A New Cardiac Motion Estimation Method Based on a Spatio-Temporal Frequency Approach and Hough Transform

N Carranza[1], G Cristóbal[1], MJ Ledesma-Carbayo[2], A Santos[2]

[1]Instituto de Óptica (CSIC), Madrid, Spain
[2]Universidad Politécnica de Madrid, Madrid, Spain

Abstract

Myocardial motion analysis and quantification is of utmost importance for analyzing contractile heart abnormalities and it can be a symptom of a coronary arterial disease. A fundamental problem in processing sequences of images is the computation of the optical flow, which is an approximation to the real image motion. This paper presents a new algorithm for optical flow estimation based on a spatiotemporal-frequency (STF) approach, more specifically on the computation of the Wigner-Ville distribution (WVD) and the Hough Transform (HT) of the motion sequences.

Experimental results are compared with variational techniques, where it is shown that the main highlight of the STF+HT approach is its accurate and robust response to noise degradations. Results obtained in real cardiac magnetic resonance images are presented.

1. Introduction

The non-invasive quantitative estimation of regional cardiac deformation has important clinical implications for the assessment of viability of the heart wall. The analysis of heart wall deformation provides quantitative estimates of the location and extent of ischemic myocardial injury. Among all available imaging methods, Cardiac Magnetic Resonance Imaging (CMR) is 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 [1].

An approximation of real image motion is the computation of optical flow, which is a fundamental topic in processing video sequences. However, the estimation of optical flow is a challenging problem in this kind of image analysis because of a wide range of possible motions and presence of noise. In addition to this, the non-rigid motion of the heart makes cardiac motion estimation a complex problem.

In this paper, we propose a new algorithm for optical flow estimation based on a spatiotemporal-frequency approach. STF representations are used for the description and understanding of signals whose frequency content is changing with time (non-stationary signals). Due to the presence of non-rigid motion in cardiac sequences, STF representations are the best candidates for motion determination. Within the field of non-stationary signals, our approach uses the Wigner-Ville distribution together with the Hough Transform. The later is a well known line and shape detection method, very robust against incomplete data and noise [2]. The rationale of using the HT in this context is because 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 detected.

The paper is organized as follows. Section 2 briefly provides an explanation of the method proposed. Section 3 evaluates the method and shows the results and Section 4 concludes the paper.

2. Methods

The motion estimation in the frequency domain is based on one fundamental property (Fourier shift theorem), which can be derived by analyzing a video sequence through a three dimensional Fourier transform.

Among the different techniques for computing optical flow using frequency-based methods, the spatiotemporal-frequency approach has been proposed. 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 [3]. In the field of non-stationary signal analysis, the WVD has been mainly used for the representation of speech and image.

The WVD distribution of a moving sequence is a 6-dimensional function defined as

$$W_i(x,y,t,w_x,w_y,w_t)=$$
$$\oiiint R_i(x,y,t,\alpha,\beta,\tau)e^{-j(\alpha w_x+\beta w_y+\tau w_t)}d\alpha d\beta d\tau \qquad (1)$$

where

$$R_i(x,y,t,\alpha,\beta,\tau)=$$
$$i(x+\alpha,y+\beta,t+\tau)i^*(x-\alpha,y-\beta,t-\tau) \qquad (2)$$

and where * denotes complex conjugation.

We assume that we can represent an image sequence through a function $i_0(x,y)$ such as

$$i(x,y,t)=i_0(x-v_xt,y-v_yt) \qquad (3)$$

where the main assumption here is that moving objects must move with a uniform velocity vector (v_x, v_y) and must have a constant illumination. The WVD of this time-varying image $i(x,y,t)$ is

$$W_i(x,y,t,w_x,w_y,w_t)=$$
$$\delta(v_xw_x+v_yw_y+w_t)W_i(x-v_xt,y-v_yt,w_x,w_y) \qquad (4)$$

From (4), the WVD of a linearly translating image with velocity (v_x,v_y) is everywhere zero except in the plane defined by

$$\{(x,y,t,w_x,w_y,w_t): v_xw_x+v_yw_y+w_t=0\} \qquad (5)$$

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 (5), which is called the motion plane [3]. 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.

The main scheme of the implementation of the algorithm for motion estimation is represented in Fig. 1.

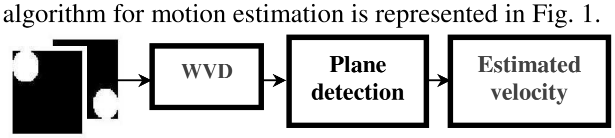

Figure 1. Scheme of the method proposed.

