

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
23 September 2004 (23.09.2004)

PCT

(10) International Publication Number  
**WO 2004/081875 A2**

(51) International Patent Classification<sup>7</sup>: **G06T 7/20**

(21) International Application Number:  
PCT/US2004/007068

(22) International Filing Date: 8 March 2004 (08.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/452,669 7 March 2003 (07.03.2003) US  
60/473,425 27 May 2003 (27.05.2003) US  
60/508,367 3 October 2003 (03.10.2003) US  
10/794,476 5 March 2004 (05.03.2004) US

(71) Applicants (for all designated States except US):  
**SIEMENS CORPORATE RESEARCH INC.** [US/US];  
755 College Road East, Princeton, NJ 08540 (US).  
**SIEMENS MEDICAL SOLUTIONS USA INC**  
[US/US]; 51 Valley Stream Parkway, Malvern, PA  
19355-1406 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **COMANICIU,**

Dorin [RO/US]; 2 Stuart Ln. West, Princeton Jct, NJ  
08550 (US). **KRISHNAN, Sriram** [US/US]; 261 Linden-  
wood Drive, Exton, PA 19341 (US). **ZHOU, Xiang Sean**  
[CN/US]; 36 Sycamore Dr., Plainsboro, NJ 08536 (US).  
**GEORGESCU, Bogdan** [RO/US]; 9 Radcliffe Ave., Apt  
1B, Highland Park, NJ 08904 (US).

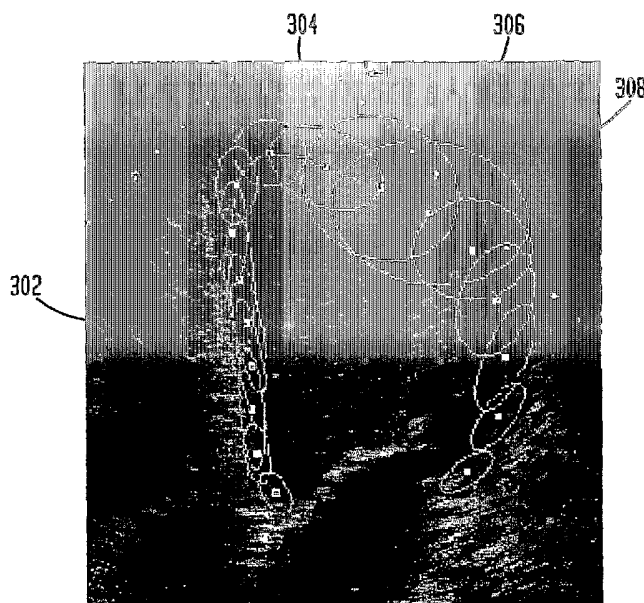
(74) Agents: **CONOVER, Michele L.** et al.; Siemens Corpora-  
tion- Intellectual Property Dept., 170 Wood Avenue South,  
Iselin, NJ 08830 (US).

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,  
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,  
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,  
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,  
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,  
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,  
ZW.

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR TRACKING A GLOBAL SHAPE OF AN OBJECT IN MOTION



(57) Abstract: A system and method for tracking a global shape of an object in motion is disclosed. One or more control points along an initial contour of the global shape are defined. Each of the one or more control points is tracked as the object is in motion. Uncertainty of a location of a control point in motion is represented using a number of techniques. The uncertainty to constrain the global shape is exploited using a prior shape model. In an alternative embodiment, multiple appearance models are built for each control point and the motion vectors produced by each model are combined in order to track the shape of the object. The movement of the shape of the object can be visually tracked using a display and color vectors.

WO 2004/081875 A2



Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

— *without international search report and to be republished upon receipt of that report*

## **SYSTEM AND METHOD FOR TRACKING A GLOBAL SHAPE OF AN OBJECT IN MOTION**

### **Cross Reference to Related Application**

This application claims the benefit of U.S. Provisional Application Serial No. 60/452,669, filed on March 7, 2003, U.S. Provisional Application Serial No. 60/473,425, filed on May 27, 2003, and U.S. Provisional Application Serial No. 60/508,367 filed on October 3, 2003 which are incorporated by reference in their entirety.

### **Field of the Invention**

The present invention is directed to a method for tracking a shape that is in motion, and more particularly, to a method for tracking a shape having linear constraints in the presence of heteroscedastic noise.

### **Background of the Invention**

For most visual tracking applications, measurement data are uncertain and sometimes missing: images are taken with noise and distortion, while occlusions can render part of the object-of-interest unobservable. Uncertainty can be globally uniform; but in most real-world scenarios, it is heteroscedastic in nature, i.e., both anisotropic and inhomogeneous. A good example is the echocardiogram (ultrasound heart data). 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.

Due to its availability, relative low cost, and noninvasiveness, cardiac ultrasound images are widely used for assessing cardiac functions. In particular,

the analysis of ventricle motion is an efficient way to evaluate the degree of ischemia and infarction. Segmentation or detection of the endocardium wall is the first step towards quantification of elasticity and contractility of the left ventricle. Examples of some existing methods include pixel-based segmentation / clustering approaches (e.g., Color Kinesis), variants of optical flow, deformable templates and Markov random process / fields, and active contours / snakes. Some methods are employed in 2-Dimensional, 3-Dimensional or 4-Dimensional (3D + time) space.

However, most existing segmentation or detection methods do not attempt to recover accurate regional motions of the endocardial wall, and in most cases, motion components along the wall are ignored. This simplified treatment is also employed by contour trackers that search only along the normals of the current contour. This is not suitable for regional wall abnormality detection, because regional motion of an abnormal left ventricle is likely to be off the normal of the contour, not to mention that global motion, such as translation or rotation (due to the sonographer's hand motion or respiratory motion the patient), causes off-normal local motion on the contour as well. It is desirable to track the global shape of endocardial wall as well as its local motion, for the detection of regional wall motion abnormalities. This information can be used for further diagnosis of ischemia and infarction.

