Robust Real-Time Myocardial Border Tracking for Echocardiography: An Information Fusion Approach

Dorin Comaniciu, *Senior Member, IEEE,* Xiang Sean Zhou, and Sriram Krishnan

***Abstract*— Ultrasound is a main non-invasive modality for the assessment of the heart function. Wall tracking from ultrasound data is, however, inherently difficult due to weak echoes, clutter, poor signal-to-noise ratio, and signal dropouts. To cope with these artifacts, pre-trained shape models can be applied to constrain the tracking. However, existing methods for incorporating subspace shape constraints in myocardial border tracking use only partial information from the model distribution, and do not exploit spatially varying uncertainties from feature tracking. In this paper, we propose a complete fusion formulation in the information space for robust shape tracking, optimally resolving uncertainties from the system dynamics, heteroscedastic measurement noise, and subspace shape model. We also exploit information from the ground truth initialization where this is available. The new framework is applied for tracking of myocardial borders in very noisy echocardiography sequences. Numerous myocardium tracking experiments validate the theory and show the potential of very accurate wall motion measurements. The proposed framework outperforms the traditional shape-space-constrained tracking algorithm by a significant margin. Due to the optimal fusion of different sources of uncertainties, robust performance is observed even on the most challenging cases.**

***Index Terms*— myocardial border tracking, subspace constraint, motion estimation with uncertainty, heteroscedastic noise, information fusion, active shape model, model adaptation.**

I. Introduction

THE tracking of the myocardial border is a very active research area. Model-based deformable templates [1], [2], [3], Markov random fields [4], optical flow methods [5], [6], [7], [8], or combinations of above [9], have been applied for tracking left ventricle (LV) from 2-D image sequences. Jacob et al. provided a brief recent review in [10]. Other related work focuses on the tracking, segmentation, or registration of heart walls in 3D or 2D+T (spatial + time) [11], [12], [13], [14] or 4-D space [15].

It is a common practice to impose model constraints in a shape tracking framework. Examples include simple models such as blobs [16] or parameterized ellipses [17], and complex models such as discriminative templates [18]. In most practical cases, a subspace model is suitable for shape tracking, since the number of modes capturing the major shape variations is limited and usually much smaller than the original number of feature components used to describe the shape [2], [1].

A straightforward treatment is to project tracked shapes into a PCA subspace [19], [1]. However, this approach cannot take advantage of the measurement uncertainty and is therefore not complete: In most real-world scenarios, measurement noise is heteroscedastic in nature (i.e., both anisotropic and inhomogeneous [20]). A good example is echocardiography (ultrasound heart images) [21]. Ultrasound is prone to reflection artifacts, e.g., specular reflectors, such as those that come from membranes. Because of the single "view direction", the perpendicular surface of a specular structure produces strong echoes, but tilted or "off-axis" surfaces may produce weak echoes, or no echoes at all (acoustic "drop out"). For an echocardiogram, the drop-out can occur at the area of the heart where the tissue surface is parallel to the ultrasound beam (See Figure 1). Intuitively, a tracking algorithm should weigh less information from such regions.

Alternatively, one could directly incorporate a PCA shape space constraint into a Kalman filter-based tracker. In [22], [23] it is suggested to set the system noise covariance matrix to be the covariance of a PCA shape model. This treatment has some limitations. First of all, it does not provide a systematic and complete fusion of the model information because, for example, the model mean is discarded. Secondly, it mixes the uncertainty from system dynamics with the uncertainty from the statistical shape constraint, while these two can be conceptually different. For example, we may want to use the dynamic model to *capture different modes* of global rigid motion, while applying a statistical shape model to *constrain the range of shape variations*. In other words, the system dynamic model can be trained specifically to deal with slow translation or rotation caused by, for example, respiration; and scale change from the contraction and expansion of the heart. Finally, existing solutions do not specifically address the issue of heteroscedastic measurement noise and its influence during the fusion with other information sources. When measurement noise is anisotropic and inhomogeneous, joint fusion of all information sources becomes critical for achieving superior performance.

In this paper we decouple the uncertainty in system dynamics and the statistical shape constraint, and introduce a unified framework for fusing *a subspace shape model with the system dynamics and the measurements with heteroscedastic noise*. Where available, the generic shape model is strongly adapted using information given about the current shape. The subspace model can take the form of a specific subspace distribution, e.g., a Gaussian, or a simple subspace constraint, e.g., the eigenspace model [24], [25]. Unlike existing *ad hoc* formulations, our framework treats the two cases in a consistent way, and combines such constraints seamlessly into the tracking framework.

Manuscript received ——–, 2003; revised ——–, 2003.

D. Coamniciu and X. S. Zhou are with Real-Time Vision and Modeling Department, Siemens Corporate Research, Princeton, NJ 08540.

S. Krishnan is with Computer Aided Diagnosis Group, Siemens Medical Solutions, Malvern, PA 19355.

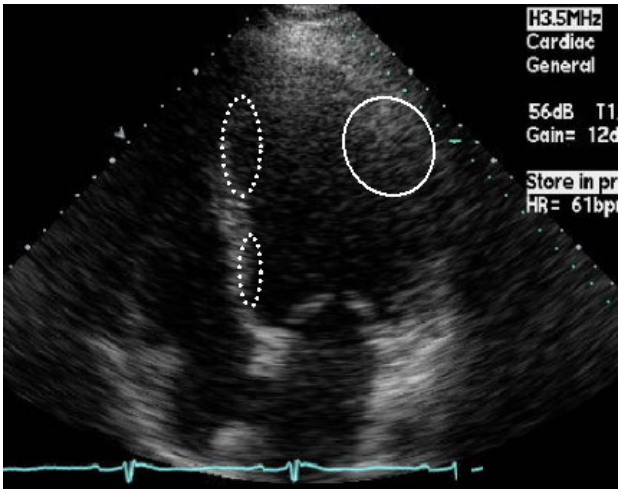


Fig. 1. Echocardiography images with area of acoustic drop-out (marked by the solid ellipse). Local wall motion estimation has covariances (depicted by dotted ellipses) that reflect heteroscedastic noise.

The new approach calls for reliable estimation of measurement uncertainties, for which we employ a recent robust solution to the motion estimation problem, which also computes the motion flow uncertainties [26].

In this paper we demonstrate the advantages of the proposed framework using ultrasound heart sequences. Ultrasound is the noisiest among common medical imaging modalities such as MRI or CT. Echocardiography is even worse due to the fast motion of the heart muscle, and respiratory interferences [21]. With spatially varying noise characteristics, echocardiograms are the most appropriate data for testing our heteroscedastic fusion framework.

The paper is organized as follows: The new shape tracking formulation using complete fusion is introduced in Section III. Section IV discusses shape model adaptation within this framework. Our robust motion computation algorithm along with measurement uncertainty estimation is presented in Section V. Section VI presents experiments on shape tracking for cardiac ultrasound. We discuss related work and future directions in Sections II and VIII.

II. Related Work

In this section, we discuss related work on tracking with subspace model constraints and heteroscedastic noise.

If the measurement is affected by heteroscedastic noise, an orthogonal projection into the constraining subspace is not only unjustified, but also very damaging in terms of information loss [27]. It can only be justified for the special case when the noise is both isotropic and homogeneous.

However, most existing work on subspace-constrained tracking did not take into account the heteroscedastic noise in the measurements [28], [24]. In the "Point Distribution Model" or "Active Shape Model" [19], [2], [1], a PCA-based subspace shape model is derived based on training shapes with landmark point correspondence. At detection time, a model is perturbed to create synthetic images for matching against the testing image at a candidate location. Measurement noise was not modeled in this process. Even when heteroscedastic noise characteristics are available, the fusion with the subspace model constraint is not complete in the existing tracking frameworks [29], [30], [22], [10]. The model mean is discarded in the formulation of [30], [31]; while in [29] a rather ad hoc thresholding (proposed in [1]) is applied so that the measurement mean is confined to a hyper-ellipsoid constraint defined by the model covariance. This operation did not even fuse information from measurement uncertainty.

Measurement uncertainty has been exploited for tracking and motion estimation in different contexts. However, none of these have put all relevant information sources into a unified fusion formulation. Jepson, Fleet and El-Maraghi [32] used a Gaussian distribution to adaptively model the appearance of the object of interest (face in their case), which is learned using the EM algorithm. As in our case, local uncertainty is captured in the covariance matrix. However, they do not use subspace models for information fusion. A rank-constrained flow estimation formulation was proposed by Bregler et al. [33]. They use constraints from both rigid and non-rigid motion represented by basis-shapes. Although occlusion is addressed, measurement uncertainty in general is not optimally exploited. Both Brand [27] and Irani [34] use measurement uncertainties, but they did not provide a complete fusion-based tracking framework that combines all the information sources including the system dynamics and the initial input. In fact, we could adopt these low-rank constrained approaches in our flow estimation sub-module.

Leedan, Matei, and Meer [35], [36] applied heteroscedastic regression for fitting ellipses and fundamental matrices. The fitting is achieved in the original space with parameterized models. In our formulation, we avoid the parameterization of shape variations — it can be very complicated and highly nonlinear. Instead, we build subspace probabilistic models through PCA and obtain closed-form solutions on both the mean and covariance of the fitted data. Although simple, this model works fine in our experiments, especially with our proposed SA-PCA model (Section IV). However, this model may be too simplistic for other applications, and more sophisticated nonlinear models may be needed (e.g., [37], [38]).

Robust model matching [39], [40] relying on M-estimators or RANSAC has been applied to limit or eliminate the influence of data components that are outliers with respect to the model. Again, the locally (in space or time) varying uncertainties are not exploited in these frameworks.

