Cardiac Image Sequence Analysis Using a Spatiotemporal Frequency Approach

WIETSKE INEKE MEYERING[1]
MARCO ANTONIO GUTIERREZ[2]
[1]LEB- Laboratório de Engenharia Biomédica, Escola Politécnica da Universidade São Paulo
Av. Prof. Luciano Gualberto, trav3, n.° 380, 05508-900, São Paulo, SP, Brasil
wietske.meyering@incor.usp.br
[2]InCor- Instituto do Coração, Serviço de Informática,
Av. Dr. Enéas de Carvalho Aguiar n.° 44, 05403-000, São Paulo, SP, Brasil
marco.gutierrez@incor.usp.br

Abstract. The estimation of left ventricle motion and deformation from series of images has been an area of attention in the medical image analysis and still remains and open and challenging problem. The proper motion tracking of left ventricle wall can contribute to isolate the location and extent of ischemic or infarcted myocardium. We present a method that automatically estimates the velocity vector field for a beating heart based on the study of variation in frequency content of a time series of non-stationary images. Results obtained with this automated method in synthetic images and cardiac gated-SPECT images are presented.

1 Introduction

Left ventricle contractile abnormalities can be an important manifestation of coronary artery disease. Wall motion changes may represent ischemia or infarction of myocardium (Marcassa et al., 1990). Quantifying the extent of regional wall motion abnormality can aid in determining the myocardial effects of coronary artery disease. It would also simplify the analysis of wall motion changes after diagnostic and therapeutic interventions and allows comparison of different imaging techniques to assess their diagnostic accuracy. For this reason the proper tracking of left ventricle wall can contribute to isolate the location and extent of ischemic or infarcted myocardium and constitutes a fundamental goal of image modalities, such as Nuclear Medicine.

The process to obtain Nuclear Medicine images involves the detection of the radiation emitted from a patient's organ or region after the administration of a radiopharmaceutical. Using a gamma camera, the detected radiation produce an image indicating the distribution of radionuclide in the body. This distribution represents the projection of a radioactive volume over the detector's face or, after tomographic reconstruction, the radioactive distribution in a volume section or slice.

Some radiopharmaceuticals such as ^{201}Tl and ^{99m}Tc-MIBI can provide information about the myocardium perfusion. Following intravenous injection, these radioactive substances are extracted almost completely from the plasma during the first pass throughout the various tissues. Defects on their distribution in the myocardium indicate a muscle hipoperfusion due to obstruction of the coronary arteries. Electrocardiographic gating of MIBI SPECT images provides the additional ability to determine the severity of abnormalities in wall motion and wall thickening associated with perfusion defects.

This work describes a new method to automatically estimate the velocity vector field for a beating heart based on the study of variation in frequency content of a series of non-stationary images as time varies. Results obtained with this automated method in synthetic images and cardiac SPECT images are presented.

2 Methods

Spatial and temporal frequency (STF) signal analysis is used for the description and understanding of signals whose frequency content is changing with time (non-stationary signals), which is the exact case when studying non-rigid motion in series of images. The other major motivation for considering the use of STF image representation as a basis for computing velocity vector field comes from the literature on mammalian vision. In particular, recent investigations have demonstrated that many neurons in various visual cortical areas of the brain behave as time-frequency bandpass filters (Adelson and Bergen, 1985, Gafni and Zeevi, 1977 and 1979).

1530-1834/01 $10.00 © 2001 IEEE

In the field of non-stationary signal analysis, the Wigner-Ville Distribution (WVD) has been used for the representation of speech and image. Jacobson and Wechsler (1987 and 1988) first suggested the use of the WVD for the representation of shape and texture information images. In particular, they formulated a theory for invariant visual pattern recognition in which the WVD plays a central role.

Given a time-varying image $f(x,y,t)$, its WVD is a 6-dimensional function defined as:

$$W_f(x,y,t,w_x,w_y,w_t) = \iiint_{-\infty}^{+\infty} R_f(x,y,t,\alpha,\beta,\tau) e^{-j(\alpha w_x+\beta w_y+\tau w_t)} d\alpha d\beta d\tau \quad (1)$$

where

$$R_f(x,y,t,\alpha,\beta,\tau) = f(x+\alpha,y+\beta,t+\tau).f^*(x-\alpha,y-\beta,t-\tau) \quad (2)$$

is the pseudocorrelation function, where * denotes complex conjugation.