In general, covariances can be assigned to image features or flow estimates that reflect underlying heteroscedastic noise. When the data is clean with a low overall noise level, the heteroscedastic nature may be ignorable, and a global uncertainty can be substituted for the local estimates. However, for very noisy inputs, especially those with spatially varying structural noise, the information encoded in the local covariance matrix becomes critical in ensuring reliable and robust inference of objects or underlying image structures.

It is a common practice to impose model constraints in a tracking framework. Examples include simple models such as blobs or parameterized ellipses, and complex models such as discriminative templates. 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. Furthermore, a Principal Component Analysis (PCA) -based eigenshape subspace can capture arbitrarily complicated shape variations, which, in the original space, even with a very simple parametric model, are highly nonlinear.

If a measurement vector 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. 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 does not take into account the heteroscedastic noise in the measurements. In the "Point Distribution Model" or "Active Shape Model", a PCA-based subspace shape model is derived based on training shapes with landmark point correspondence. The resulting subspace of eigenshapes captures the most significant variations in the training data set. At detection time, a model is perturbed to create synthetic images for matching against the testing image at a candidate location. However, the measurement noise was not modeled in this process.

Even when heteroscedastic noise characteristics are available, they were typically disregarded during the subspace model fitting. For example, in one known approach where full covariance matrix was captured in the measurements, a rather ad hoc thresholding is applied so that the measurement mean is confined to a hyper-ellipsoid constraint defined by the model covariance. This operation is independent of the measurement noise.

Another known approach applies 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. Therefore, this approach leads to information loss during the projection.

Another known approach uses 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 the present invention, local uncertainty is captured in the covariance matrix. The difference is that the present invention specifically studies the subspace model constraints and the critical choice of intersection over projection when anisotropic uncertainty is present.

Another known approach uses a subspace constraint implicitly during the optical flow estimation, which also utilizes flow uncertainties. Although in a different framework for a different application, present invention recognizes that “more reliable flow-vectors will have more influence in the subspace projection process.”

Another known approach applies heteroscedastic regression for fitting ellipses and fundamental matrices. The fitting is achieved in the original space with parameterized models. In the present invention, parameterization of shape variations is avoided — it can be very complicated and highly nonlinear. Instead, the present invention builds subspace linear probabilistic models through, e.g., PCA, and obtain closed-form solutions on both the mean and covariance of the fitted data.

Robust model matching relying on M-estimators or RANSAC has also 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.

Other related approaches include data imputation, the practice of “filling in” missing data with plausible values. Work in this area is rooted in statistics with broad applications toward speech recognition, medical image analysis, and social science, etc. However, the formulation of data imputation problems typically assumes 0-1 availability, i.e., a data component is either missing or available. There is a need for a unified framework for fusing subspace model constraints with information about the shape dynamic and the heteroscedastic nature of the measurement noise and about the shape dynamics.

### **Summary of the Invention**

The present invention is directed to a system and method for tracking a global shape of an object in motion. One or more control points along an initial contour of the global shape are defined. Each of the one or more control points is tracked as the object is in motion. Uncertainty of a location of a control point in motion is estimated. One form to represent the uncertainty is a covariance matrix. When employing a subspace shape constraint model, the uncertainty is exploited using a non-orthogonal projection and/or information fusion. Each subsequent contour is displayed.

The present invention is also directed to a system for visually tracking movement of a shape of an object. One or more first color vectors are generated to represent contraction of control points along the contour of the shape. One or more second control vectors are generated to represent dilation of control points along the contour of the shape. The first and second color vectors are periodically displayed thereby displaying movement of the shape.

The present invention is also directed to a method for tracking a global shape of an object in motion. One or more control points are defined along the global shape. Each of the one or more control points are tracked as the object is in motion. Multiple appearance models are built for each control point. Motion vectors produced by each model for tracking the shape are combined.

### **Brief Description of the Drawings**

Preferred embodiments of the present invention will be described below in more detail, wherein like reference numerals indicate like elements, with reference to the accompanying drawings:

FIG. 1 is a block diagram of an exemplary system for implementing a method for shape tracking in accordance with the present invention;

FIG. 2 is an echocardiography image of a heart that illustrates areas of acoustic drop out and estimated local wall motion uncertainties;

FIG. 3 is an echocardiography image of a left ventricle that illustrates an endocardial contour with localization uncertainties of its control points that are strongly anisotropic and inhomogeneous;

FIG. 4 illustrates examples of an incremental PCA model and a strongly adapted PCA (SA-PCA) model with different  $\alpha$  values;

FIGs. 5a-5c illustrate echocardiography images of a left ventricle in which the endocardial wall is initialized (a) and tracked using IPCA (b) and SA-PCA (c);

FIGs. 6a and 6b illustrate echocardiography images of a left ventricle wherein the movement of the endocardial wall is tracked in accordance with the present invention;

FIGs. 7a-7d illustrate test contours representing echocardiography images of an endocardial wall of a left ventricle from an apical four-chamber view in accordance with the present invention;

FIGs. 8a-8d illustrate test contours representing echocardiography images of an endocardial wall of a left ventricle from a parasternal short axis view in accordance with the present invention;

FIG. 9 illustrates echocardiography images which exemplify a method for visualizing the images in accordance with the present invention; and

FIG. 10 illustrates a multi-model control point based tracker in accordance with an embodiment of the present invention.

### **Detailed Description**

The present invention is directed to a method for tracking shapes with linear constraints in the presence of heterodastic noise. An example where such a method would be utilized is for tracking the global shape of a myocardial wall as well as its local motion to detect regional wall motion abnormalities in the heart. The method may also be used to track the endocardial wall or epicardial wall of the heart. It is to be understood by those skilled in the art that the present invention may be used in other applications where shape tracking is useful such as, but not limited to, recognizing movement of human features such as head movements, facial features, hand movements or other body movements.