There is much research work done in medical domain that tracks heart motion using various techniques (e.g., [5], [41], [42], [22], [43], [44], [45]). However, they did not address the issue of heteroscedastic noise and its fusion with other information sources. Jacob et al. [10] apply a two-step approach to impose a *shape space* constraint in a Kalman filtering framework. The *shape space* is a linearly transformed affine subspace or eigen-subspace. However, the projection into the shape space is orthogonal, without taking into account the heteroscedastic noise of the measurement. Another difference is that they train a model for every sequence using several contours from an expert user, while we will only use one initial

contour to start the tracker because a generic shape model is trained across patients (but updated for the current patient).

III. Subspace Fusion-Based Shape Tracking

Throughout this paper, we represent shapes by control or landmark points, assuming correspondence. These points are fitted by splines before shown to the user. The analysis is based on the vectors formed by concatenating the coordinates of all control points [10], [19]. For more implementation details, please refer to Section VI.

A typical tracking framework fuses information from the prediction defined by a dynamic process and from noisy measurements. When applied to shape tracking, additional global constraints are necessary to stabilize the overall shape in a feasible range. Some existing frameworks have addressed this issue by projecting the tracked shape or the driving noise process of the dynamic system into an eigenshape space [10], [29], [2], [30]. However, such treatments *failed to exploit the heteroscedastic measurement noise during projection* [19], or *did not fully exploit the model distribution* (as a result, the projection can be arbitrarily far from the model mean in the subspace). We propose in this section a *complete* fusion framework that addresses these shortcomings.

A. Kalman Filter as Fusion in the Information Space

For a Kalman filter, the measurement update equation has the following form [46], [47]:

$$\mathbf{x}_{k+1|k+1} = \mathbf{x}_{k+1|k} + \mathbf{K}(\mathbf{z}_{k+1} - \mathbf{H}\mathbf{x}_{k+1|k}) \tag{1}$$

where

$$\mathbf{K} = \mathbf{P}_{k+1|k}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T + \mathbf{R})^{-1} \tag{2}$$

$$\mathbf{P}_{k+1|k+1} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k} \tag{3}$$

$$\mathbf{P}_{k+1|k} = \mathbf{S}\mathbf{P}_{k|k}\mathbf{S}^T + \mathbf{Q} \tag{4}$$

Here $\mathbf{P}$ is the state covariance, $\mathbf{H}$ is the measurement matrix, $\mathbf{S}$ is the system matrix, and $\mathbf{x}_{i|j}$ is the state estimate at time i given the state at time j. The measurement model is $\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{r}_k$, where $\mathbf{r}_k$ represents measurement noise with covariance $\mathbf{R}$. The system/process model is $\mathbf{x}_{k+1} = \mathbf{S}\mathbf{x}_k + \mathbf{q}_k$, where $\mathbf{q}_k$ represents system noise with covariance $\mathbf{Q}$.

Kalman filter has many other forms, one of which being the *information filter* [47]. The *information space* is the space obtained by multiplying a vector by its corresponding *information matrix*, which is, in the Gaussian case, the inverse of the error covariance matrix. *Information filter* propagates the *information state* instead of the original state (see Appendix for derivation):

$$\begin{aligned}&\mathbf{x}_{k+1|k+1} = \\ &(\mathbf{P}^{-1}_{k+1|k} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}(\mathbf{P}^{-1}_{k+1|k}\mathbf{x}_{k+1|k} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}_{k+1})\end{aligned} \tag{5}$$

For the special case where $\mathbf{H}$ is a square matrix and admits an inverse, we can see Eq. (5) in a strict information fusion form, namely, the fusion of prediction and measurement in the information space:

$$\begin{aligned}&\mathbf{x}_{k+1|k+1} = \\ &(\mathbf{P}^{-1}_{k+1|k} + \mathbf{R}_{\mathbf{x}}^{-1})^{-1}\left[\mathbf{P}^{-1}_{k+1|k}\mathbf{x}_{k+1|k} + \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{x}_{\mathbf{z},k+1}\right]\end{aligned} \tag{6}$$

where $\mathbf{R}_{\mathbf{x}} = \mathbf{H}^{-1}\mathbf{R}(\mathbf{H}^{-1})^T$ and $\mathbf{x}_{\mathbf{z},k+1} = \mathbf{H}^{-1}\mathbf{z}_{k+1}$.

B. Static Subspace Model Fusion

Let us now turn our attention to the problem of information fusion with one of the sources in a subspace. We will derive in the sequel a unified way for incorporating subspace (model) constraints using information fusion.

Given two noisy measurements of the same n-dimensional variable $\mathbf{x}$, each characterized by a multidimensional Gaussian distribution, $\mathcal{N}(\mathbf{x}_1, \mathbf{C}_1)$ and $\mathcal{N}(\mathbf{x}_2, \mathbf{C}_2)$, the maximum likelihood estimate of $\mathbf{x}$ is the point with the minimal sum of Mahalanobis distances, $D^2(\mathbf{x}, \mathbf{x}_i, \mathbf{C}_i), i = 1, 2$, to the two centroids, i.e., $\mathbf{x}^* = \textit{argmin}\ d^2$ with

$$\begin{aligned}d^2 =& D^2(\mathbf{x}, \mathbf{x}_1, \mathbf{C}_1) + D^2(\mathbf{x}, \mathbf{x}_2, \mathbf{C}_2) \\ =& (\mathbf{x} - \mathbf{x}_1)^T\mathbf{C}_1^{-1}(\mathbf{x} - \mathbf{x}_1) + (\mathbf{x} - \mathbf{x}_2)^T\mathbf{C}_2^{-1}(\mathbf{x} - \mathbf{x}_2)\end{aligned} \tag{7}$$

Taking derivative with respect to $\mathbf{x}$ and setting it to zero, we get the best linear unbiased estimate (BLUE) of $\mathbf{x}$ ([48], [49]):

$$\mathbf{x}^* = \mathbf{C}(\mathbf{C}_1^{-1}\mathbf{x}_1 + \mathbf{C}_2^{-1}\mathbf{x}_2) \tag{8}$$

$$\mathbf{C} = (\mathbf{C}_1^{-1} + \mathbf{C}_2^{-1})^{-1} \tag{9}$$

Now, assume that one of the Gaussians is in a subspace of dimension p, e.g., $\mathbf{C}_2$ is singular. With the singular value decomposition of $\mathbf{C}_2 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$, with $\mathbf{u}_i$'s orthonormal and $\mathbf{\Lambda} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_p, 0, \ldots, 0\}$, we rewrite Mahalanobis distance to $\mathbf{x}_2$ in Eq. (7) in the canonical form:

$$\begin{aligned}D^2(\mathbf{x}, \mathbf{x}_2, \mathbf{C}_2) &= (\mathbf{x} - \mathbf{x}_2)^T\mathbf{C}_2^{-1}(\mathbf{x} - \mathbf{x}_2) \\ &= \sum_{i=1}^{n}\lambda_i^{-1}[\mathbf{U}^T(\mathbf{x} - \mathbf{x}_2)]^2\end{aligned} \tag{10}$$

When λ_i tends to 0, $D^2(\mathbf{x}, \mathbf{x}_2, \mathbf{C}_2)$ goes to infinity, unless $\mathbf{U}_0^T\mathbf{x} = 0$, where $\mathbf{U}_0 = [\mathbf{u}_{p+1}, \mathbf{u}_{p+2}, \ldots, \mathbf{u}_n]$. Here we have assumed, without loss of generality, that the subspace passes through the origin of the original space. Since $\mathbf{x}_2$ resides in the subspace, ${\mathbf{U}_0}^T\mathbf{x}_2 = \mathbf{0}$. As a result, we only need to retain in Eq. (10) those terms corresponding to a non-zero λ_i:

$$d_{m,2} = \sum_{i=1}^{p}\lambda_i^{-1}[\mathbf{U}_p^T(\mathbf{x} - \mathbf{x}_2)]^2 \equiv (\mathbf{x} - \mathbf{x}_2)^T\mathbf{C}_2^{+}(\mathbf{x} - \mathbf{x}_2) \tag{11}$$

where $\mathbf{C}_2^{+}$ is the pseudoinverse of $\mathbf{C}_2$([50], p. 49), and $\mathbf{U}_p = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p]$.

Taking into account that ${\mathbf{U}_0}^T\mathbf{x} = \mathbf{0}$, $\mathbf{x}$ can be expressed in another form to reflect this constraint:

$$\mathbf{x} = \mathbf{U}\mathbf{U}^T\mathbf{x} = \mathbf{U}([\mathbf{U}_p|\mathbf{U}_0]^T\mathbf{x}) = \mathbf{U}\begin{bmatrix}\mathbf{y}\\ \mathbf{0}\end{bmatrix} = \mathbf{U}_p\mathbf{y} \tag{12}$$

where $\mathbf{y}$ is a $1\times p$ vector. Eq. (7) now takes the following general form:

$$\begin{aligned} d^2 =&(\mathbf{U}_p\mathbf{y}-\mathbf{x}_1)^T\mathbf{C}_1^{-1}(\mathbf{U}_p\mathbf{y}-\mathbf{x}_1)+ \\ &(\mathbf{U}_p\mathbf{y}-\mathbf{x}_2)^T\mathbf{C}_2^{+}(\mathbf{U}_p\mathbf{y}-\mathbf{x}_2) \end{aligned} \tag{13}$$

Taking derivative with respect to $\mathbf{y}$ yields the fusion estimator for the subspace:

$$\mathbf{y}^* = \mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T(\mathbf{C}_1^{-1}\mathbf{x}_1+\mathbf{C}_2^{+}\mathbf{x}_2) \tag{14}$$

$$\mathbf{C}_{\mathbf{y}^*} = [\mathbf{U}_p^T(\mathbf{C}_1^{-1}+\mathbf{C}_2^{+})\mathbf{U}_p]^{-1} \tag{15}$$

Equivalent expressions can be obtained in the original space:

$$\mathbf{x}^* = \mathbf{U}_p\mathbf{y}^* = \mathbf{C}_{\mathbf{x}^*}(\mathbf{C}_1^{-1}\mathbf{x}_1+\mathbf{C}_2^{+}\mathbf{x}_2) \tag{16}$$

$$\mathbf{C}_{\mathbf{x}^*} = \mathbf{U}_p\mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T \tag{17}$$

It can be shown that $\mathbf{C}_{\mathbf{x}^*}$ and $\mathbf{C}_{\mathbf{y}^*}$ are the corresponding covariance matrices for $\mathbf{x}^*$ and $\mathbf{y}^*$ (see Appendix). Notice that this solution is not a simple generalization of Eq. (8) by substituting pseudoinverses for regular inverses, which will not constrain $\mathbf{x}^*$ to be in the subspace.

