

A novel dynamic texture based approach for segmentation and registration of liver ultrasound

Sergiy Milko¹, Eigil Samset^{2,1}, Timor Kadir³

¹ University in Oslo, Department of Informatics, Gaustadaleen 23, N-0371 Oslo, Norway

² Rikshospitalet, Intervensjonssenteret, Songsvannsveien 20, N-0027 Oslo, Norway

³ Siemens Molecular Imaging, 23-38, Hythe Bridge Street, Oxford, OX1 2EP, UK
sergiym@ifi.uio.no, eigilsa@ifi.uio.no, timor.kadir@siemens.com

Abstract. Ultrasound examination has become a vital part of many interventional procedures, including Radio Frequency Ablation (RFA) of liver cancer – the application of focus in this paper. Automated analysis of ultrasound is challenging due to shadowing effects, a limited field of view and, most problematic, time varying speckle. We propose a novel method to model the time dependent appearance of ultrasound using dynamic textures. We mobilize autoregressive models to describe and differentiate between dynamic appearances of several anatomical structures, such as hepatic vasculature and liver parenchyma. The discriminative power of the method is demonstrated by two applications: segmentation of dynamic liver ultrasound and registration of intra-operative ultrasound to preoperative MR during RFA procedure. Quantitative assessment of the method is reported by comparing the manual and automatic segmentations of liver ultrasound using Dice Similarity Coefficients.

Introduction

Ultrasound is a widely available, cost effective and highly interactive imaging modality. The need for automated analysis of ultrasound (US) arises from several clinical applications [1, 2]. The application of interest for our work is ultrasound guidance during Radio Frequency Ablation (RFA) of liver cancer. Two precursors for the successful RFA are the registration of preoperative plan to intra-operative US and identification of critical structures, such as large vessels, on ultrasound images. Hence in this paper we focus on two applications: segmentation of hepatic vessels and registration of preoperative MR to intra-operative ultrasound.

Whereas a number of robust methods have been developed for the analysis of CT [3], MR and other imaging modalities [4], ultrasound received far less attention. Automated analysis of ultrasound can be more challenging due to a number of artefacts, which are invariably present on US images: shadowing, intensity inhomogeneity and, most problematic, time varying speckle.

Most approaches for ultrasound analysis to date, with few exceptions [5, 6], are typically based on the assumption that the dominant source of information for tissue discrimination in US images lies at the interface between tissue classes. Therefore, considerable efforts were made to detect and track the sparse edges, with or without

shape models, with the corollary that tissue regions were largely ignored and considered to be primarily a source of noise.

A radically different approach is adopted in this paper. We note that anatomical structures are more clearly visible in US when viewed as a continuous sequence; indeed, to facilitate interpretation, clinicians tend to view US as a sequence. Furthermore, the US appearance, including speckle, is highly correlated at least in the short term. Therefore we propose to analyze and to interpret the dynamic content of ultrasound directly – in particular we assume that different anatomical structures exhibit different appearances in US including for instance, speckle. This approach enables us to study and utilize the appearance of the tissue region to drive image analysis in addition to the boundary information. This is beneficial as the dominant source of “noise” in many traditional approaches becomes an additional source of information in the proposed method.

Analysis of the US speckle was addressed in few recent studies, with applications including registration of consecutive US images [7, 8], characterization of scattering media [9], correlation between linear tissue transformations [10] and velocity estimation [11]. In [12] the principles of speckle formation were investigated for tissue characterization problem. The speckle dynamics was modeled based on different assumptions including additive Gaussian [7, 10] and multiplicative Rayleigh [8, 11] noise.

To capture the behavior of the appearance of the speckle we consider the US sequence as a dynamic texture. In the computer vision literature, dynamic textures have been proposed as means of analyzing and synthesizing image sequences [13, 14]. Segmentation of a single [15, 16] and multiple moving dynamic textures [17], registration of dynamic scenes [14], and estimation of optical flow [17] are among most studied applications. Medical image analysis involving dynamic textures was reported in [18] for the segmentation of beating heart from MRI. In that work, however, the appearance of the structures examined differed largely making segmentation relatively straightforward – even with standard techniques. In our application the variations in appearance between the structures of interest are quite subtle.