The present invention can also be used in 2 dimensional, 3 dimensional and 4 dimensional (3D + time) medical analyses of anatomical structures such as the heart, lungs or tumors that are evolving over time.

For purposes of describing the present invention, an example will be described for tracking the endocardial wall of the left ventricle. FIG.1 illustrates an exemplary architecture of an echocardiograph system that uses a method for tracking the shape of an endocardial wall of a left ventricle in accordance with the present invention. A medical sensor 102, such as an ultrasound transducer is used to perform an examination on a patient. The sensor 102 is used to obtain medical measurements consistent with a particular medical examination. For example, a patient experiencing heart problems may have an echocardiogram performed to help diagnose the particular heart ailment. An ultrasound system provides two-, three-, and four (3D + time)-dimensional images of the heart from various perspectives.

The information obtained by the sensor 102 is communicated to a processor 104 which may be a workstation or personal computer. The processor 104 converts the sensor data into an image that is communicated to display 108. The display 108 may also communicate other graphical information or tables of information relating to the image. In accordance with the present invention, the processor 104 is also provided with data representing an initial contour of the endocardial wall. The data may be provided manually by a user such as a physician or sonographer, or automatically by the processor 104. The contour comprises a series of individual points, the movement of which is tracked by the processor 104 and illustrated on display 108. The specifics regarding how the individual points are tracked will be described in greater detail hereinafter.

In addition to data from the medical sensor 102, the processor 104 may also receive other data inputs. For example, the processor may receive data from a database 106 associated with the processor 104. Such data may include subspace models that represent potential contour shapes for the endocardial wall. These subspace models may be images of left ventricles that are

representative of a plurality of patients or may be computer generated models of contour shapes based on statistical information. The processor 104 tracks the individual points of the contour shape using known approaches such as Bayesian kernel matching or optical flow-based methods. Error accumulation during tracking is remedied by using a multi-template adaptive matching framework. Uncertainty of tracking is represented at each point in the form of a covariance matrix, which is subsequently fully exploited by a subspace shape constraint using a non-orthogonal projection.

FIG. 2 illustrates a typical echocardiograph image of a heart. The part of the endocardial wall of the left ventricle that has acoustic drop out is marked by solid ellipse 208. Estimations of local wall motion are depicted by the dotted ellipses 202, 204. Because of the acoustic drop out, the endocardium wall is not always at the strongest edge in the image. In accordance with the present invention, a unified framework is used for fusing subspace model constraints with information about the shape dynamic and the heteroscedastic nature of the measurement noise and about the shape dynamics. 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.

The present invention tracks individual control points on a contour that represents the endocardium wall. The tracking may be performed using a Bayesian kernel matching approach or a flow-based approach. An example of a Bayesian kernel matching approach is described in the article authored by co-inventor, Dorin Comaniciu and titled: *Bayesian Kernel Tracking*, Annual Conf. of the German Society for Pattern Recognition (DAGM'02), Zurich, Switzerland, 438-445, 2002 which is incorporated by reference in its entirety. An example of an optical flow based method for tracking the individual points is described in co-pending application serial no. 10/681,702 entitled "Density Estimation-Based Information Fusion for Multiple Motion Computation" which is incorporated by reference in its entirety. In accordance with the present invention, error accumulation during tracking is remedied using a multi-template adaptive matching framework, taking advantage of the periodic nature of cardiac motion.

Uncertainty of tracking is represented at each point in the form of a covariance matrix, which is subsequently fully exploited by a subspace shape constraint using a non-orthogonal projection.

The tracking framework is a two-step, iterative process over the image sequences. An initial contour with control points is drawn on the first image (either automatically or manually); then, for every subsequent image in order, every control point is first tracked independently, with anisotropic uncertainty also recorded. As the second step, the new contour is projected into a feasible subspace using non-orthogonal projection. The feasible subspace is learned based training contours and also adapted to the current case using the initial contour available to the tracker, also taking into account the confidence of initialization (i.e., a manual initialization carries a higher confidence than a fully automatic one).

As indicated above, during the tracking process multiple templates are employed. The use of the multiple templates results in a more accurate representation of the shape statistics for the current case. In the Bayesian kernel matching approach, a first template is taken from the initialized frame. Subsequent templates are added when they are both sufficiently different from the existing templates, and also sufficiently informative for localization purpose, which is measured by kernel matching with itself: a more informative patch is one that has a higher confidence when matched to itself.

The decision to use more than one templates is based on the observation that the heart motion is periodic, therefore, different appearance patterns within one cycle will all reappear in later cycles. To match with multiple templates, match is made to each template and the one with the best match is selected. Or, to save computation, only the neighboring templates of the previous matching template may be matched, exploiting again the periodic motion.

Uncertainty on matching location is calculated in the neighborhood of the optimal location, with the likelihood map calculated in the same way using the kernel matching process. The likelihood surface is then used for estimating a

covariance matrix, through, e.g., fitting with a two-dimensional Gaussian distribution or the inversion of a weighted estimation of the Hessian matrix.

In the case of the flow-based approach, a typical optical flow implementation uses only neighboring frames, making the tracking process over long sequences susceptible to error accumulation and drifting. Multiple templates are employed for flow calculation, much the same way as for kernel matching. A new template is added whenever the flow uncertainty is high, while the local gradient is distinctive from existing templates.

After the new locations are obtained for every control point, the next step is to constrain the overall shape by a statistical model that captures the "legitimate" shape variations of a human heart. PCA-based shape models, which are also known as point distribution models, or active shape models are used.

Because the uncertainties over the heart contour are neither homogeneous, some areas are worse than others due to, say, signal dropouts; nor isotropic, for example, there is worse localization along an edge than along the gradient direction. FIG. 3 shows an example of such anisotropic and inhomogeneous noise across a contour. As can be seen, individual points, such as points 302, 304, 306 and 308 have been initially identified along the contour of the endocardial wall. A certainty measurement is taken for each point and an ellipsoid is created around each point which indicates the level of certainty that that particular point is in the correct location; the larger the ellipsoid, the greater the level of uncertainty as to the location of a particular point. As seen in FIG. 3, point 302 has a relatively small ellipsoid surrounding it, thereby indicating that there is a high degree of certainty as to the location of the point on the contour. Points 306 and 308 have relatively large ellipsoids surrounding them, thereby indicating a large degree of uncertainty as to the location of these points.