When only the subspace constraint is valid, without assuming any model distribution in the subspace, the solution becomes (by setting λ_i to ∞):

$$\mathbf{y}^* = \mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T\mathbf{C}_1^{-1}\mathbf{x}_1 \tag{18}$$

$$\mathbf{C}_{\mathbf{y}^*} = (\mathbf{U}_p^T\mathbf{C}_1^{-1}\mathbf{U}_p)^{-1} \tag{19}$$

When only the subspace (defined by $\mathbf{U}_p$) is considered, and the covariance $\mathbf{C}_1 = c\mathbf{I}$ is isotropic, we have

$$\mathbf{y}^* = (c^{-1}\mathbf{U}_p^T\mathbf{I}\mathbf{U}_p)^{-1}\mathbf{U}_p^T(c\mathbf{I})^{-1}\mathbf{x}_1 = \mathbf{U}_p^T\mathbf{x}_1 \tag{20}$$

$$\mathbf{C}_{\mathbf{y}^*} = c\mathbf{I}_p \tag{21}$$

This is the orthogonal projection used in the literature as subspace model constraint [24], [28]

Figure 2 depicts the difference among three subspace constraining approaches, namely, *orthogonal projection* (Eq. (20)), *non-orthogonal projection* (Eq. (18)), and *non-orthogonal projection with fusion* (Eq. (14)), synthetically using a three-dimensional feature space with a two-dimensional subspace model.

C. Tracking Shapes through Subspace Fusion

Applying the subspace fusion formula, Eq. (14) and (15), on the Kalman fusion result of Eq. (5) (In general $\mathbf{H}$ is not invertible; otherwise, Eq. (6) can be used.) and a subspace Gaussian source $\mathcal{N}(\mathbf{x}_2, \mathbf{C}_2)$, we obtain a complete fusion framework:

$$\begin{aligned} &\mathbf{x}_{k+1|k+1} = \\ &\mathbf{P}_{k+1|k+1}((\mathbf{S}\mathbf{P}_{k|k}\mathbf{S}^T+\mathbf{Q})^{+}\mathbf{x}_{k+1|k}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}_{k+1}+\mathbf{C}_2^{+}\mathbf{x}_2) \end{aligned} \tag{22}$$

$$\begin{aligned} &\mathbf{P}_{k+1|k+1} = \\ &\mathbf{U}_p[\mathbf{U}_p^T((\mathbf{S}\mathbf{P}_{k|k}\mathbf{S}^T+\mathbf{Q})^{+}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}+\mathbf{C}_2^{+})\mathbf{U}_p]^{-1}\mathbf{U}_p^T \end{aligned} \tag{23}$$

where we have assumed, without loss of generality, that the system noise covariance $\mathbf{Q}$ can be contained in the subspace.

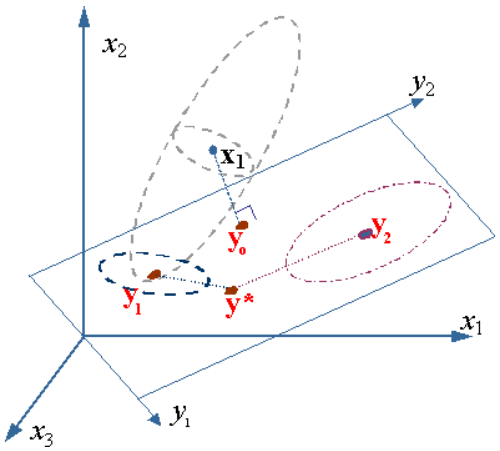


Fig. 2. Comparing *orthogonal projection* ($\mathbf{y_o}$), *non-orthogonal projection* ($\mathbf{y}_1$), and *non-orthogonal projection with fusion* ($\mathbf{y}^*$). Here $\mathbf{x}_1$ is the input vector with a covariance matrix, $\mathbf{y}_2$ is the model mean in the subspace.

Observe the symmetry of the solution which *combines all the available knowledge in the information space*. These equations provide a unified fusion of the system dynamics, a subspace model, and measurement noise information. They completely represent the various uncertainties that affect the tracking system.

Compared to a PCA *shape space* representation [30], [22], the above formulation uses not only the model subspace (the eigenvectors), but also the actual model distribution, in a unified fusion framework. On the other hand, if only a subspace constraint is desired, we can simply apply the special case of Eq. (18) and (19) on Eq. (5), and the resulting fusion is still within the same analytical framework. The traditional subspace constraining technique of orthogonal projection into the eigenspace will discard information carried in the heteroscedastic noise of the measurements.

IV. Strongly-Adapted-PCA Model

For endocardium (the inner layer of the heart muscle) tracking, the use of a statistical shape model learned from a large pool of training samples to guide the contours from a specific heart is problematic. Theoretically, what we really need is *the statistical shape model of the current heart* instead of a generic heart. Therefore, there is a strong motivation for us to adapt the generic model toward what is known for the current case. A natural choice is to use the initial contour (manual or through automatic detection) to adapt the existing PCA model.

Naturally, incremental PCA (IPCA) [51] comes into the picture, but this strategy is not strongly adapted toward the current case and is often too weak (we will elaborate more in the sequel). We apply a strongly-adapted-PCA (SA-PCA) model as follows:

We assume that the PCA model (excluding the current case) and the initialized contour for the current case *jointly* represent the variations of the current case, but with relative energy, or representative power, being α and $(1-\alpha)$, respectively, with $0<\alpha<1$. In other words, a portion of the shape variations of the current case is captured by the generic model, while

the rest is captured *in the direction of* the initial contour in the model space.

The PCA model has its mean, eigenvalue matrix, and eigenvector matrix, denoted by $\mathbf{x}_m$, $\mathbf{\Lambda}$, and $\mathbf{U}$, respectively. If the original full covariance matrix $\mathbf{C}$ is stored (this would be the case when the original dimensionality is not forbiddingly high), the adapted mean and covariance matrix are simply the weighted sum of the two contributing sources:

$$\mathbf{x}_{m,new} = \alpha\mathbf{x}_m + (1-\alpha)\mathbf{x} \tag{24}$$

$$\begin{aligned}\mathbf{C}_{new} =& \alpha(\mathbf{C} + (\mathbf{x}_m - \mathbf{x}_{m,new})(\mathbf{x}_m - \mathbf{x}_{m,new})^T) \\ &+ (1-\alpha)(\mathbf{x} - \mathbf{x}_{m,new})(\mathbf{x} - \mathbf{x}_{m,new})^T \\ =& \alpha\mathbf{C} + \alpha(1-\alpha)(\mathbf{x} - \mathbf{x}_m)(\mathbf{x} - \mathbf{x}_m)^T\end{aligned} \tag{25}$$

Eigenanalysis can be performed on $\mathbf{C}_{new}$ to obtain the new subspace model.

In the case that $\mathbf{C}$ is not stored but only $\{\mathbf{x}_m, \mathbf{\Lambda}, \mathbf{U}\}$ is available *in the subspace*, through straight algebraic manipulations we can arrive at the adapted eigenanalysis results $\{\mathbf{x}_{m,new}, \mathbf{\Lambda}_{new}, \mathbf{U}_{new}\}$ as follows: The initial contour $\mathbf{x}$ has a subspace component as $\mathbf{x}_s = \mathbf{U}^T\mathbf{x}_d$, where $\mathbf{x}_d = \mathbf{x} - \mathbf{x}_m$, and a residual vector as $\mathbf{x}_r = (\mathbf{x} - \mathbf{x}_m) - \mathbf{U}\mathbf{x}_s$. Let $\mathbf{x}_{ru}$ be the normalized version of $\mathbf{x}_r$ with norm 1 (or zero if $\mathbf{x}_r$ is zero norm). The adapted eigenvector matrix that represents the combined energy will have the following form:

$$\mathbf{U}_{new} = [\mathbf{U}, \mathbf{x}_{ru}]\mathbf{R} \tag{26}$$

$\mathbf{R}$ and $\mathbf{\Lambda}_{new}$ will be the solutions to the following eigenanalysis problem:

$$\left(\alpha \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} + \alpha(1-\alpha) \begin{bmatrix} \mathbf{x}_s\mathbf{x}_s^T & e_r\mathbf{x}_s \\ e_r\mathbf{x}_s^T & e_r^2 \end{bmatrix}\right)\mathbf{R} = \mathbf{R}\mathbf{\Lambda}_{new} \tag{27}$$

where $e_r = \mathbf{x}_{ru}^T(\mathbf{x} - \mathbf{x}_m)$ is the residual energy.