The primary focus of this paper is to investigate the discriminative power of the dynamic texture model. We propose to model the temporal evolution of the US appearance with dynamic textures - the outputs of time invariant linear dynamic systems. We rely on the fact that different structures produce different parameters of the corresponding dynamic textures.

Our objective is to mobilize this information for labeling the anatomy comprising the US liver, such as the hepatic vessels and parenchyma, through simultaneous analysis of both the dynamics of the images and the static US image intensity. For the registration problem the anatomical structures derived from US and corresponding features in MR drive the registration between two modalities.

Methods

Dynamic textures

A detailed analysis of the dynamic textures comprising learning, modeling, synthesis and recognition is given in [13, 19]. Some definitions required to apply the dynamic texture theory to our problem are given here.

Let $\{I(t) | t = 1..\tau\}, I(t) \in R^m$ be a sequence of images. At every time point t we measure a noisy version of the image, $y(t) = I(t) + w(t)$ where $w(t) \in R^m$ is an independent and identically distributed (*IID*) sequence drawn from a known distribution. This gives a positive measured sequence $y(t) \in R^m, t = 1..\tau, m$ - number of image pixels. We call the sequence $\{I(t)\}$ a (*linear*) *dynamic texture* if there exists a set of n spatial filters $\phi_\alpha : R \rightarrow R^m, \alpha = 1..n$ and a state variable $\bar{x}(t) \in R^n$ such that

$I(t) = \sum_{\alpha=1}^n \phi_\alpha(x_\alpha(t))$, and the state variable evolves according to $\bar{x}(t) = \sum_{i=1}^k A_i \bar{x}(t-i) + Bv(t)$, for some choice of matrices $A_i \in R^{n \times n}, i = 1..k$, $B \in R^{n \times n_v}$, initial condition $\bar{x}(0) = \bar{x}_0$, and $v(t) \in R^{n_v}$ - an *IID* realization from some stationary distribution. Note, that $x_\alpha(t)$ refers to the α -th component of the vector $\bar{x}(t)$. A dynamic texture can then be associated with an auto-regressive (AR) process

$$\begin{cases} \bar{x}(t) = \sum_{i=1}^k A_i \bar{x}(t-i) + Bv(t) \\ y(t) = \sum_{\alpha=1}^n \phi_\alpha(x_\alpha(t)) + w(t) \end{cases} \quad (1)$$

with $\bar{x}(0) = \bar{x}_0$. The choice of filters $\phi_\alpha, \alpha = 1..n$ can be either biologically motivated, or related to computational efficiency [13, 14].

Application to the characterization of liver structures

In this work we disregard the filters $\{\phi_\alpha\}$ and focus on modeling the dynamics of individual pixels. We consider the intensity function at each pixel p as a state variable $x_p(t) \in R^1$. A sequence $\{x_p(t), t = 1..\tau\}$ can be treated as a time series, a sequence of 1D observations. Ultimately, the time varying appearance of each pixel is modeled as an output of an *Auto Regressive (AR) process*:

$$x_p(t) = A_1 x_p(t-1) + A_2 x_p(t-2) + \dots + A_k x_p(t-k) + Bv(t) \quad (2)$$

where $\{A_i, i = 1..k\}$ and B are real numbers characterizing the sequence. The term $v(t)$ represents a residual noise and is modeled as a Gaussian with zero mean and unit variance. Using least squares fitting, one can estimate the order k , coefficients $\{A_i, i = 1..k\}$ and B , that best explain the measurements $\{x_p(t), t = 1..\tau\}$.

In the formulation above, the pixel observations, $\{x_p(t), t = 1..\tau\}$, are naturally non-negative with positive mean Mx_p . To utilize standard techniques for the estimation of AR models with zero mean and following [14], we subtracted Mx_p from the sequence $\{x_p(t), t = 1..\tau\}$ before the least square fitting.