Taking into account a general uncertainty model over the new contour points, the optimal projection will no longer be orthogonal. The optimal solution in the subspace is in fact the maximal likelihood shape on the intersected distribution in the shape model subspace. Furthermore, if there is a distribution model in the subspace, there is no reason for ignoring such extra information. In

the following the detailed analysis is provided as to how to perform non-orthogonal projection; and, if the subspace model distribution is available, how to fuse the information of such a model with the uncertain input.

The present invention uses a subspace model-based fusion approach. Given two noisy measurements of the same  $n$ -dimensional variable  $\mathbf{x}$ , each characterized by a multidimensional Gaussian distribution,  $p_1$  and  $p_2$ , the maximum likelihood estimate of  $\mathbf{x}$  is the point with the minimal sum of Mahalanobis distances to the two centroids:  $\mathbf{x}^* = \operatorname{argmin} d_m$  with

$$d_m = (\mathbf{x} - \mathbf{x}_1)^T \mathbf{C}_1^{-1} (\mathbf{x} - \mathbf{x}_1) + (\mathbf{x} - \mathbf{x}_2)^T \mathbf{C}_2^{-1} (\mathbf{x} - \mathbf{x}_2) \quad (1)$$

Taking derivative with respect to  $\mathbf{x}$ , we get:

$$\begin{aligned} \mathbf{x}^* &= \mathbf{C}(\mathbf{C}_1^{-1} \mathbf{x}_1 + \mathbf{C}_2^{-1} \mathbf{x}_2) \\ \mathbf{C} &= (\mathbf{C}_1^{-1} + \mathbf{C}_2^{-1})^{-1} \end{aligned} \quad (2)$$

This is also known as the best linear unbiased estimate (BLUE) of  $\mathbf{x}$ .

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  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ , with  $\mathbf{u}_i$ 's orthonormal and  $\mathbf{\Lambda} = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_p, 0, \dots, 0\}$ , we rewrite Mahalanobis distance to  $\mathbf{x}_2$  in Eq. (1) in the canonical form:

$$d_{m,2} = (\mathbf{x} - \mathbf{x}_2)^T \mathbf{C}_2^{-1} (\mathbf{x} - \mathbf{x}_2) = \sum_{i=1}^n \lambda_i^{-1} [\mathbf{U}^T (\mathbf{x} - \mathbf{x}_2)]^2 \quad (3)$$

When  $\lambda_i$  tends to 0,  $d_{m,2}$  goes to infinity, unless  $\mathbf{U}_0^T \mathbf{x} = 0$ , where  $\mathbf{U}_0 = [\mathbf{u}_{p+1}, \mathbf{u}_{p+2}, \dots, \mathbf{u}_n]$ . (Here we have assumed, without loss of generality, the subspace

passes through the origin of the original space. Since  $\mathbf{x}_2$  resides in the subspace,  $\mathbf{U}_0^T \mathbf{x}_2 = \mathbf{0}$ .)

Therefore, we only need to retain in Eq. (3) those terms corresponding to a non-zero  $\lambda_i$ :

$$d_{m,2} = \sum_{i=1}^p \lambda_i^{-1} [\mathbf{U}_p^T (\mathbf{x} - \mathbf{x}_2)]^2 \equiv (\mathbf{x} - \mathbf{x}_2)^T \mathbf{C}_2^+ (\mathbf{x} - \mathbf{x}_2), \quad (4)$$

where  $\mathbf{C}^+$  is the pseudoinverse of  $\mathbf{C}$ , and  $\mathbf{U}_p = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p]$ .

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

$$\text{for } \mathbf{x} = \mathbf{U} \mathbf{U}^T \mathbf{x} = \mathbf{U} ([\mathbf{U}_p \mid \mathbf{U}_0]^T \mathbf{x}) = \mathbf{U} [\mathbf{y} \mid \mathbf{0}] = \mathbf{U}_p \mathbf{y}, \quad (5) \quad \text{a } 1 \times p$$

vector  $\mathbf{y}$ . Eq. (1) now takes the following general form despite singularity:

$$d_m = (\mathbf{U}_p \mathbf{y} - \mathbf{x}_1)^T \mathbf{C}_1^{-1} (\mathbf{U}_p \mathbf{y} - \mathbf{x}_1) + (\mathbf{U}_p \mathbf{y} - \mathbf{x}_2)^T \mathbf{C}_2^+ (\mathbf{U}_p \mathbf{y} - \mathbf{x}_2) \quad (6)$$

Taking derivative with respect to  $\mathbf{y}$  yields

$$\mathbf{y}^* = \mathbf{C}_{y^*} \mathbf{U}_p^T (\mathbf{C}_1^{-1} \mathbf{x}_1 + \mathbf{C}_2^+ \mathbf{x}_2) \quad (7)$$

$$\mathbf{C}_{y^*} = [\mathbf{U}_p^T (\mathbf{C}_1^{-1} + \mathbf{C}_2^+) \mathbf{U}_p]^{-1} \quad (8)$$

$$\mathbf{x}^* = \mathbf{U}_p \mathbf{y}^* = \mathbf{C}_{x^*} (\mathbf{C}_1^{-1} \mathbf{x}_1 + \mathbf{C}_2^+ \mathbf{x}_2) \quad (9)$$

$$\mathbf{C}_{x^*} = \mathbf{U}_p \mathbf{C}_{y^*} \mathbf{U}_p^T \quad (10)$$

Eq. (7) indicates a fusion in the information space, followed by a transformation into the subspace. Eq. (9) is a coordinate transform back into the

original space. It can be shown that  $\mathbf{C}_{\mathbf{x}^*}$  and  $\mathbf{C}_{\mathbf{y}^*}$  are the corresponding covariance matrices for  $\mathbf{x}^*$  and  $\mathbf{y}^*$ .