The above formulas are extensions of IPCA or eigenspace merging formula of [51], with tunable energy ratios between the new data and the old data. These become equivalent to IPCA if we set α to be the fraction of points in the model versus the total number of points. Typically, this will be a number very close to 1 since the number of contours in the training set is usually large. With α set at a smaller value (we use 0.5), the PCA model is strongly adapted toward the current case, hence the name.

Contours from the current heart are more likely to resemble the initial contour of the same heart than those contours in the generic training set, especially if the current heart has an irregular shape that is not represented in the training set. Figure 3 shows a comparison of IPCA and SA-PCA. This parasternal short axis view has an irregular shape (with an upper concave part). The incremental PCA model, taking in the initial contour (Figure 3a) but with a very small weight ($< 0.01\%$), fails to capture the concave nature of the current shape; and has constrained the contours to a typical circular shape (Figure 3b). SA-PCA yields a contour that fits much better to the true borders (Figure 3c).

With SA-PCA, our framework now incorporates four information sources: the system dynamic, measurement, subspace model, and the initial contour. This last addition is especially useful for periodic shape deformations such as the cardiac motion.

V. Estimating Motion

To measure the motion of each of the control points we use an adaptation of the frame-to-frame motion estimation algorithm described in [26], which has been shown to be very competitive in terms of performance evaluation using standard sequences. We present in the sequel a summary of the algorithm. For more details, see [26].

The main idea is that the *motion in a certain neighborhood can be robustly estimated as the most significant mode of some initial motion estimates*, expressed by mean vectors and associated covariance matrices. The most significant mode of the underlying multimodal density function of initial estimates is detected by mode tracking across scales. We first perform mode detection at a large scale, while in the next stages the analysis scale is gradually reduced and the mode detection is performed again. At each scale the mode detection algorithm is initialized with the convergence location from the previous scale. The underlying mechanism for mode detection relies on the variable-bandwidth mean shift [52].

To be more explicit, assume a constant velocity model and denote by $(\hat{\mathbf{z}}_i, \mathbf{R}_i)$, $i = 1 \ldots n$ the initial motion estimates computed at each spatial coordinate of a neighborhood of dimension $n = M \times M$. The initial estimates are derived through the traditional Least Squares [53] with some upper bound on the covariance eigenvalues to avoid instabilities, or through Biased Least Squares [54], [55].

Using the robust fusion framework, we determine the location $\hat{\mathbf{z}}_m$ of the most significant mode in the 2-dimensional velocity space by performing mode tracking across scales. At the mode location, the following equations are valid

$$\hat{\mathbf{z}}_m = \mathbf{R}(\hat{\mathbf{z}}_m)\sum_{i=1}^{n}\omega_i(\hat{\mathbf{z}}_m)\mathbf{R}_i^{-1}\hat{\mathbf{z}}_i \tag{28}$$

$$\mathbf{R}(\hat{\mathbf{z}}_m) = \left(\sum_{i=1}^{n}\omega_i(\hat{\mathbf{z}}_m)\mathbf{R}_i^{-1}\right)^{-1} \tag{29}$$

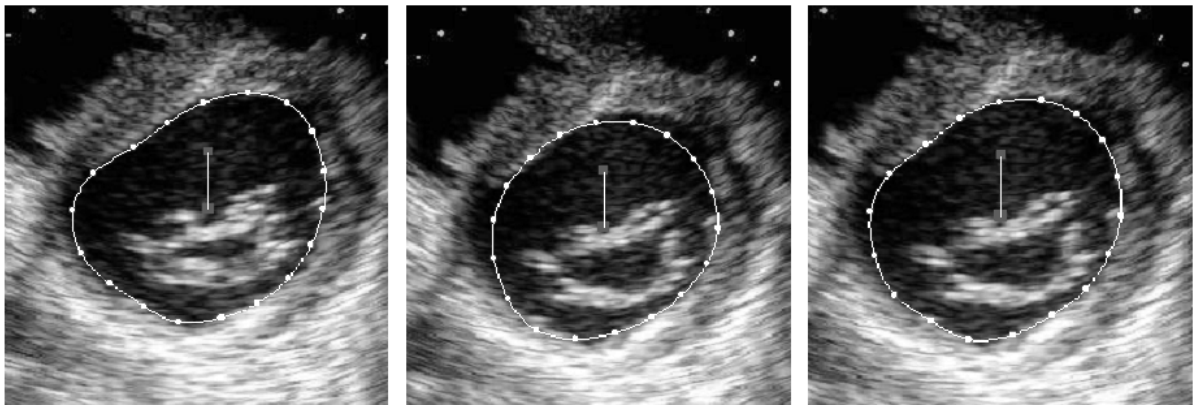
where the weights are given by

$$\omega_i(\hat{\mathbf{z}}_m) = \frac{\frac{1}{|\mathbf{R}_i|^{1/2}}\exp\left(-\frac{1}{2}D^2\left(\hat{\mathbf{z}}_m, \hat{\mathbf{z}}_i, \mathbf{R}_i\right)\right)}{\sum_{i=1}^{n}\frac{1}{|\mathbf{R}_i|^{1/2}}\exp\left(-\frac{1}{2}D^2\left(\hat{\mathbf{z}}_m, \hat{\mathbf{z}}_i, \mathbf{R}_i\right)\right)} \tag{30}$$

Equations (28) and (29) define the estimate of the local motion vector. The procedure is robust and can accommodate multiple apparent motions, since it relies on the multiscale computation of the mode of initial estimates.

In the current work, for each control point we compute initial estimates using 17×17 windows and fuse the results on $n = 5 \times 5$ neighborhoods. A pyramid of three levels is employed with covariance propagation across levels. Figure 4 depicts the uncertainty calculated at the bottom of the pyramid for the contour points.

To avoid error accumulation from frame to frame, the motion is always computed with reference to the neighborhood of the control point in the first frame of the sequence (i.e., the



(a) (b) (c)

Fig. 3. SAPCA versus incremental PCA. (a) is the initial contour; (b) is the 5th contour tracked and constrained using an incremental PCA model [51]; (c) is the same frame tracked and constrained using an adaptive PCA model with $\alpha = 0.5$.

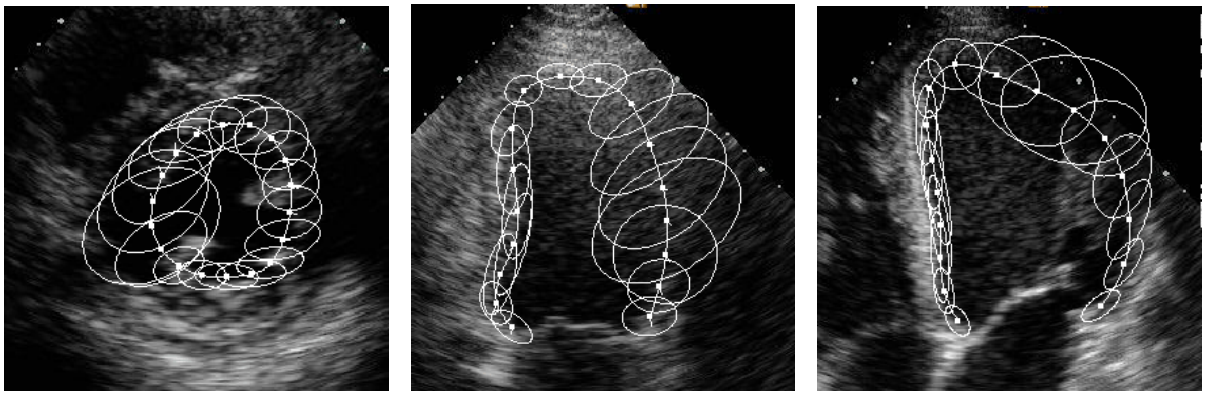


Fig. 4. The 95% confidence ellipses corresponding to the local measurement uncertainty. Frames sampled from Case #6, #25, and #32.

current frame is always compared to a model extracted from the first frame). Since we update the location of the model at each frame, the motion estimation process always starts with a good initialization. The overall procedure is suitable for the tracking of periodic sequences such as the heart ultrasound data. It resembles a template-based tracker that benefits from the fast computation of frame-to-frame motion.

VI. Experiments and Analysis

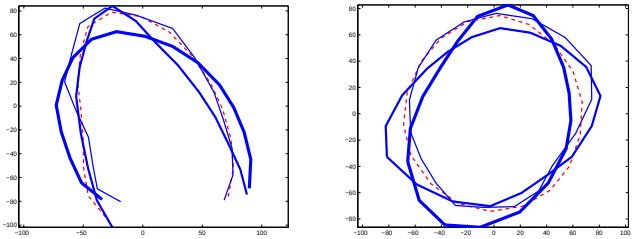
In this section we will apply and evaluate the new framework to track heart contours using very noisy echocardiography data. The tracker was implemented in C++ and is running at about 20 frames per second on a single 2GHz Pentium 4 PC. Our data were selected by a cardiologist to represent normals as well as various types and stages of cardiomyopathies, with sequences varying in length from 18 frames to 90 frames. All data were collected by Siemens Acuson machines (Sequoia model), using tissue harmonic imaging. The images are through all parts of the heart cycle. Both training and test data were traced by experts, and confirmed by one cardiologist.