The proposed tissue characterization algorithm utilizes the AR model coefficients $\{A_i, i = 1..k\}$ and B to label each pixel with a tissue class. Initial experiments demonstrated that the majority of information is contained in A_1 and A_2 , whereas $\{A_k | k > 2\}$ do not reliably differentiate among underlying anatomical structures. Intuitively, this means that the two US time-points following the current time-point contain most of the information of the dynamic variation. Therefore we explicitly selected the order $k = 2$ in (2).

To cope with motions caused by heart beating and tremor of the hand-held probe we “averaged” the dynamics over the pixel neighbourhood by concatenating the observations of neighbouring pixels and estimating the AR model for the resulting 1D sequence (fig.1). This proved more stable than per-pixel AR fitting. The estimation of the AR model coefficients was performed with Neumaier and Schneider’s `ARfit` algorithm [20], explicitly specifying the order $k = 2$.

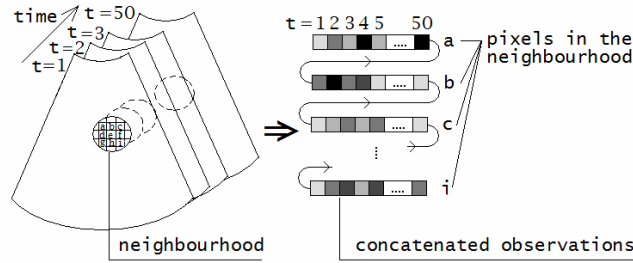


Fig. 1. Concatenated observations for AR model fitting

Segmentation

AR coefficients A_1 , A_2 and B estimated for each pixel p together with its mean intensity $I_p = Mx_p$ produce a sample 4D feature space. Analysis of the empirical distribution of sample points in this space suggested that differentiation of four anatomical structures could be feasible (fig.2): liver parenchyma, hepatic vessels, external (w.r.t liver) vascular structures and compressed tissues between the US transducer head and liver surface.

Our objective was to develop a supervised classification algorithm in 4D feature space, which would produce an automatic classification of parametric vectors (A_1, A_2, B, I_p) generated from the US sequences. For training, ten ultrasound sequences ($\tau =$

50) were manually segmented. During the acquisition of each sequence the US probe was held still with respect to the anatomy. This meant that a single segmentation per sequence could be used since boundaries were largely stationary and time independent.

Another ten sequences, not included into the training set, were utilized to assess the segmentation performance. We used two classification techniques: a K-nearest neighbour ($k=20$) and a maximum a posteriori (MAP) classifier. Automatic and manual segmentations were compared by calculating Dice Similarity Coefficients (DSC) for vessels, parenchyma and compressed area.

Registration

Following [2] we registered the MR to a sparse set of segmented US images in 3D. An optimal (rigid) registration of the liver from the US and MR can be achieved by aligning the vessels derived from both the US and MR datasets. To this end hepatic vessels were manually segmented from the MR scan. Manual MR segmentation was preferred to facilitate the examination of the efficacy of the proposed technique separately from any MR based processing.

The US data consisted of sequences acquired with a stationary US probe. Segmentations were related in 3D space by utilizing the position and orientation of the US transducer equipped with a tracking sensor. The initial alignment was achieved by the Iterative Closest Point method by matching seven external fiducial markers. Registration was performed by Powell optimization of the rigid transformation maximizing the overlap between hepatic vessels derived from both modalities.

Data

We used a GE System Five US scanner with a 3.5 MHz abdominal probe to acquire B-mode images of the liver. To facilitate access to the liver, sequences were captured at a maximum inhalation phase. To cover a sufficient portion of the organ yet to ensure stable tissue boundaries on each US sequence, we adopted a “move-stop” acquisition protocol. The operator was sweeping over the liver stopping every ~1cm and keeping the transducer stationary for ~3 seconds while the next sequence was recorded. The scanning depth was set to 13cm to ensure a good field of view and to keep the capture frequency at 22.8fps. It proved sufficient to capture the dynamics and gained enough images (>50) in each sequence for the AR model fitting.