Notice that this solution is not a simple generalization of Eq. (2) by substituting pseudoinverses for regular inverses, which will not constrain  $\mathbf{x}^*$  to be in the subspace.

Alternatively, we can write out Eq. (7) and (8) as

$$\mathbf{y}^* = (\mathbf{U}_p^T \mathbf{C}_1^{-1} \mathbf{U}_p + \mathbf{\Lambda}_p^{-1})^{-1} (\mathbf{U}_p^T \mathbf{C}_1^{-1} \mathbf{x}_1 + \mathbf{\Lambda}_p^{-1} \mathbf{y}_2) \quad (11)$$

Here  $\mathbf{y}_2$  is the transformed coordinates of  $\mathbf{x}_2$  in the subspace spanned by  $\mathbf{U}_p$ , and  $\mathbf{\Lambda}_p = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_p\}$ .

Eq. (11) can also be shown as a BLUE fusion in the subspace of two distributions, one is  $N(\mathbf{y}_2, \mathbf{\Lambda}_p)$  and the other is the "intersection" (not projection!) of  $N(\mathbf{x}_1, \mathbf{C}_1)$  in the subspace,  $N((\mathbf{U}_p^T \mathbf{C}_1^{-1} \mathbf{U}_p)^{-1} \mathbf{U}_p^T \mathbf{C}_1^{-1} \mathbf{x}_1, (\mathbf{U}_p^T \mathbf{C}_1^{-1} \mathbf{U}_p)^{-1})$ .

The use of a statistical shape model learned from a large pool of training samples to guide the contours from a specific heart can sometimes be problematic. In accordance with the present invention, the training samples are used to obtain a shape model of the current heart, and not a generic heart. Therefore, there is a strong motivation to adapt the generic model toward what is known for the current case. In accordance with the present invention, the initial contour (manual or through automatic detection) of the endocardial wall of the heart of the patient is examined to adapt the existing PCA model.

In determining the actual contour of the endocardial wall and tracking its movement, two approaches are considered: one with the initial contour being deterministic, and the other with the initial contour being uncertain (this can be the case when the initial contour is from an automatic detection algorithm, which also provides uncertainties).

When the initial contour is assumed to be a point (deterministically certain), a strongly-adapted-PCA (SA-PCA) model is used to track the movement

of the point. It is assumed that the old 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 (i.e., representative power) being  $\alpha$  and  $(1 - \alpha)$ , respectively, with  $0 < \alpha < 1$ . In other words, we assume that a portion of shape variations of the current case will be represented in the generic model, while the rest is captured in the direction of the initial contour.

The PCA model is denoted by its mean, eigenvalue matrix, and eigenvector matrix 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:

$$\mathbf{x}_{m,new} = \alpha\mathbf{x}_m + (1 - \alpha)\mathbf{x} \quad (12)$$

$$\begin{aligned} \mathbf{C}_{new} &= \alpha(\mathbf{C} + (\mathbf{x}_m - \mathbf{x}_{m,new})(\mathbf{x}_m - \mathbf{x}_{m,new})^T) \\ &\quad + (1 - \alpha)(\mathbf{x} - \mathbf{x}_{m,new})(\mathbf{x} - \mathbf{x}_{m,new})^T \\ &= \alpha\mathbf{C} + \alpha(1 - \alpha)(\mathbf{x} - \mathbf{x}_m)(\mathbf{x} - \mathbf{x}_m)^T \end{aligned} \quad (13)$$

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:

$$\mathbf{U}_{new} = [\mathbf{U}, \mathbf{x}_{ru}] \mathbf{R} \quad (14)$$

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



$$\left( \alpha \begin{bmatrix} \Lambda & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} + \alpha(1-\alpha) \begin{bmatrix} \mathbf{x}_s \mathbf{x}_s^T & e_r \mathbf{x}_s \\ e_r \mathbf{x}_s^T & e_r^2 \end{bmatrix} \right) \mathbf{R} = \mathbf{R} \Lambda_{new} \quad (15)$$

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

Above formulae are more general than IPCA, with tunable energy ratios between the new data and the old data. These become equivalent to IPCA if we set  $\alpha$  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  $\alpha$  set at a smaller value (say, 0.5), the PCA model is strongly adapted toward the current case, hence the name. FIG. 4 shows a simple two-dimensional illustration of IPCA and SA-PCA with different  $\alpha$  values.

Point 404 represents the current case. Each x 402 represents a training point corresponding to a particular model. Ellipse 406 shows the original model distribution. Ellipse 408 illustrates an Incremental PCA model which corresponds to a strongly adapted PCA model in which  $\alpha = 0.99$ . Ellipse 410 illustrates a strongly adapted PCA model in which  $\alpha = 0.5$ . Ellipse 412 illustrates a strongly adapted PCA model in which  $\alpha = 0.1$ . Each ellipse depicts a 90% equal-probable contour of the corresponding distribution.

Indeed, the contours from the current heart are much 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 well represented in the training set. In our system, we set  $\alpha$  at 0.5. This permits a strong influence from the initial contour: 50% of model energy is from the initial contour (this reduces to only 1% if IPCA is applied with a model trained with 99 examples).

FIGs. 5a-5c show a comparison of IPCA and SA-PCA. This parasternal short axis view has an irregular shape (with a concave part), while the training set is overwhelmingly populated with circular shapes. The incremental PCA

model, taking in the initial contour (FIG. 5a) 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 (FIG. 5b). This result is in fact the same as that obtained using the old PCA model without an incremental step. Our adaptive PCA model with  $\alpha = 0.5$ , fits much better to the true borders (FIG. 5c).