A. Implementation

We use manually traced contours of the endocardium in echocardiography images as the training set. Both apical two- or four-chamber views (open contour with 17 control points) and parasternal short axis views (closed contour with 18 control points) are trained and tested. Apical and short axis views are trained separately using contours outside the test set. The tracing is done by either adjusting generic contour templates or drawing a new contour using the mouse. In the latter case, the contour is drawn with a pre-defined starting point based on anatomy so that landmark points can be automatically assigned. Typical landmarks are the apex, the papillary muscles, and the basal end of the septum. We number the control points in clockwise order, with the first point from the mid inferior region for the short axis views and from the basal septal/inferior region for apical views. Furthermore, the mid-point (the 9th point) for the apical view is assigned at the apex. Our system automatically places the other points equally on the contour based on arc length of the fitted spline.

The algorithm can tolerate some variability on the location of the landmark points, partly due to the application of SA-PCA.

The training contours are aligned using the iterative Procrustes analysis approach proposed by Cootes and Taylor [19] to cancel out global translation, rotation and scaling. PCA is then performed and the original dimensionality of 34 and 36 is reduced to 7 and 8, respectively. Figure 5 shows the dominant eigenshapes (without splining) for the two views along with



(a) (b)

Fig. 5. The dominant eigenshapes for: (a) Apical views; (b) Short axis views. The dashed curve is the model mean.

their model means trained on about 200 contours for each view. The training set consists of sequences from 11 subjects, some of which have different degrees of cardiomyopathy.

In this paper we use a simple 0th order dynamic model to impose a temporal smoothness constraint [1]. We employ a diagonal matrix to model the uncertainty in system dynamics, and set the relative confidence of this model empirically .

In practice, we perform alignment on the training shapes before the PCA, hence, at each frame the learned shape model, $\{\mathbf{x}_2, \mathbf{C}_2\}$ in Section III-B, needs to be aligned to the fusion result $\{\tilde{\mathbf{x}}, \tilde{\mathbf{C}}\}$ from the measurement and the dynamic model. We adopt the optimal transformation $\mathcal{T}_o$ which minimizes a weighted sum-of-squares measure of point difference subject to translation, rotation and scaling ([56], p. 102), with the weighting matrix being $\tilde{\mathbf{C}}^{-1}$. The system transforms the model mean as well as *the model covariance* using $\mathcal{T}_o$ before the final fusion.

B. Echocardiogram Tracking

For systematic evaluation, we use a set of 32 echocardiogram sequences outside of the training set for testing, with 18 parasternal short-axis (PS) views and 14 apical two- or four-chamber (AC) views, all with expert-annotated ground-truth contours.

Figure 6 shows snapshots from two tracked sequences. Notice that the endocardium is not always on the strongest edge. Sometimes it manifests itself only by a faint line; sometimes it is completely invisible or buried in heavy noise; sometimes it will cut through the root of the papillary muscles where no edge is present. Our multiscale measurement process follows whatever patterns the initial contour is attached to, with varying levels of confidence. This is shown in Figure 4. Notice the strong anisotropy and inhomogeneity associated with the measurement uncertainty. The full exploitation of this information resulted in robust tracking over high noise and drop-out regions. Figure 7 shows frames from four sequences (first row), along with ground-truth contours (second row), and our tracking results (last row). The difference is barely noticeable even in the drop-out regions.

When the measurement process makes a large error in a drop-out or high-noise region, the corresponding localization uncertainty is usually higher as well, because of the lack of trackable patterns. In such cases, our fusion will correct the error to a larger extent than the effect of an orthogonal projection into the shape space. An example is shown in Figure 8. The fusion will not correct the error completely, but note that this correction step is accumulative so that the overall effect at a later frame in a long sequence can be very significant.

To compare performance of different methods, we need proper distance measures. We use both a Mean Sum of Squared Distance (MSSD) (*cf.* [45]) and a Mean Absolute Distance (MAD) (*cf.* [43]). For the tracked sequence S_i with m frames/contours, $\{c_1, c_2, ..., c_m\}$, where each contour c_j has n points $\{(x_{j,1}, y_{j,1}), (x_{j,2}, y_{j,2}), ..., (x_{j,n}, y_{j,n})\}$, the distances of S_i from the ground truth sequence S_i^0 are

$$\begin{aligned} MSSD_i &= \frac{1}{m}\sum_{j=1}^{m} MSSD_{i,j} \\ &= \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{k=1}^{n}((x_{j,k}-x_{j,k}^0)^2+(y_{j,k}-y_{j,k}^0)^2) \end{aligned} \tag{31}$$

$$\begin{aligned} MAD_i &= \frac{1}{m}\sum_{j=1}^{m} MAD_{i,j} \\ &= \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{k=1}^{n}\sqrt{(x_{j,k}-x_{j,k}^0)^2+(y_{j,k}-y_{j,k}^0)^2} \end{aligned} \tag{32}$$

The overall performance measure for a particular method is the averaged distance on the whole test set of l sequences: $MSSD = \frac{1}{l}\sum_{i=1}^{l} MSSD_i$ and $MAD = \frac{1}{l}\sum_{i=1}^{l} MAD_i$. These measures do *not* give higher weights to longer sequences. We also record the standard deviation of the two distances for each frame and average the results for each sequence first, and then across sequences. A critical difference between our distance measures and those of [45] or [43] is that we have the point correspondence through tracking, and we want to measure the performance of these correspondences. In [45], [43], no correspondence is assumed and the nearest point from the other contour is taking as the corresponding point–As a result, motion component along the tangent of the contour *cannot* be evaluated. In reality, global or regional tangent motion are common during a cardiac cycle, and they reveal crucial information regarding cardiac function.

With the same MADs, a contour with impulsive errors has a higher MSSD than one with uniform errors. We test the performance of our approach under both measures. Our fusion framework is compared to three methods. The first method is a tracker based on our optical flow algorithm without shape constraint ("Flow") [26]. The second method is the same tracker but adding orthogonal PCA shape space constraints ("FlowShapeSpace") (*cf.* [31], [19], [23]). The third is "FlowShapeSpace" using our SA-PCA model ("FlowSAPCA")–we show the best empirical performance by tuning α values. Figure 9, 10 and Table I show the comparison using the two distance measures. Our proposed method ("Proposed") significantly outperforms others, with lower average distances and lower standard deviations for such distances. Figure 12

[1]Possible extensions include the constant-velocity model or sinusoidal dynamic model which captures the periodic expansion and contraction of the heart

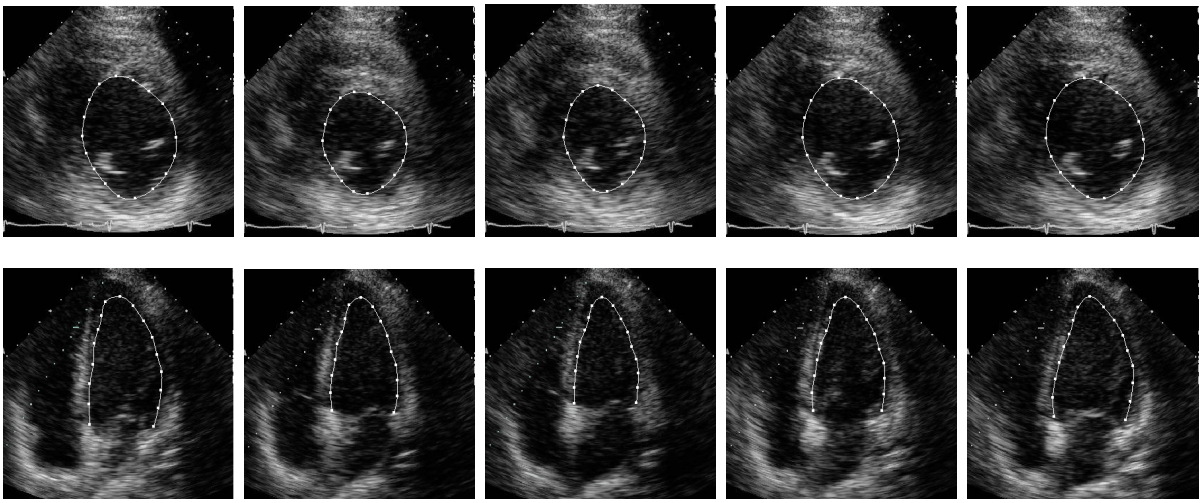


Fig. 6. Two tracking examples in rows, with 5 snapshots per sequence. Shown on the first row are Frame 1, 10, 15, 20 and 25 from Case #24 with 25 frames, and on the second row are Frame 1, 8, 12, 16 and 20 from Case #17 with 20 frames.