Results

An example of the segmented US sequence is presented on fig.2. MAP (fig.2.a) and Knn (fig.2.b) segmentations in 4D parametric space are shown together with MAP segmentation based purely on pixel intensity (fig.2.c). Segmentation done by the

expert is shown as a reference (fig.2.d). Vascular structures are delineated with white contours. A compressed tissue area and a liver parenchyma are separated with black boundaries.

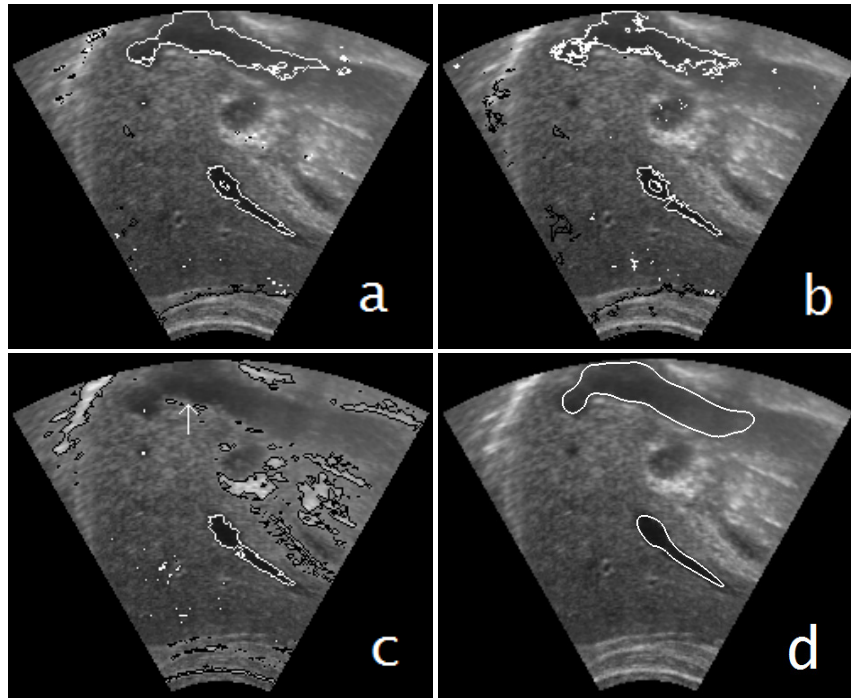


Fig. 2. Automatic segmentation of the US sequence of the liver: a - MAP classifier in 4D feature space; b - Knn classifier ($k = 20$); c - MAP classifier based purely on US intensity; d - segmentation by the expert.

Table 1. Comparison of automatic and manual segmentation: Dice Similarity Coefficients

	Vasculature			Parenchyma			Compressed area		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
MAP, 4D	0.52	0.32	0.78	0.92	0.88	0.95	0.79	0.58	0.89
Knn, 4D	0.51	0.31	0.74	0.92	0.88	0.94	0.79	0.63	0.89
MAP, intensity	0.16	0.08	0.44	0.88	0.86	0.92	0.26	0.17	0.32

Table 1 shows quantitative results of automatic and manual segmentations of ten ultrasound sequences. Dice Similarity Coefficients (DSC) were calculated for vasculature, liver parenchyma and compressed tissue area. The internal hepatic and external vasculature were unified for conciseness.