For the special case where a time-varying image is uniformly translating at some constant velocity (v_x,v_y), the image sequence can be expressed as a convolution between a static image and a translating delta function:

$$f(x,y,t) = f(x,y) * \delta(x-v_x t, y-v_y t) \quad (3)$$

Using the convolution and windowing properties of the WVD, we obtain

$$W_f(x,y,t,w_x,w_y,w_t) = \delta(v_x w_x + v_y w_y + w_t) W_f(x-v_x t, y-v_y t, w_x, w_y) \quad (4)$$

The WVD of a linearly translating image with constant velocity is everywhere zero except in the plane defined by

$$\{(x,y,t,w_x,w_y,w_t) : v_x w_x + v_y w_y + w_t) = 0\} \quad (5)$$

for fixed (v_x,v_y). Equivalently, for an arbitrary pixel at x,y and t, each local spatial and temporal frequency spectrum of the Wigner-Ville distribution is zero everywhere except on the plane defined by

$$\{(w_x,w_y,w_t) : v_x w_x + v_y w_y + w_t) = 0\} \quad (6)$$

From (1) the Wigner-Ville Distribution assigns a three-dimensional spatiotemporal-frequency spectrum to each pixel x,y,t, over which the image is defined (Jacobson and Wechsler, 1987). However, the WVD assigns a 3D spectrum with interference due cross correlation (Cohen, 1995) when more than one frequency is present. In order to smooth the spectrum a filter must be introduced. In this work the Hanning filter was used:

$$h = 0{,}5 * \left[1 - \cos\left(\frac{2\pi n}{N}\right)\right] \quad \text{for} \quad 0 \le n \le N-1 \quad (7)$$

After obtaining the smoothed 3D-frequency spectrum, it is possible to estimate the velocity vector for each pixel by determining the best fitting plane to the points of WVD. The orientation of this plane in the space is also directly related to the pixel velocity. The plane can be obtained through a multiple linear regression (Chapra and Canale, 1998 and Kleinbaum et al., 1988):

$$z = b_0 + b_1 x + b_2 y + e \quad (8)$$

The equation (8) is a linear regression extension where z is a linear function of two independent variables. The values of the coefficients b_0, b_1, b_2 are achieved by solving the following linear system:

$$\begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \sum y_i & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum z_i \\ \sum x_i z_i \\ \sum y_i z_i \end{bmatrix} \quad (9)$$

The coefficients b_1, b_2 are related to the movement on x and y directions, respectively.

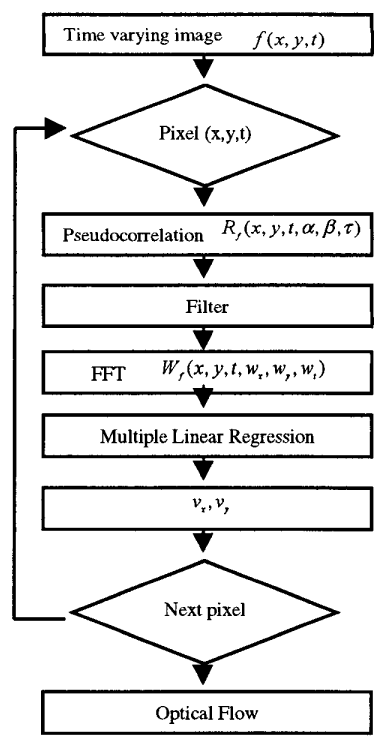

Figure 1 Flow diagram for the methodology.

Figure 1 represents the flow diagram of the methodology applied in this work.

3 Results

Synthetic images

The synthetic image consists of a texture distributed in a grid of 32x32, where dark pixels simulate hot regions in SPECT images. Figure 2 represents one frame of the synthetic image.

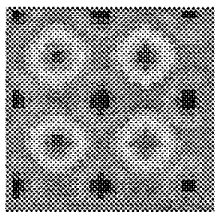

Figure 2 One slice of the synthetic image.

The velocity vector field was accomplished with the Hanning filter. The simulations were done by translations and rotations of the synthetic image.

Translation simulations were performed by translating the entire image in x and y directions. As an example of a translation simulation we obtain Table 1. In this Table, $\hat{v}$ and STD are mean and standard deviation of the estimated velocity in each direction, respectively. The Root Mean Square Error (RMSE), in percentage, was used to measure the global difference between real and estimated velocities.