Here, we propose the use of the HT for detecting the position and slope of the plane. The HT detects an object and estimates the pose parameters by computing the largest subset of image features fitting a rigid template. Its strong points are the ability to discard features belonging to other objects and the robustness against incomplete data and noise [2]. These characteristics are very important in our problem due to the fact of the presence of cross-terms of the Wigner-Ville distribution, so we can use a more efficient method for determining the motion plane.

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, those pixels of the WVD belonging to cross-terms might difficult the final result interpretation, and even sometimes can produce 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 way discarding the information from cross-terms pixels. Furthermore, this implementation is computationally less demanding. At the end, our problem can be reduced to find one straight line in each temporal frame.

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

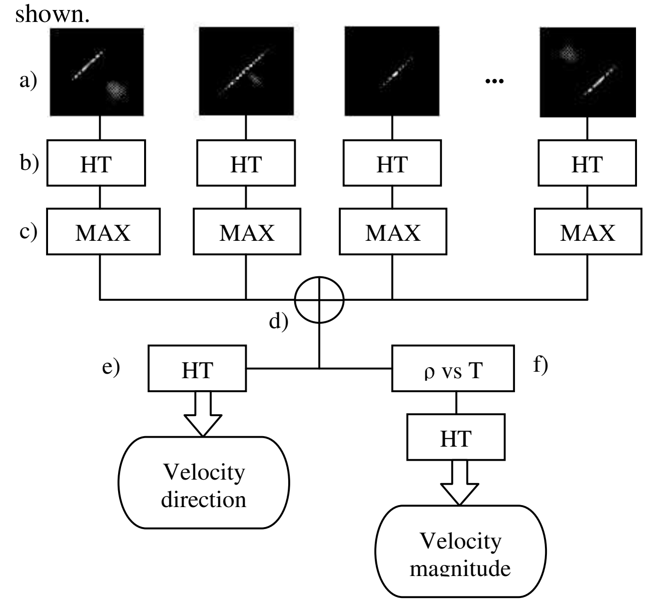

Figure 2. Flow diagram of the plane detection stage.

The next step of the algorithm is performing the HT for each of the frames of the spectrum (see Fig.2-b). One example of the result obtained after this stage is shown in Fig.3-a. Taking the maximum value of this HT (Fig 2-c), 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 the HT lie all in 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.2-d), it is straightforward to find that all these maxima belong to one line (see Fig.3-b)

Actually, the information provided by the angle of one of the peaks would be enough for estimating the direction of the velocity. But most of the times, ideal conditions are not met (for example, the presence of cross-terms induced by the WVD), and in some cases, by 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. 3-b, which is that all the maxima form a straight line and by applying the HT to the summation of all the peaks (Fig 2-e), 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. 3-c shows, the values of ρ lie along a line, whose slope can be measured by means of another HT (Fig 2-f).

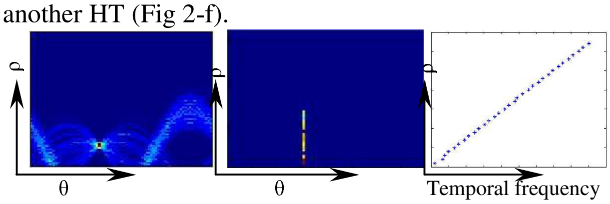

Figure 3. Left: HT of the first frame shown in Fig. 2-a. Center: Summation of all maxima of the HT of all the frames of the WVD; Right: Values of ρ of the HT for the different frames of the WVD

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 before is executed for each of the windows. However, using only one fixed size of window can lead to several problems. For this reason, the technique above described has been extended in a hierarchical coarse-to-fine framework. The problems which have being solved with the hierarchical implementation are the aliasing effect, the apperture efect and the problem of measuring large or too small image motions

3. Results

The new methodology proposed here was applied for evaluation purposes to synthetic images of a moving circular object with constant intensity.

The method was applied to distinct values of radius of the circle, initial position and velocity. Some results are shown on 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.

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	-9.49e-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

A further step on the analysis has been done, estimating the motion locally by means of the hierarchical implementation above explained. An example with the same synthetic image of a moving circular object with v_x=1 and v_y=0 is shown in Fig 4. Fig 4-a represents the optical flow obtained with the method proposed, estimating the motion locally with two hierarchical levels, where the image size is 128x128x25 pixels, and window sizes are 25 and 50 pixels. Fig 4-b represents the optical flow obtained for the same sequence with a variational approach.

With the method proposed, for these sizes of window we can observe that the optical flow estimated in the regions near to the border of the circle is very accurate (compare with Fig. 4-b). Only regions inside the circle provide uncertainty due to the apperture problem, and therefore values for optical flow are less accurate. In the case of the variational approach, 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.