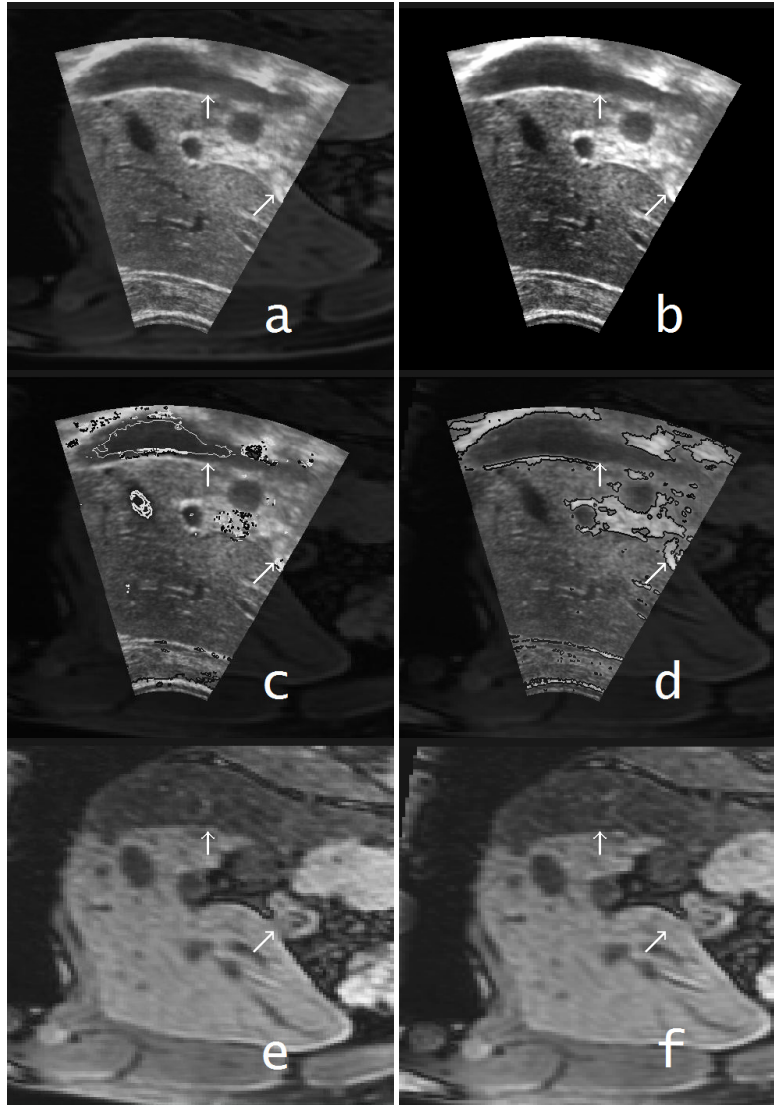


Fig. 3. Registration of the MR and US: a – initial marker-based alignment; b – US image; c, d – registrations of the MR and segmented US: in 4D feature space (c) and purely by intensity (d); e, f – MR cross-sections corresponding to registrations c, d, respectively.

Maximal, minimal and average DSCs are shown for MAP and Knn classifications in a 4D feature space as well as for purely intensity based MAP segmentation. Proximity of the DSC to 1 indicates a good correspondence between manual and automatic segmentations.

Two registrations of the MR to the segmented US are introduced on fig.3. Images show a sample US image and a corresponding cross-section of the registered MR. Fig.3.a demonstrates the marker-based initialization for both registrations. An automatic registration of the MR and US-segmentation in 4D parametric space is displayed on fig.3.c, e. The MR registered to purely intensity-based US-segmentation is shown on fig.3.d, f. White arrows point to the anatomical structures on the US and to the corresponding locations on the registered MR. Segmentations are delineated on fig.3.c, d following the conventions used on fig.2.

Table 2. Computational time for segmentation/registration steps, in seconds (Processor: 2GHz, RAM: 2GB)

	One US sequence		Entire US dataset	
AR model fitting (MATLAB)	342		11628	
	MAP	Knn	MAP	Knn
Tissue classification in 4D param. space (C++)	5	38	162	1292
	+In 4D param. space		+Purely by intensity	
Registration of the US-segmentation ^{*+} and the MR (C++)	56		30	
Total segmentation ^{*+} and registration time	11846		12950	

* - by MAP classifier

Segmentation, registration and total computational time is summarized in Table 2. Note, that registration was performed for the entire set of segmented US sequences.