Table 1 Data for the simulation of translation of 1 pixel/frame on x and y directions of the synthetic image.

r	$\hat{v}$ x	$\hat{v}$ y	STD	RMSE
1	0,9629	0,8950	0,0169	8,67
2	0,9490	0,8862	0,0287	10,32
3	0,9358	0,8781	0,0358	11,58
4	0,9206	0,8741	0,0383	12,49
5	0,9169	0,8719	0,0477	13,00
6	0,9168	0,8644	0,0546	13,75
7	0,9117	0,8593	0,0583	14,49
8	0,9036	0,8556	0,0600	15,19
9	0,8918	0,8480	0,0687	16,63
10	0,8841	0,8360	0,0763	18,30
11	0,8749	0,8258	0,0813	20,19
12	0,8629	0,8138	0,0893	22,90
13	0,8428	0,7930	0,1084	27,55
14	0,8179	0,7653	0,1362	33,57
15	0,7834	0,7297	0,1789	41,52

To verify the influence of the image border, we measured the RMSE from the center to the border of the image. Figure 3 shows the RMSE along the radii, from the center of the image to the edge for the case of translation.

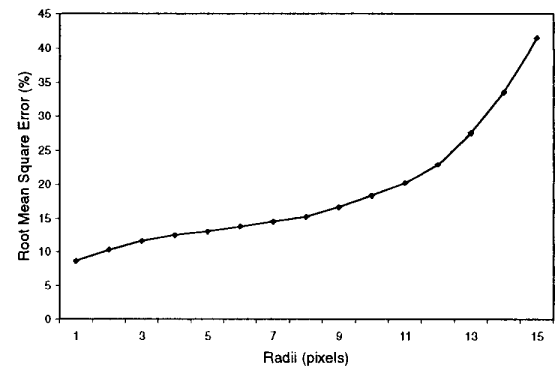

Figure 3 RMSE along the radii observed for translation of 1 pixel/frame on x and y directions.

Table 2 shows the results for rotation where $\hat{w}$ and STD are the mean and the standard deviation of the estimated angular velocity, respectively. The RMSE is also used to verify the difference between the real and the estimated angular velocities.

Table 2 Data for rotation of 6 degrees/frame on counter-clockwise direction of the synthetic image.

r	$\hat{w}$	STD	RMSE
1	5,6895	0,0352	29,76
2	4,7441	0,0263	38,50
3	4,8644	0,0276	37,58
4	5,2254	0,0255	30,25
5	5,3572	0,0213	24,98
6	5,3113	0,0185	23,30
7	5,2769	0,0166	22,19
8	5,3285	0,0151	20,16
9	5,4087	0,0139	18,04
10	5,4374	0,0129	16,83
11	5,4374	0,0122	16,31
12	5,3858	0,0118	16,49
13	5,2254	0,0117	17,77
14	5,0363	0,0120	19,39
15	4,7269	0,0136	22,83

Figure 4 shows the graphic for the RMSE along the radii from the center of the image. The error is higher at the center of the image due to the low tangential velocity of the pixels and at the edges due to the border effect.

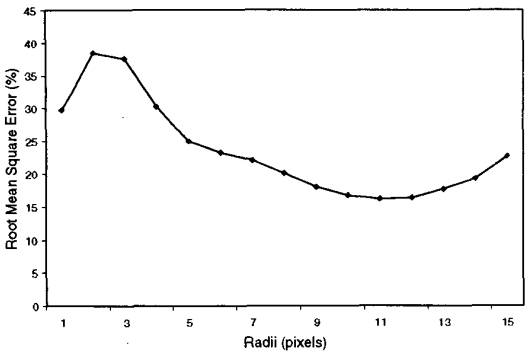

Figure 4 RMSE along the radii, for the case of rotation of 6 degrees/frame on counter-clockwise direction.

Figure 5 shows the velocity vector field obtained after translation and rotation movements imposed to the synthetic phantom.