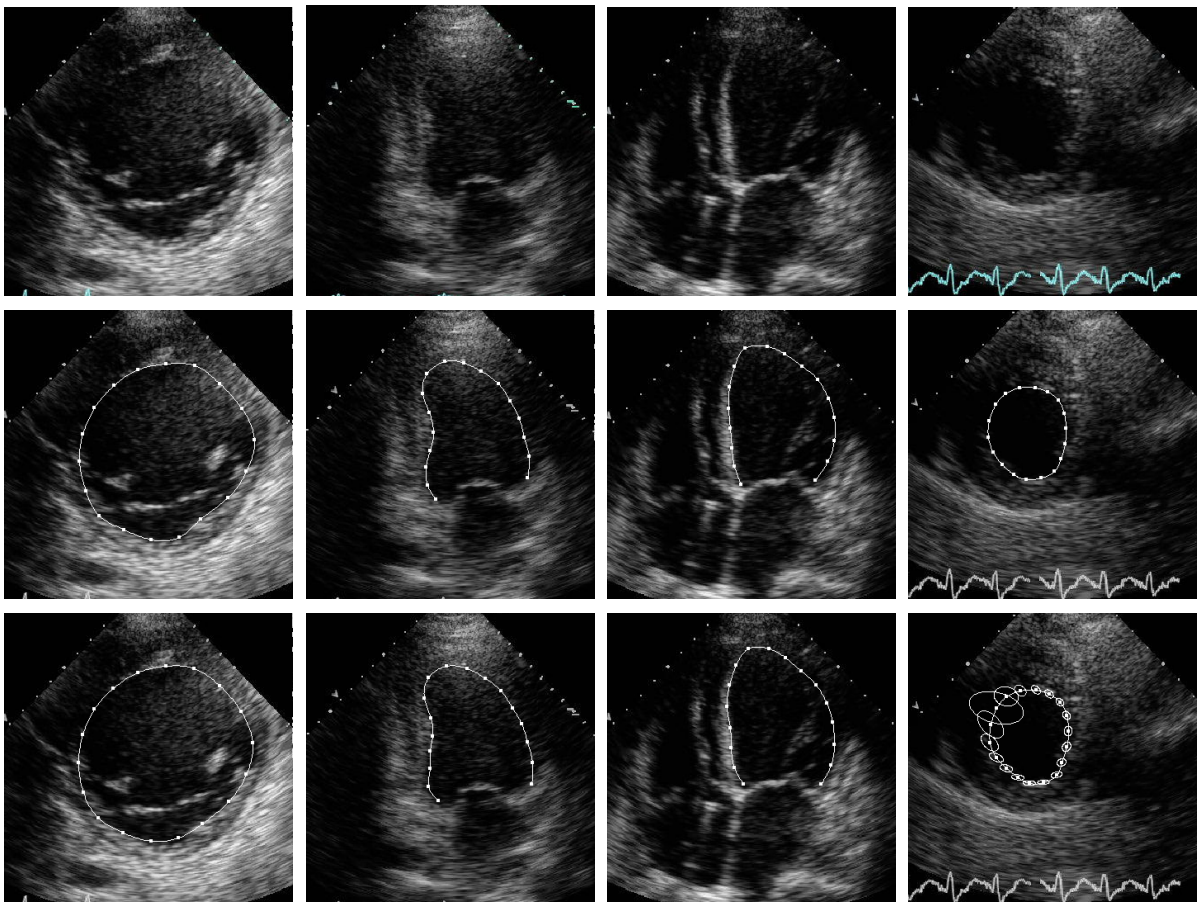
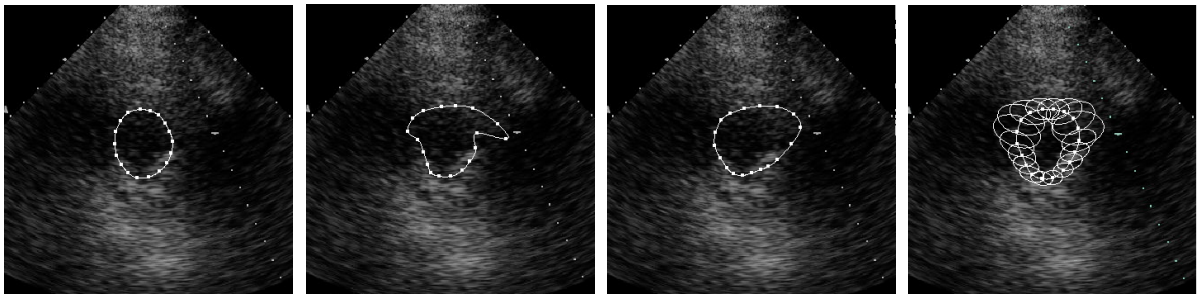


Fig. 7. Four more tracking examples in columns. First row shows the images, second row are the expert-drawn contours for the image directly above, and the third row shows our tracking results. On the last contour we also show the uncertainty ellipses.



(a) (b) (c) (d)

Fig. 8. Comparison of orthogonal projection with our proposed fusion approach in handling measurement errors (Frame 15 from Case #11). (a) the expert-drawn contour; (b) the un-constrained flow results; (c) the constrained flow using orthogonal projection into the SA-PCA eigenshape space; and (d) the contour by our fusion framework along with the instantaneous uncertainty ellipses used during the fusion. Notice the strong correction our method brought over the local measurement errors.

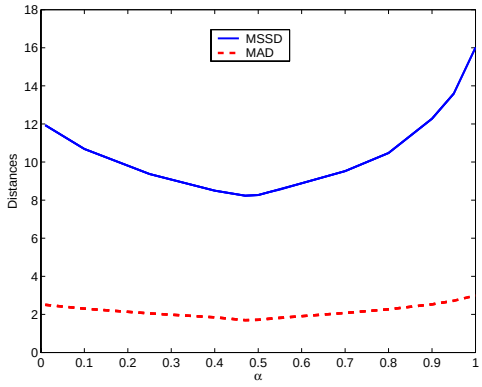


Fig. 12. The sensitivity of the proposed method to different values of α, which is the energy ratio of the original PCA model in the SA-PCA model.

shows the sensitivity of the proposed approach regarding the values of α, which is the energy ratio of the original PCA model in the SA-PCA model. It shows that for a wide range of α values, the proposed framework outperforms the other methods. The optimal value for α depends on the datasets and is 0.49 in our case.

These distances are based on the unit of pixels displayed on the screen. As references for understanding these results: the full echocardiogram image is displayed at a size of about 640 $\times$ 480 pixels. It should be noted that our results are not indicative for *border localization* accuracies, but rather for *motion tracking* performances given an initial contour.

In Figure 11, we show the averaged distances (in terms of MSSD) for different control points for the two views. Notice that the errors are not uniform across control points, especially those from the "Flow" approach: for apical views, the high-error regions are along the lateral wall and around the apex; while for short axis views, errors are larger on the upper left and upper right around mid septal and mid anterior regions. This is in agreement with the typical drop-out regions in these two views (*cf.* Figure 4 and 8). The proposed approach corrected such large errors completely.

It is worth pointing out that our proposed method performs relatively well even on the noisiest cases. (See the last four columns in Table I).

VII. Discussions

We have set our goal to track control points on the endocardium, with anisotropic confidence estimated at each point at any given time step, and exploit this information when consulting a prior shape model as a constraint. However, we should point out that this goal may not be achievable for some sequences due to the nature of echocardiography and the way it is taken: The heart is moving in 3D while 2D echocardiography captures only a slice of the heart which may change from frame to frame. To make things worse, the heart also has a small twisting motion around the apex during systole. Therefore, our goal of tracking point correspondences can only be achieved to a certain degree. In our dataset, the 3D motion of the heart is not significant and the tracking is reasonable for most cases..

VIII. Conclusions and Future Work

This paper presented a joint information fusion framework to track shapes under heteroscedastic noise with a probabilistic subspace model constraint. The major contributions of our work are three-fold: first, we formulate the tracking framework as an information fusion problem to take into account full information from the subspace model, where the model mean is retained; secondly, we estimate and exploit uncertainties from the optical-flow during the shape model constraining process, which weighs more on flows with high confidence while discounts flows estimated from uncertain areas such as drop-out regions; finally, we propose a model adaptation technique that uses the initial contour to update a PCA model, so that a generic shape model can be applied to constrain the tracking of myocardial borders of a specific shape. This approach has been shown to perform robustly even on difficult ultrasound heart images with spatially varying noise.

Our framework is general and can be applied to other modalities. Future potential applications include tracking in MR, perfusion, and extensions to tracking using 3-D or 4-D data.

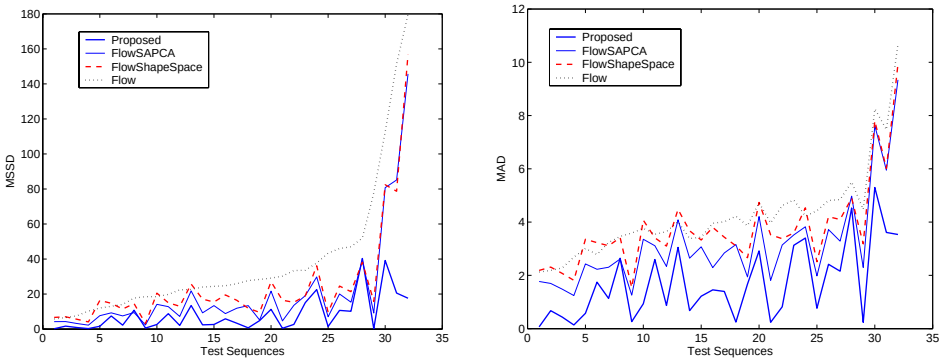


Fig. 9. Mean distances between tracked points and the ground truth location for 32 test sequences. (a) Mean Sum of Squared Distances (MSSD_i); (b) Mean Absolute Distance (MAD_i). The test sequences are sorted by the ascending order of the MSSD of "Flow".

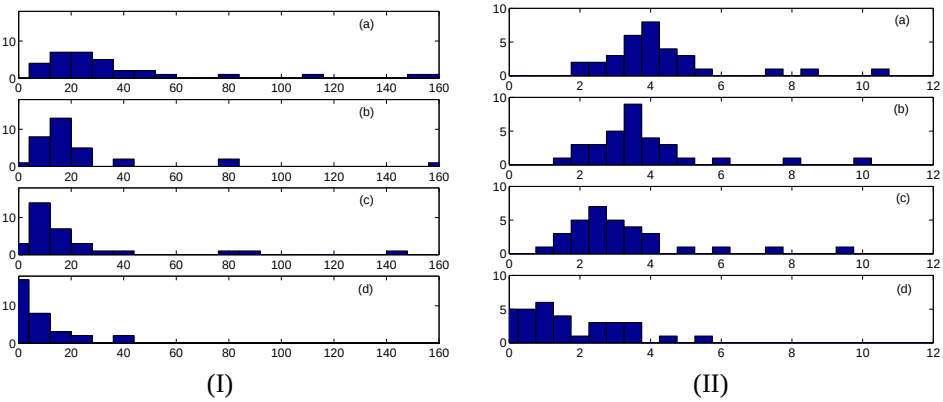


Fig. 10. The error histograms based on (I) MSSD, and (II) MAD, for the four approaches: (a) Flow; (b) FlowShapeSpace; (c) FlowSAPCA; and (d) Proposed.