A subtle while important interaction between the non-orthogonal projection with fusion approach and the SA-PCA model adaptation is as follows: the fusion with the model mean and covariance is necessary to filter out contours that are in the subspace but too far from the model distribution. However, this stronger constraint (than orthogonal projection) will inevitably alter low likelihood, or outlier heart contours. The SA-PCA model mediates this dilemma through a strong model shift toward the direction of the current case using the extra information provided in the initial contour given.

A Kalman filter fuses information from the prediction defined by a dynamic process and from noisy measurements. When applied for shape tracking, additional global constraints are necessary to stabilize the overall shape in a feasible range. The present invention utilizes a unified fusion framework to incorporate subspace model constraints into a Kalman filter.

For a Kalman filter, the measurement update equation has the following form:

$$\mathbf{x}_{k+1|k+1} = \mathbf{x}_{k+1|k} + \mathbf{K}(\mathbf{z}_{k+1} - \mathbf{H}\mathbf{x}_{k+1|k}) \quad (16)$$

where

$$\mathbf{K} = \mathbf{P}_{k+1|k} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k+1|k} \mathbf{H}^T + \mathbf{R})^{-1} \quad (17)$$

$$\mathbf{P}_{k+1|k+1} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k} \quad (18)$$

$$\mathbf{P}_{k+1|k} = \mathbf{S}\mathbf{P}_{k|k}\mathbf{S}^T + \mathbf{Q} \quad (19)$$

Here  $\mathbf{P}$  is the estimation error covariance, and  $x_{ij}$  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  $\mathbf{Q}$ .

Using the approach above, the update equations of the Kalman filter with subspace constraint and heteroscedastic noise are given by

$$\mathbf{x}_{k+1|k+1} = \mathbf{P}_{k+1|k+1} ((\mathbf{P}_{k|k} + \mathbf{Q})^+ \mathbf{x}_{k|k} + \mathbf{R}^{-1} \mathbf{z} + \mathbf{C}_2^+ \mathbf{x}_2) \quad (20)$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{U}_p [\mathbf{U}_p^T ((\mathbf{P}_{k|k} + \mathbf{Q})^+ + \mathbf{R}^{-1} + \mathbf{C}_2^+) \mathbf{U}_p]^{-1} \mathbf{U}_p^T \quad (21)$$

where it is assumed that the system noise covariance  $\mathbf{Q}$  is contained 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, subspace constraint, and noise information. They represent the complete representation of various uncertainties that affect the tracking system.

In accordance with an alternative embodiment of the present invention, a tracking technique is used that relies on control point based object representations and on robust fusion to integrate model information across frames. A set of control point based object representations are maintained and acquired at different time instances. The estimated motion suggested by the control points are robustly combined to determine the next position of the object. Visual tracking of the object is achieved by maintaining several models over time. The result is a nonparametric representation of the probability density function that characterizes the object appearance. Tracking is performed by obtaining independently from each model a motion estimate and its uncertainty through optical flow. The final estimate for each control point is computed using a robust fusion technique such as a Variable-Bandwidth Density-based Fusion (VBDF) procedure. VBDF computes the location of the most significant mode of the

displacements density function while taking into account their uncertainty. The VBDF procedure manages the multiple data sources and outliers in the displacement estimates. Occlusions are naturally handled through the estimate uncertainty for large residual errors. The model is divided in several regions for which the flow is computed independently. The residual alignment error is used to compute the scale of the covariance matrix of the estimate, therefore reducing the influence of the unreliable displacements.

Object tracking challenges due to occlusions and appearance variations are handled in the present invention through a multi-modal control point based approach. Maintaining several representatives for a 2 dimensional appearance model does not restrict it to a unimodal distribution and the VBDF mechanism robustly integrates multiple estimates to determine the most dominant motion for each control point. To model changes during tracking, several exemplars of the object appearance over time are maintained.

FIG. 10 illustrates an exemplary multi-model control point based tracker in accordance with the present invention. The top three frames 1002, 1004 and 1006 illustrate the current exemplars in the model set, each having associated a set of overlapping control points. A control point based approach is more robust than a global representation and therefore less sensitive to illumination changes and pose. Another advantage of the present invention is that partial occlusion can be handled at the control point level by analyzing the matching likelihood.

Each control point is processed independently; its location and covariance matrix is estimated in the current image with respect to all of the model templates. For example, one of the control points is illustrated by gray rectangle 1014 and its location and uncertainty with respect to each model is shown in  $I_{new}$  1008. The VBDF robust fusion procedure is applied to determine the most dominant motion (mode) with the associated uncertainty as illustrated in frame 1010. Note the variance in the estimated location of each control point due to occlusion or appearance change.

The location of the control points in the current frame is further constrained by a global parametric motion model. A similarity transformation model and its parameters are estimated using the confidence in each control point location. Therefore, the reliable control points contribute more to the global motion estimation. The current frame is added to the model set if the residual error so the reference appearances is relatively low. The threshold is chosen such that the images are not added where the object has significant occlusion. The number of templates in the model can be variable or fixed. If the number of templates are fixed, a scheme can be provided in which to discard certain templates (e.g., discard the oldest template).

A VBDF estimator is based on nonparametric density estimation with adaptive kernel bandwidths. The VBDF estimator is defined as the location of the most significant mode of the density function. The mode computation is based on the variable-bandwidth mean shift technique in a multiscale optimization framework.