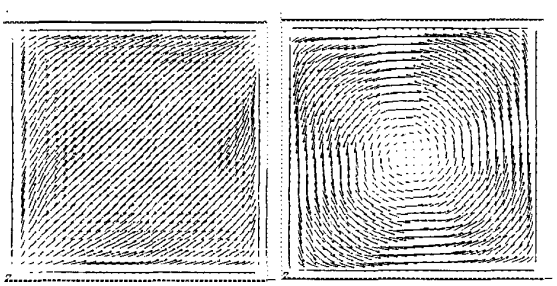

Figure 5 Velocity vector field for the translation simulation of 1 pixel/frame on x and y direction (left) and rotation of 6 degrees/frame on counter-clockwise direction (right).

The results for all simulations of translation and rotation are shown on Table 3 and Table 4, respectively. To remove the contribution of the border effects, only the pixels within 8 pixels radii are considered.

Table 3 Translation simulations data.

vx	vy	$\hat{v}$ x	$\hat{v}$ y	STD	RMSE
1	0	0,9029	0,0011	0,0310	11,69
2	0	1,7980	-0,0018	0,0762	12,58
0	1	0,0004	0,8717	0,0386	15,57
0	2	0,0001	1,7430	0,0790	15,64
1	1	0,9036	0,8556	0,0060	15,19
1	2	0,9023	1,7194	0,0985	16,42
2	1	1,7959	0,8631	0,0978	13,67
1	-2	0,9072	-1,7216	0,0962	16,28
-2	1	-1,7981	0,8675	0,0941	13,41

Table 4 Rotation simulations data.

w	$\hat{w}$	STD	RMSE
2	1,80	0,0065	22,65
4	3,59	0,0119	21,65
6	5,23	0,0151	20,16
8	7,08	0,0223	21,84
10	8,87	0,0264	20,85
12	10,67	0,0323	20,90
14	12,38	0,0401	22,15
16	14,01	0,0450	22,56
18	15,74	0,0522	23,06
20	17,68	0,0561	21,70

The errors associated with the velocity measurement are around 14,5% for translation simulations and around 22% for rotation simulations. These results show that the methodology performs well for translation and rotation applied to the phantom.

Real Images

The procedure to estimate the motion was also applied to gated-SPECT perfusion study ^{99m}Tc-MIBI obtained from a dual head rotating gamma camera (ADAC VertexPlus with a LEAP Collimator). The acquisition process is synchronized with the electrocardiogram and the cardiac cycle can be divided into 8 or 16 frames per cycle. A total of 64 projections were obtained over a semi-circular 180° orbit. All projections images were stored using a 64x64, 16 bits matrix. Transverse tomograms were reconstructed with a thickness of one pixel per slice (resolution of 6,47mm). The volume of transverse tomograms was re-oriented, and sets of slices perpendicular to the long axis (oblique transverses) and of slices parallel to the long axis (oblique coronals and sagittals) were created.

The velocity vector fields were obtained from a series of 2D gated-SPECT slices. Figure 6 depicts one oblique SPECT slice at systole and diastole and the superimposition of the velocity vector field.

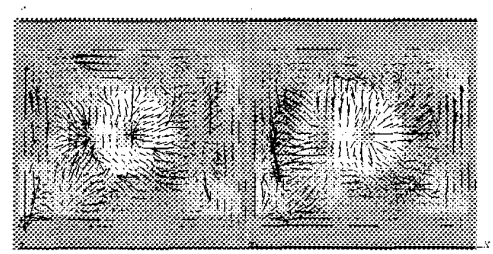

Figure 6 Oblique SPECT slice at systole (left) and diastole (right), superimposed by their velocity vector fields that delineates the non-rigid motion performed by the cardiac structures.

The computational time spend in the process for obtaining the velocity vector field for an image 32x32x32 is around 2:40 hours. One way to decrease the computational time is to use a threshold eliminating the pixels with lower counts, mainly the background. On figure 7 we use a threshold of 30% of the maximum value of the image, the computational time was reduced to 35 minutes for the same data.

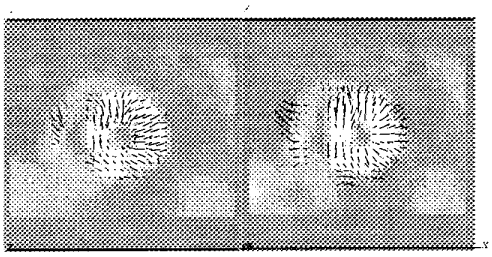

Figure 7 Oblique SPECT slice at systole (left) and diastole (right), superimposed by their velocity vector fields calculated over the pixels greater than 30% of the maximum value of the image.