TABLE I

PERFORMANCE OF PROPOSED METHOD. THE "MOST DIFFICULT CASES" CONSIST OF THE LAST THREE CASES IN FIGURE 9A. THESE ARE THE MOST DIFFICULT CASES FOR "FLOW".)

	All Cases				*Most Difficult Cases*			
Methods	*MSSD*	$\bar{\sigma}_{MSSD}$	*MAD*	$\bar{\sigma}_{MAD}$	*MSSD*	$\bar{\sigma}_{MSSD}$	*MAD*	$\bar{\sigma}_{MAD}$
Flow	38.1	82.9	4.3	3.6	147.9	325.0	8.8	8.2
FlowShapeSpace	24.7	35.5	3.8	2.4	106.0	181.2	7.9	6.3
FlowSAPCA	17.5	21.7	3.1	1.9	78.8	98.6	7.1	4.7
Proposed	8.3	14.3	1.7	1.6	25.8	34.8	4.1	2.8

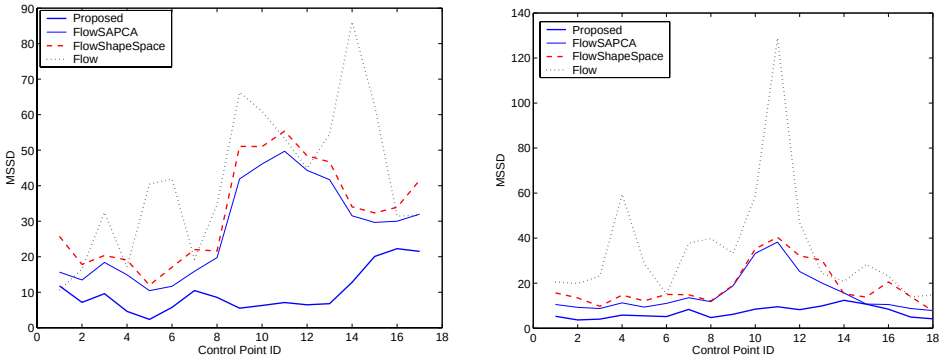


Fig. 11. Averaged MSSD for different control points. (a) Apical 2- or 4-chamber views; (b) Parasternal short axis views. It shows that for apical views, the high-error regions are along the lateral wall and around the apex; while for short axis views, errors are larger on the left and right around mid septal and mid anterior regions.

APPENDIX I
KALMAN FILTER AS INFORMATION FILTER

We first rewrite $\mathbf{P}_{k+1|k+1}$ in Eq. (3) as follows:

$$\begin{aligned}&\mathbf{P}_{k+1|k+1} = (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k}\\ =&(\mathbf{P}_{k+1|k}-\mathbf{P}_{k+1|k}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T+\mathbf{R})^{-1}\mathbf{H}\mathbf{P}_{k+1|k})\end{aligned} \tag{A.1}$$

It can be shown that

$$\mathbf{P}^{-1}_{k+1|k+1} = \mathbf{P}^{-1}_{k+1|k} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{A.2}$$

because (we drop the subscripts for simplicity):

$$\begin{aligned}&[\mathbf{P}-\mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T+\mathbf{R})^{-1}\mathbf{H}\mathbf{P}][\mathbf{P}^{-1}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}]\\ &=\mathbf{I}+\mathbf{P}\mathbf{H}^T[\mathbf{R}^{-1}-(\mathbf{H}\mathbf{P}\mathbf{H}^T+\mathbf{R})^{-1}(\mathbf{I}+\mathbf{H}\mathbf{P}\mathbf{H}^T\mathbf{R}^{-1})]\mathbf{H}\\ &=\mathbf{I}\end{aligned} \tag{A.3}$$

We can now rewrite $\mathbf{x}_{k+1|k+1}$ of Eq. (1):

$$\begin{aligned}&\mathbf{x}_{k+1|k+1}\\ &=(\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k}(\mathbf{P}^{-1}_{k+1|k}\mathbf{x}_{k+1|k}+\\ &\quad\mathbf{P}^{-1}_{k+1|k}(\mathbf{I}-\mathbf{K}\mathbf{H})^{-1}\mathbf{K}\mathbf{z}_{k+1})\\ &=\mathbf{P}_{k+1|k+1}(\mathbf{P}^{-1}_{k+1|k}\mathbf{x}_{k+1|k}+\underbrace{\mathbf{P}^{-1}_{k+1|k}(\mathbf{I}-\mathbf{K}\mathbf{H})^{-1}\mathbf{K}}\mathbf{z}_{k+1})\end{aligned} \tag{A.4}$$

The under-braced term in the parenthesis can be simplified as follows:

$$\begin{aligned}&\mathbf{P}^{-1}_{k+1|k}(\mathbf{I}-\mathbf{K}\mathbf{H})^{-1}\mathbf{K}\\ &=\left[(\mathbf{K}^{-1}\mathbf{P}_{k+1|k}-\mathbf{H}\mathbf{P}_{k+1|k}\right]^{-1}\\ &=\left[[\mathbf{P}_{k+1|k}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T+\mathbf{R})^{-1}]^{-1}\mathbf{P}_{k+1|k}-\mathbf{H}\mathbf{P}_{k+1|k}\right]^{-1}\\ &=\mathbf{H}^T\mathbf{R}^{-1}\end{aligned} \tag{A.5}$$

Finally we obtain:

$$\mathbf{x}_{k+1|k+1} = \mathbf{P}_{k+1|k+1}(\mathbf{P}^{-1}_{k+1|k}\mathbf{x}_{k+1|k}+\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}_{k+1}) \tag{A.6}$$

APPENDIX II
PROOF THAT EQ. (15) AND (17) ARE COVARIANCE MATRICES

According to Eq. (14), we estimate the covariance of $\mathbf{y}^*$:

$$\begin{aligned}&E[\mathbf{y}^*\mathbf{y}^{*T}] =\\ &\mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T(\mathbf{C}_1^{-1}E[\mathbf{x}_1\mathbf{x}_1^T]\mathbf{C}_1^{-1}+\mathbf{C}_2^{+}E[\mathbf{x}_2\mathbf{x}_2^T]\mathbf{C}_2^{+})\mathbf{U}_p\mathbf{C}_{\mathbf{y}^*}\end{aligned} \tag{B.1}$$

Since

$$E[\mathbf{x}_i\mathbf{x}_i^T] = \mathbf{C}_i, \tag{B.2}$$

$$\mathbf{C}_2^{+}\mathbf{C}_2 = \mathbf{U}_p\mathbf{U}_p^T, \tag{B.3}$$

$$\mathbf{U}_p^T\mathbf{U}_p = \mathbf{I} \tag{B.4}$$

We have

$$\begin{aligned}E[\mathbf{y}^*\mathbf{y}^{*T}] =&\mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T(\mathbf{C}_1^{-1}+\mathbf{U}_p\mathbf{U}_p^T\mathbf{C}_2^{+})\mathbf{U}_p\mathbf{C}_{\mathbf{y}^*}\\ =&\mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T(\mathbf{C}_1^{-1}+\mathbf{C}_2^{+})\mathbf{U}_p\mathbf{C}_{\mathbf{y}^*}\\ =&\mathbf{C}_{\mathbf{y}^*}\mathbf{C}_{\mathbf{y}^*}^{-1}\mathbf{C}_{\mathbf{y}^*}\\ =&\mathbf{C}_{\mathbf{y}^*}\end{aligned} \tag{B.5}$$

And subsequently, for $\mathbf{x}^*$ we have

$$E[\mathbf{x}^*\mathbf{x}^{*T}] = E[\mathbf{U}_p\mathbf{y}^*\mathbf{y}^{*T}\mathbf{U}_p^T] = \mathbf{U}_p\mathbf{C}_{\mathbf{y}^*}\mathbf{U}_p^T \tag{B.6}$$

ACKNOWLEDGMENT

We have enjoyed many times the inspiring discussions with and guidances from Dr. Randolph P. Martin from Emory Heart Center. We are grateful for the generous support from Alok Gupta of Siemens Medical Solutions. We would like to thank Visvanathan Ramesh from Siemens Corporate Research and David Gustafson, Bhavani Duggirala, and Diane Paine from Siemens Ultrasound for helpful discussions on the subject.

REFERENCES

[1] T. Cootes and C. Taylor, "Active shape models-'smart snakes'," in *Proc. British Machine Vision Conference*, 1992, pp. 266–275.

[2] ——, "Active shape models - Their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[3] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Medical Imaging*, vol. 15, no. 3, pp. 290–298, 1996.

[4] M. Mignotte, J. Meunier, and J.-C. Tardif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable templates and markov random fields," *Pattern Analysis and Applications*, vol. 4, no. 4, pp. 256–271, November 2001.

[5] G. E. Mailloux, F. Langlois, P. Y. Simard, and M. Bertrand, "Restoration of the velocity field of the heart from two-dimensional echocardiograms," *IEEE Trans. Medical Imaging*, vol. 8, no. 2, pp. 143–153, 1989.