Discussion & Conclusion

Fig. 2 clearly demonstrates that simultaneous analysis of dynamics and intensity of US has sufficient power to discriminate a number of tissue classes of interest. Comparison between purely intensity-based (fig.2.c) and combined intensity and dynamic texture based (fig.2.a, b) segmentations illustrates the improved discriminative power brought about through the dynamic texture representation. Even though internal vasculature was usually extracted by the purely intensity-based classification, it often failed to capture the hepatic vein located further from the probe. This is demonstrated on fig.2.c (white arrow). Moreover, Table 1 shows that despite distant vascular structures and liver parenchyma exhibiting overlapping intensities, they can be separated by analyzing their dynamic appearance in US: average vascular DSC is three times larger for the dynamics-based segmentations. This emphasizes the added value of considering the dynamics when addressing the problem of US analysis.

The benefit of utilizing the dynamics for the US analysis is shown in another application. Fig.3 illustrates two registrations of the preoperative MR to intra-

operative US based on the proposed method. Automatic registration of the MR and US segmented in 4D parametric space (fig.3.c, e) is compared with that of the MR and purely intensity-based US-segmentation (fig.3.d, f). All figures show the same sample US image and a corresponding cross-section of the registered MR. Visual assessment of both registrations across all US images testified that the dynamic texture based US-segmentation produced more accurate alignment with the MR than purely intensity-based tissue labeling. This is due to a larger basis of vessels extracted by the segmentation in 4D space; five times bigger than by the intensity-only classification. We believe that this is large distant vessels captured by the analysis of the US dynamics (white contours on fig.3.c) that significantly increase the registration accuracy. White arrows on fig.3 illustrate the improved registration by pointing to the anatomical structures on the US and corresponding locations on the registered MR. Different amounts of vasculature extracted by the dynamics-based and pure intensity-based classifications are also reflected in Table 2. Even though taking less than a minute, the registration of the dynamics-based segmentation is twice longer than that of the intensity-based segmentation.

Our experiments suggested that the choice of the classification technique is important for both applications. Two classification algorithms were compared in this paper: a maximum a posteriori (MAP) and a Knn classifier. Although Table 1 suggests a comparable performance of both, the MAP classifier generally produced smoother segmentations (fig.2.a) than the Knn classifier (fig.2.b). Furthermore, the MAP-classification is almost eight times faster than the Knn-segmentation, as shown in Table 2. This is due to the pre-computed distribution of tissue classes in the 4D parametric space achieved by the MAP algorithm. The MAP classifier was therefore our choice for the registration workflow.

Although novel and promising, the US representation introduced in this paper has certain limitations in its current form. Its major assumption is stationarity of structure boundaries throughout the image sequence. As a result a continuous sweep over the liver with constantly changing field of view cannot be reliably segmented. Therefore we adopted a “move-stop” acquisition protocol when doing the MR/US registration. We plan to address this issue in future work by optimizing of the acquisition protocol and adapting the representation to include spatial components.

Table 2 highlights the bottleneck in the current implementation of the proposed method. MATLAB-based AR-model fitting takes ~6 min for one US sequence and over 3 hours for the entire US dataset. This prevents the approach from being used in the clinical environment in its current form. We believe that effective implementation of the `ARfit` [20] package in a compiled language such as C++ will significantly reduce the computational time. Other directions for future work include the development of fully automated registration method and a thorough evaluation of the proposed approach.

Acknowledgement

This research was conducted within the framework of the ARIS*ER (Augmented Reality in Surgery, Education & Research) project funded by Marie Curie RTN.