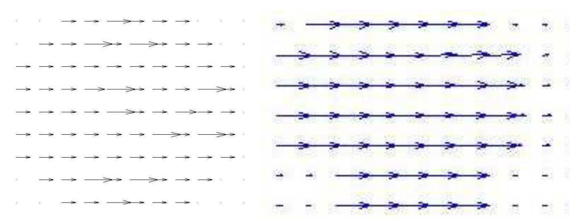

Figure 4 – Left: optical flow with local estimation using the method based on WVD-HT; Right: optical flow obtained using variational methods.

The method proposed has been applied to real CMR sequences in order to estimate the myocardial deformation, and its performance has been evaluated under these non ideal cases. In this case, the analysis is more complex because the initial conditions are not the ideal ones. The first condition is that objects must move under a uniform velocity. In a real sequence, a uniform velocity cannot be guaranteed.

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. In order to try to make uniform these small changes between the images of the sequence, a preprocessing stage has been performed, by filtering and thresholding the image. Nevertheless, the problem of the presence of a deformable object is going to be the main one.

Fig. 5-a presents several frames of one of the windows of the original sequence (extracted from a CMR), where it is shown how the object doesn't preserve the form. Therefore, the correlation of this sequence (Fig. 5-b) is variable with time. For that reason, as seen in Fig. 5-c, the spectrum doesn't have a clear plane detectable by the HT. This fact is due to the non-uniformity 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. Information about the position and shape of each frame of correlation is extracted and processed so as to get a uniform correlation where the main information from the original one is preserved. Therefore, the plane is detectable (Fig. 5-d) and the accuracy of the motion detection is improved.

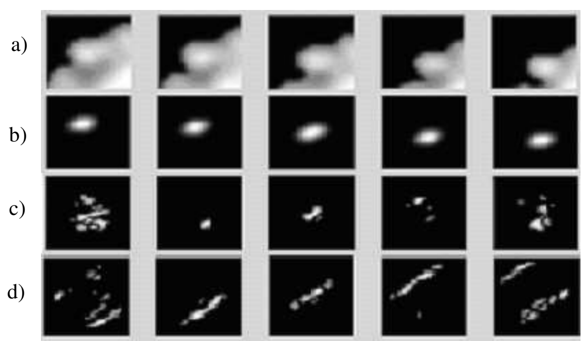

Figure 5 - a) Frames of a cropped region of a cardiac magnetic resonance. b) Correlation of the sequence. c) Spectrum obtained. d) Spectrum after processing the correlation.

The optical flow obtained for a sequence, taking frames of a systolic period is shown in Fig.6.

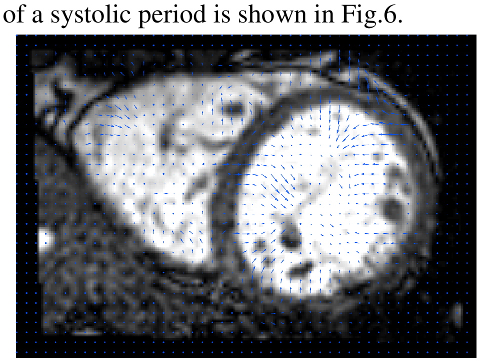

Figure 6 – Optical flow obtained during a period of systole superimposed on the first image of the sequence.

Unfortunately, the ground truth is not available in these real sequences, so it is not possible to achieve a quantitative result, but a qualitatively assessment of the direction of the optical flow obtained confirms the movement of the myocardium.

4. Discussion and conclusions

In this paper, a frequency-based method for motion estimation 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. In the case of global motion estimation, results are very close to the actual ones. For local motion estimation, results are very accurate and similar to those obtained with the variational approach.

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

Acknowledgements

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

References

[1] V. Sinitsyn, "Magnetic resonance imaging in coronary heart disease", European Journal of Radiology, Vol. 38, Issue 3, pp. 191-199, June 2001.

[2] J. S. Marques, "A probabilistic framework for the Hough Transform and Least Squares Pose Estimation", *EUSIPCO 1998*, Island of Rhodes, Greece, September 8-11. 1998.

[3] M. Pingault, D. Pellerin, "Motion estimation of transparent objects in the frequency domain", *Signal Processing*, vol. 84, pp. 709 – 719, April 2004.

Address for correspondence

Noemí Carranza Herrezuelo
Instituto de Óptica (CSIC)
Dpto. Imágenes y Visión
C/Serrano 121, 28006
Madrid (SPAIN)
noemi@optica.csic.es