[6] D. Adam, O. Hareuveni, and S. Sideman, "Semiautomated border tracking of cine echocardiographic ventricular images," *IEEE Trans. Medical Imaging*, vol. 6, pp. 266–271, September 1987.

[7] P. Baraldi, A. Sarti, C. Lamberti, A. Prandini, and F. Sgal-lari, "Semiautomated border tracking of cine echocardiographic ventricular images," *IEEE Trans. Biomedical Eng.*., vol. 43, no. 3, pp. 259–272, March 1986.

[8] S. Lai and B. Vemuri, "Reliable and efficient computation of optical flow," *Intl. J. of Computer Vision*, vol. 29, no. 2, pp. 87–105, 1998.

[9] K. Althoff, G. Hamarneh, and T. Gustavsson, "Tracking contrast in echocardiography by a combined snake and optical flow technique," in *IEEE Computers in Cardiology*, Cambridge, USA, vol. 3653, September 2000, pp. 29–32.

[10] G. Jacob, J. Noble, C. Behrenbruch, A. Kelion, and A. Banning, "A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography," *IEEE Trans. Medical Imaging*, vol. 21, no. 3, pp. 226–238, 2002.

[11] A. Amini, Y. Chen, M. Elayyadi, and P. Radeva, "Tag surface reconstruction and tracking of myocardial beads from spamm-mri with parametric b-spline surfaces," *IEEE Trans. Medical Imaging*, vol. 20, no. 2, pp. 94–103, 2001.

[12] M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, R. H. Mohiaddin, and D. Rueckert, "Atlas-based segmentation of 4d cardiac mr sequences using non-rigid registration," in *Proc. of Medical. Image Computing and Computer Assisted Intervention (MICCAI)*, Tokyo, Japan, 2002, pp. 642–650.

[13] S. Mitchell, B. P. Lelieveldt, R. van der Geest, H. G. Bosch, J. H. Reiber, and M. Sonka, "Time-continuous segmentation of cardiac mr image sequences using active appearance motion models," in *Proc. SPIE Medical Imaging*, San Diego, CA, USA, vol. 4322, 2001, pp. 249–256.

[14] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3d ultrasound with mr images: a new approach combining intensity and gradient information," *IEEE Trans. Medical Imaging*, vol. 20, no. 10, pp. 1038–1049, 2001.

[15] A. Montillo, D. Metaxas, and L. Axel, "Automated segmentation of the left and right ventricles in 4d cardiac spamm images," in *Proc. of Medical. Image Computing and Computer Assisted Intervention (MICCAI)*, Tokyo, Japan, 2002, pp. 620–633.

[16] N. Jojic, M. Turk, and T. S. Huang, "Tracking self-occluding articulated objects in dense disparity maps," in *Proc. Intl. Conf. on Computer Vision*, Kerkyra, Greece, 1999, pp. 123–130.

[17] Y. Q. Chen, Y. Rui, and T. S. Huang, "Parametric contour tracking using unscented kalman filter," in *Proc. IEEE Intl. Conf. on Image Processing*, Rochester, NY, 2002, pp. 613–616.

[18] S. Avidan, "Support vector tracking," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, vol. I, 2001, pp. 184–191.
[19] T. Cootes and C. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Proc. SPIE Medical Imaging*, 2001, pp. 236–248.
[20] Y. Kanazawa and K. Kanatani, "Do we really have to consider covariance matrices for image features?" in *Proc. Intl. Conf. on Computer Vision*, Vancouver, Canada, vol. II, 2001, pp. 586–591.
[21] J. K. Oh, J. B. Seward, and A. J. Tajik, *The Echo Manual*. Lippincott Williams & Wilkins, Philadelphia,, 1999.
[22] G. Jacob, A. Noble, and A. Blake, "Robust contour tracking in echocardiographic sequence," in *Proc. Intl. Conf. on Computer Vision*, Bombay, India, 1998, pp. 408–413.
[23] A. Blake, M. Isard, and D. Reynard, "Learning to track the visual motion of contours," *Artificial Intelligence*, vol. 78, pp. 101–133, 1995.
[24] M. A. Turk and A. P. Pentland, "Face recognition using eigen-face," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, 1991, pp. 586–591.
[25] M. Black and A. Jepson, "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation," in *Proc. European Conf. on Computer Vision*, Cambridge, UK, 1996, pp. 610–619.
[26] D. Comaniciu, "Nonparametric information fusion for motion estimation," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Madison, Wisconsin, vol. I, 2003, pp. 59–66.
[27] M. Brand and R. Bhotika, "Flexible flow for 3D nonrigid object tracking and shape recovery," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, vol. I, 2001, pp. 315–322.
[28] H. Murase and S. K. Nayar, "Visual learning and recognition of 3-d objects from appearance," *Intl. J. of Computer Vision*, vol. 14, no. 1, pp. 5–24, 1995.
[29] A. Baumberg and D. Hogg, "An efficient method for contour tracking using active shape models," in *Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 1994, pp. 194–199.
[30] A. Blake and M. Isard, *Active contours*. Springer Verlag, 1998.
[31] G. Jacob, A. Noble, M. Mulet-Parada, and A. Blake, "Evaluating a robust contour tracker on echocardiographic sequences," *Med. Image Analysis*, vol. 3, no. 1, pp. 63–75, 1999.
[32] A. Jepson, D. Fleet, and T. El-Maraghi, "Robust online appearance models for visual tracking," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, vol. I, 2001, pp. 415–422.
[33] C. Bregler, A. Hertzmann, and H. Biermann, "Recovering non-rigid 3d shape from image streams," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hilton Head, SC, vol. II, 2000, pp. 690–696.
[34] M. Irani, "Multi-frame optical flow estimation using subspace constraints," in *Proc. Intl. Conf. on Computer Vision*, Kerkyra, Greece, Sept. 1999, pp. 626–633.
[35] Y. Leedan and P. Meer, "Heteroscedastic regression in computer vision: Problems with bilinear constraint," *Intl. J. of Computer Vision*, vol. 37, no. 2, pp. 127–150, 2000.
[36] B. Matei and P. Meer, "Reduction of bias in maximum likelihood ellipse fitting," in *Int'l Conf. on Pattern Recog.*, 2000, pp. 802–806.
[37] J. Rehg and T. Kanade, "visual tracking of high dof articulated structures: An application to human hand tracking," in *Proc. European Conf. on Computer Vision*, Stockholm, Sweden, 1994, pp. 35–46.
[38] B. C. Vemuri, Y. Guo, and Z. Wang, "Deformable pedal curves and surfaces: Hybrid geometric active models for shape recovery," *Intl. J. of Computer Vision*, vol. 44, no. 2, pp. 137–155, 2001.
[39] M. Black and P. Ananadan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Computer Vision and Image Understanding*, vol. 63, no. 1, pp. 75–104, 1996.
[40] M. Rogers and J. Graham, "Robust active shape model search," in *Proc. European Conf. on Computer Vision*, Copenhagen, Denmark, 2002, pp. 517–530.
[41] P. Clarysse, D. Friboulet, and I. E. Magnin, "Tracking geometrical descriptors on 3-d deformable surfaces: application to the left-ventricular surface of the heart," *IEEE Trans. Medical Imaging*, vol. 16, no. 4, pp. 392–404, 1997.
[42] J. McEachen and J. Duncan, "Shape-based tracking of left ventricular wall motion," *IEEE Trans. Medical Imaging*, vol. 16, pp. 270–283, 1997.
[43] I. Mikić, S. Krucinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates," *IEEE Trans. Medical Imaging*, vol. 17, no. 2, pp. 274–284, 1998.
[44] M. Sühling, M. Arigovindan, P. Hunziker, and M. Unser, "Motion analysis of echocardiograms using a local-affine, spatio-temporal model," in *Proc. IEEE Int'l Symp. on Biomedical Imaging: Macro to Nano (ISBI'02)*, Washington DC, 2002, pp. 573–576.
[45] Y. Akgul and C. Kambhamettu, "A coarse-to-fine deformable contour optimization framework," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 2, pp. 174–186, 2003.
[46] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the Am. Soc. of Mechanical Eng., D, Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.
[47] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*. Prentice Hall, 1993.
[48] Y. Bar-Shalom and L. Campo, "The effect of the common process noise on the two-sensor fused track covariance," *IEEE Trans. Aero. Elect. Syst.*, vol. AES-22, no. 22, pp. 803–805, 1986.
[49] X. Li, Y. Zhu, and C. Han, "Unified optimal linear estimation fusion - part i: Unified models and fusion rules," in *Proc. of 3rd Intl. Conf. on Information Fusion*, Paris, *France*, 2000, pp. MoC2–10–MoC2–17.
[50] L. L. Scharf, *Statistical Signal Processing*. Addison Wesley, Reading, MA, 1991.
[51] P. Hall, D. Marshall, and R. Martin, "Merging and splitting eigenspace models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 9, pp. 1042–1048, 2000.
[52] D. Comaniciu, "An algorithm for data-driven bandwidth selection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 2, pp. 281–288, 2003.
[53] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. DARPA Imaging and Understanding Workshop*, 1981, pp. 121–130.
[54] A. Hoerl and R. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
[55] D. W. Marquardt, "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation," *Technometrics*, vol. 12, no. 3, pp. 591–612, 1970.
[56] T. Cootes and C. Taylor, "Statistical models for appearance for computer vision," 2001, unpublished manuscript. Available at `http://www.wiau.man.ac.uk/~bim/Models/app_model.ps.gz`.