Let  $x_i \in \mathbf{R}^d$ ,  $i = 1 \dots n$  be the available  $d$ -dimensional estimates, each having an associated uncertainty given by the covariance matrix  $C_i$ . The most significant mode of the density function is determined iteratively in a multiscale fashion. A bandwidth matrix  $H_i = C_i + \alpha^2 I$  is associated with each point  $x_i$ , where  $I$  is the identity matrix and the parameter  $\alpha$  determines the scale of the analysis. The sample point density estimator at location  $x$  is defined by

$$\hat{f}(x) = \frac{1}{n(2\pi)^{d/2}} \sum_{i=1}^n \exp\left(-\frac{1}{2}D^2(x, x_i, H_i)\right) \quad (22)$$

where  $D$  represents the Mahalanobis distance between  $x$  and  $x_i$

$$D^2(x, x_i, H_i) = (x - x_i)^\top H_i^{-1} (x - x_i) \quad (23)$$

The variable bandwidth mean shift vector at location  $x$  is given by

$$m(x) = H_h(x) \sum_{i=1}^n \omega_i(x) H_i^{-1} x_i - x \quad (24)$$

where  $H_h$  represents the harmonic mean of the bandwidth matrices weighted by the data-dependent weights  $\omega_i(x)$

$$H_h(x) = \left( \sum_{i=1}^n \omega_i(x) H_i^{-1} \right)^{-1} . \quad (25)$$

The data dependent weights computed at the current location  $x$  have the expression

$$\omega_i(x) = \frac{\frac{1}{|H_i|^{1/2}} \exp\left(-\frac{1}{2} D^2(x, x_i, H_i)\right)}{\sum_{i=1}^n \frac{1}{|H_i|^{1/2}} \exp\left(-\frac{1}{2} D^2(x, x_i, H_i)\right)} \quad (26)$$

and note that they satisfy  $\sum_{i=1}^n \omega_i(x) = 1$ .

It can be shown that the density corresponding to the point  $x + m(x)$  is always higher or equal to the one corresponding to  $x$ . Therefore iteratively updating the current

location using the mean shift vector yields a hill-climbing procedure which converges to a stationary point of the underlying density.

The VBDF estimator finds the most important mode by iteratively applying the adaptive mean shift procedure at several scales. It starts from a large scale by choosing the parameter  $\alpha$  large with respect to the spread of the points  $x_i$ . In this case the density surface is unimodal therefore the determined mode will correspond to the globally densest region. The procedure is repeated while reducing the value of the parameter  $\alpha$  and starting the mean shift iterations from the mode determined at the previous scale. For the final step the bandwidth matrix associated to each point is equal to the covariance matrix, i.e.  $H_i = C_i$ .

The VBDF estimator is a powerful tool for information fusion with the ability to deal with multiple source models. This is important for motion estimation as points in a local neighborhood may exhibit multiple motions. The most significant mode corresponds to the most relevant motion.

Consider that we have  $n$  models  $M_0, M_1, \dots, M_n$ . For each image we maintain  $c$  components with their location denoted by  $x_{ij}$ ,  $i = 1 \dots n, j = 1 \dots c$ . When a new image is available, the location and the uncertainty for each component and for each model is estimated. This step can be done using several techniques such as ones based on image correlation, spatial gradient or regularization of spatio-temporal energy. Using the VBDF technique, the result is the motion estimate  $x_{ij}$  for each component and

its uncertainty  $C_{ij}$ . Thus  $x_{ij}$  represents the location estimate of component  $j$  with respect to model  $i$ . The scale of the covariance matrix is also estimated from the matching residual errors. This will increase the size of the covariance matrix when the respective control point is occluded therefore occlusions are handled at the control point level.

The VBDF robust fusion technique is applied to determine the most relevant location  $x_j$  for component  $j$  in the current frame. The mode tracking across scales results in

$$\begin{aligned}\hat{x}_j &= C(\hat{x}_j) \sum_{i=1}^n \omega_i(\hat{x}_j) \hat{C}_{ij}^{-1} \hat{x}_{ij} \\ C(\hat{x}_j) &= \left( \sum_{i=1}^n \omega_i(\hat{x}_j) \hat{C}_{ij}^{-1} \right)^{-1}.\end{aligned}\quad (27)$$

with the weights  $\omega_i$  defined as in (26)

$$\omega_i(\hat{x}_j) = \frac{\frac{1}{|C_{ij}|^{1/2}} \exp\left(-\frac{1}{2} D^2(\hat{x}_j, \hat{x}_{ij}, C_{ij})\right)}{\sum_{i=1}^n \frac{1}{|C_{ij}|^{1/2}} \exp\left(-\frac{1}{2} D^2(\hat{x}_j, \hat{x}_{ij}, C_{ij})\right)} \quad (28)$$

and note that they satisfy  $\sum_{i=1}^n \omega_i(x) = 1$ .

It can be shown that the density corresponding to the point  $x + m(x)$  is always higher or equal to the one corresponding to  $x$ . Therefore iteratively updating the current location using the mean shift vector yields a hill-climbing procedure which converges to a stationary point of the underlying density.



The VBDF estimator finds the most important mode by iteratively applying the adaptive mean shift procedure at several scales. It starts from a large scale by choosing the parameter  $\alpha$  large with respect to the spread of the points  $x_i$ . In this case the density surface is unimodal therefore the determined mode will correspond to the globally densest region. The procedure is repeated while reducing the value of the parameter  $\alpha$  and starting the mean shift iterations from the mode determined at the previous scale. For the final step, the bandwidth matrix associated to each point is equal to the covariance matrix, i.e.,  $H_i = C_i$ .

The VBDF estimator is a powerful tool for information fusion with the ability to deal with multiple source models. This is important for motion estimation as points in a local neighborhood may exhibit multiple motions. The most significant mode corresponds to the most relevant motion.

An example for tracking multiple control point models will now be described in accordance with the present invention. There are  $n$  models  $M_0, M_1, \dots, M_n$ . For each image  $c$  control points with their location denoted by  $x_{i,j}$ ,  $i=1 \dots n$ ,  $j=1 \dots c$  are maintained. When a new image is available, the location and uncertainty for each control point and for each model are estimated. This step can be done using several techniques such as ones based on image correlation, spatial gradient or regularization of spatio-temporal energy. Using the VBDF technique, the result is the motion estimate  $x_{ij}$  for each control point and its uncertainty  $C_{ij}$ . Thus  $x_{ij}$  represents the location estimate of component  $j$  with respect to model  $i$ . The scale of the covariance matrix is also estimated from the matching residual errors. This will increase the size of the covariance matrix when the respective control point is occluded therefore occlusions are handled at the control point level.