References

1. Wein, W., Röper B., Navab N.: Integrating diagnostic B-mode ultrasonography into CT-based radiation treatment planning. *IEEE Trans. on Med. Imag.*, Vol. 26, No. 6 (2007) 866–879
2. Blackall, J.M., Penney, G.P., King, A.P., Hawkes, D.J.: Alignment of Sparse Freehand 3-D Ultrasound With Preoperative Images of the Liver Using Models of Respiratory Motion and Deformation. *IEEE Transactions on Medical Imaging*, Vol. 24, No. 11 (2005) 1405–1416
3. Soler, L., Delingette, H., Malandain, G., Montagnat, J., Ayache, N., Koehl, C., Dourthe, O., Malassagne, B., Smith, M., Mutter, D., Marescaux, J.: Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, Vol. 6, No. 3 (2001) 131–142
4. Duncan, J.S., Ayache, N.: Medical image analysis: progress over two decades and the challenges ahead. *IEEE Trans on Pattern Anal. and Mach. Intell.*, Vol. 22, No. 1 (2000) 85–106
5. Zhan, Y., Shen, D.: Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. *IEEE Trans. on Med. Imag.*, Vol. 25, No. 3 (2006) 256–272
6. Zhan, Y., Shen, D.: Automated Segmentation of 3D US Prostate Images Using Statistical Texture-Based Matching Method. *Lecture Notes in Computer Science*, Vol. 2878 (2003) 688–696
7. Mellor, M., Brady, M.: Fluid Registration of Ultrasound using Multi-scale Phase Estimates. In *Proceedings of BMVA'05* (2005)
8. Dinstein, I., Cohen, B.: New maximum likelihood estimation schemes for noisy ultrasound images. *Pattern Recognition*, Vol. 35 (2002) 455–463
9. Dutt, V., Greenleaf, J.F.: Ultrasound echo envelope analysis using a homodyned k distribution signal model. *Ultrasonic Imaging*, Vol. 16, No. 4 (1994) 265–287
10. Bertrand, M., Meunier, J.: Ultrasonic texture motion analysis: Theory and simulation. *IEEE Trans. on Med. Imag.*, Vol. 14, No. 2 (1995) 293–300
11. Boukerroui, D.: Velocity estimation in ultrasound images: A block matching approach. *Lecture Notes in Computer Science*, Vol. 2732 (2003) 586–598
12. Thijssen, J. M.: Ultrasonic speckle formation, analysis and processing applied to tissue characterization. *Pattern Recognit. Lett.*, Vol. 24, No. 4-5 (2003) 659–675
13. Doretto, G., Chiuso, A., Wu, Y.N., Soatto, S.: Dynamic Textures. *International Journal of Computer Vision*, Vol. 51, No. 2 (2003) 91–109
14. Fitzgibbon, A.: Stochastic rigidity: image registration for nowhere-static scenes. In *Proc. of IEEE Intern. Conference on Computer Vision*, Vol. 1 (2001) 662–669
15. Doretto, G., Cremers, D., Favaro, P., Soatto, S.: Dynamic Texture Segmentation. In *Proceedings of the 9th IEEE International Conference on Computer Vision*, Vol. 2 (2003) 1236–1242
16. Ghoreyshi, A., Vidal, R.: Segmenting Dynamic Textures with Ising Descriptors, ARX Models and Level Sets. *Lecture Notes in Computer Science*, Vol. 4358 (2007) 127–141
17. Vidal, R., Ravichandran, A.: Optical flow estimation & segmentation of multiple moving dynamic textures. *Computer Vision and Pattern Recognition*, Vol. 2, No. 20-25 (2005) 516–521
18. Ravichandran, A., Vidal, R., Halperin, H.: Segmenting a beating heart using polysegment and Spatial GPCA. *3rd IEEE Intern. Symp. on Biomedical Imag.: Nano to Macro*(2006)634–637
19. Doretto, G.: *Dynamic Textures: Modeling, Learning, Synthesis, Animation, Segmentation, and Recognition*. Ph.D. Thesis, University of California, Los Angeles, CA (2005)
20. Schneider, T., Neumaier, A.: Algorithm. ARfit – A Matlab Package for the Estimation of Parameters and Eigenmodes of Multivariate Autoregressive Models. *ACM Transactions on Mathematical Software*, Vol. 27 (2001) 58–65