown in Fig. 5. Although the ground truth is not provided as was mentioned before, a qualitative assessment of the direction of the optical flow obtained indicates that it properly fits the movement of the myocardium.

5. CONCLUSIONS

In this paper, a frequency-based method for the estimation of motion based on the computation of the Wigner–Ville distribution together with the Hough Transform has been presented. Results from synthetic sequences have been shown, evaluated, and compared with an implementation based on the variational

method. The proposed method has been applied also in real sequences of CMR.

The estimation of global motion yields values of optical flow very close to the actual ones. In this case, we can say that the solution is more accurate and robust to noise than the one obtained with the variational method. Nevertheless, a more fair comparison between both methods has been performed with a local estimation of motion. By selecting an appropriate window size of analysis and using a hierarchical approach, we showed that the proposed method gives similar results to the variational approach.

The advantages brought by the WVD-HT method are the noise robustness and the accuracy of the global estimation of motion. In both the WVD-HT and variational methods, one of the major drawbacks is the selection of the optimal parameters for the algorithms.

The algorithm has been applied to real cardiac magnetic resonance sequences. Motion estimation in these sequences is very important for a better diagnosis of cardiac diseases. Results obtained have been evaluated qualitatively. Further evaluation with clinical data should be addressed in the future for confirming current promising results.

ACKNOWLEDGMENTS

Financial support of this research was provided by the following research grants: TEC2004-00834; TEC2005-24739-E; TEC2005-24046-E; 20045OE184 and PI040765.

REFERENCES

1. E. Simoncelli, *Distributed Representation and Analysis of Visual Motion*, Thesis of Massachusetts Institute of Technology, 1993.
2. S. S. Beauchemin and J. L. Barron, *ACM Computing Surveys* **27**, 433 (1995).
3. W. I. Meyering, M. A. Gutierrez, S. S. Furuie, et al., *Computers in Cardiology 2000* (IEEE Press, Cambridge, 2000), pp. 619–622.
4. B. K. P. Horn and B. G. Schunck, *Artificial Intell.* **17**, 185 (1981).
5. G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations* (Springer-Verlag, New York, 2001).
6. M. Pingault, D. Pellerin, *Signal Processing* **84**, 709 (2004).
7. L. Jacobson and H. Wechsler, *Comput. Vis. Graph. Image Proc.* **38**, 29 (1987).
8. P. A. Laplante and A. D. Stoyenko, *Real-Time Imaging: Theory, Techniques, and Applications* (IEEE Press, 1996).
9. E. Wigner, *Phys. Rev.* **40**, 749 (1932).
10. E. Ville, *Cables et Transmission* **2A**, 61 (1948).
11. J. Hormigo and G. Cristobal, *Adv. Imag. Electron. Phys.* **131**, 65 (2003).
12. J. Bigun, *Vision with Direction: A Systematic Introduction to Image Processing and Computer Vision* (Springer-Verlag, Berlin, 2006).
13. N. Aggarwal and W. C. Karl, *IEEE Trans. Image Proc.* **15** (3), 582 (2006).
14. J. S. Marques, in *Proceedings of the 1998 European Signal Processing Conference* (Island of Rhodes, Greece, 1998).
15. R. O. Duda and P. E. Hart, *Commun. ACM* **15** (1), (1972).