4 Conclusions

This work described a new methodology to determine velocity vector field in cardiac images based on spatial and temporal frequency (STF) analysis. The majors motivations for considering the use of STF representation as a basis for computing non-rigid motion are: (1) the fact that it is used for the description and understanding of signals whose frequency content is changing with time (non-stationary signals), which is the case when studying non-rigid motion in a series of images; (2) some investigations on mammalian vision have demonstrated that many neurons in various visual cortical areas of the brain behave as spatiotemporal-frequency band-pass filters.

The method was applied in synthetic images in experiments involving translation and rotation and in real images of a gated-SPECT study yielding the velocity vector field that describes the corresponding motion.

In order to decrease the computational time a threshold was applied to the cardiac image reducing the number of pixels that has to be calculated. This solution has decreased significantly the computational time.

Further research shall include:

1. Optimization of the computational time by the segmentation of the cardiac images before the processing, eliminating the pixels without movement information.
2. Extension of the methodology to 4D images.
3. Assessment of normality patterns.

Acknowledgements

This work has been supported by Grant n°.97/14207-1 of the Foundation of Aid for Research of São Paulo State (FAPESP) and the Zerbini Foundation.

References

[1] E. H. Adelson, J. R. Bergen, "Spatiotemporal energy models for the perception of motion", *Journal of the Optical Society of America* A/12 (1985), 284--299.

[2] M. Bennamoun, "Application of time-frequency signal analysis to motion estimation". *Proceedings IEEE ICIP* (1997), 148--151.

[3] S. C. Chapra, R. P. Canale, *Numerical methods for engineers with programming and software applications.* 3rd edn. McGraw-Hill, New York, 1998.

[4] L. Cohen, *"Time-frequency analysis"*. New Jersey, Prentice Hall PTR, 1995.

[5] C. Fennema, W. Thompson, "Velocity determination in scenes containing several moving objects". *Computer Vision Graphics and Image Processing*, v. 9 (1979), 301--315.

[6] H. Gafni, Y. Y. Zeevi, "A model for separation of spatial and temporal information in the visual system". *Biological Cybernetics*, 28 (1977), 73--82.

[7] J. Gafni, Y. Y. Zeevi, "A model for processing of movement in the visual system". *Biological Cybernetics*, 32 (1979), 165--173.

[8] J. J. Gibson, *The perception of the visual world.* Boston, Houghton Mifflin, 1950.

[10] M. A. Gutierrez et al., "Computing Optical Flow in Cardiac Images for 3D Motion Analysis". In: *Computers in Cardiology 1993*. Los Alamitos: IEEE Computer Society Press., (1994), 37--40.

[11] B. K. P. Horn, B. G. Schunck, "Determining optical flow". *Artificial Intelligence*, v. 17 (1981), 185--203.

[12] L. D. Jacobson, H. Wechsler, "Derivation of optical flow using a spatiotemporal-frequency approach". *Computer Vision, Graphics and Image Processing*; v. 38 (1987), 29--65.

[13] L. D. Jacobson, H. Wechsler, "Derivation of optical flow using a spatiotemporal-frequency approach". *Signal Processing*, v. 14 (1988), 37--68.

[14] D. G. Kleinbaum, L. L. Kupper, K. E. Muller, *Applied regression analysis and other multivariable*

methods. 2nd edn, Boston, PWS-KENT Publishing Company, 1988.

[15] P. A. Laplante, A. D. Stoyenko, *Real-time imaging: theory, techniques and applications.* IEEE Press, 1996.

[16] C. Marcassa, et al. "A New Method for Noninvasive Quantitation of Segmental Myocardial Wall Thickening Using Tc -99m -2-methoxy –isobutyl -isonitrile: Scintigraphic Results in Normal Subjects". *Journal of Nuclear Medicine*, v. 31 (1990), 173--177.

[17] T. R. Reed, H. Wechsler, "Segmentation of textured images and gestalt organization using spatial/spatial-frequency representations". *IEEE Transactions on Patterns Analysis and Machine Intelligence*, v. 12 (1990), 1--12.

[18] A. B. Watson, A. J. Ahumada, "Model of human visual-motion sensing". *Journal of the Optical Society of America* A/12(1985), 322--342.