The VBDF robust fusion technique is applied to determine the most relevant location  $x_j$  for component  $j$  in the current frame. The mode tracking across scales results in

$$\hat{\mathbf{x}}_j = \mathbf{C}(\hat{\mathbf{x}}_j) \sum_{i=1}^n \omega_i(\hat{\mathbf{x}}_j) \hat{\mathbf{C}}_{ij}^{-1} \hat{\mathbf{x}}_{ij}$$

$$\mathbf{C}(\hat{\mathbf{x}}_j) = \left( \sum_{i=1}^n \omega_i(\hat{\mathbf{x}}_j) \hat{\mathbf{C}}_{ij}^{-1} \right)^{-1} \quad (29)$$

with the weights  $\omega_i$  as defined in (28)

$$\omega_i(\hat{\mathbf{x}}_j) = \frac{\frac{1}{|\mathbf{C}_{ij}|^{1/2}} \exp\left(-\frac{1}{2} D^2(\hat{\mathbf{x}}_j, \hat{\mathbf{x}}_{ij}, \mathbf{C}_{ij})\right)}{\sum_{i=1}^n \frac{1}{|\mathbf{C}_{ij}|^{1/2}} \exp\left(-\frac{1}{2} D^2(\hat{\mathbf{x}}_j, \hat{\mathbf{x}}_{ij}, \mathbf{C}_{ij})\right)} \quad (30)$$

Following the computation of each control point, a weighted rectangle fitting is carried out with the weights given by the covariance matrix of the estimates. The image patches are related by a similarity transform  $T$  defined by 4 parameters. The similarity transform of the dynamic control point location  $\mathbf{x}$  is characterized by the following equations.

$$T(\mathbf{x}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mathbf{x} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (31)$$

where  $t_x$ ,  $t_y$  are the translational parameters and  $a$ ,  $b$  parameterize the 2D rotation and scaling.

The minimized criterion is the sum of Mahalanobis distances between the reference location  $\mathbf{x}_j^0$  and the estimated ones  $\mathbf{x}_j$  ( $j^{\text{th}}$  control point location in the current frame).

$$(\hat{\mathbf{x}}_j - T(\mathbf{x}_j^0))^T \mathbf{C}(\hat{\mathbf{x}}_j)^{-1} (\hat{\mathbf{x}}_j - T(\mathbf{x}_j^0)) . \quad (32)$$

Minimization is done through standard weighted least squares. Note that because the covariance matrix is used for each control point, the influence of points with high uncertainty is reduced.

After the rectangle is fitted to the tracked control points, the dynamic component candidate inside the rectangle is uniformly resampled. It is assumed that the relative position of each control point with respect to the rectangle does not change a lot. If the distance of the resample position and the track position computed by the optical flow of a certain control point is larger than a tolerable threshold, the track position is regarded as an outlier and replaced with a resampled point. The current image is added to the model set if sufficient control points have low residual error. The median residual error between the models and the current frame is compared with a predetermined threshold  $T_h$ .

Given a set of models  $M_0, M_1, \dots, M_n$  in which the component  $j$  has location  $\mathbf{x}_{ij}$  in frame  $i$ , our object tracking algorithm can be summarized by the following steps:

1. Given a new image  $I_f$  compute  $\hat{\mathbf{x}}_{ij}^{(f)}$  through robust optical flow [8] starting from  $\hat{\mathbf{x}}_{ij}^{(f-1)}$ , the location estimated in the previous frame;
2. For  $j = 1 \dots c$ , estimate the location  $\hat{\mathbf{x}}_j^{(f)}$  of component  $j$  using the VBDF estimator (Subsection 3.2) resulting in (6);
3. Constrain the component location using the transform computed by minimizing (9);
4. Add the new appearance to the model set if its median residual error is less than  $T_h$ .

The proposed multi-template framework can be directly applied in the context of shape tracking. If the tracked points represent the control points of a shape modeled by splines, the use of the robust fusion of multiple position estimates increases the reliability of the location estimate of the shape. It also results in smaller corrections when the shape space is limited by learned subspace constraints. If the contour is available, the model templates used for

tracking can be selected online from the model set based on the distance between shapes.

FIGs. 6a and 6b illustrate echocardiography images of a left ventricle in which movement of the endocardial wall is tracked in accordance with the present invention. Frame 1 for each figure 602, 610 illustrates the initial contour of the wall as shown by the dots in the image. The subsequent frames 604, 606, 608, 612, 614, 616 illustrate how the movement of the wall is tracked over time. Measurements for each dot are made in accordance with the present invention.

### Example

Hand traced heart contours in echocardiography images are used as the training set in this example. Both apical four-chamber (open contour) and parasternal short axis (closed contour) views are tested as illustrated in FIGs. 7a-7d and FIGs. 8a-8d. Landmark points are also annotated for each contour. FIG. 7b and FIG. 8b show the set of training contours drawn together. The eigenanalysis approach is applied where every contour is a vector with coordinates of the ordered landmark points as its components (34 components for the open contour and 36 for the closed contour). Then PCA is performed on the matrix whose columns are the training vectors. And the eigenvalues are used to form a diagonal matrix as the model covariance matrix in the subspace. The measurement covariance matrix is adjusted to test different scenarios.

FIGs. 7a and 8a illustrate test contours where the solid curves 704, 802 are the ground truth and the dashed curves 702, 804 are the noisy measurements. It is desired to find a contour in the eigen-contour subspace that is the closest to the current noisy contour, with more uncertain point adjusted more. FIG. 7c and FIG. 8c show the result when orthogonal projection is applied to the contours of FIG. 7a and FIG. 8a (with isotropic covariance depicted by small circles 712, 812 around the landmark points). FIG. 7d and FIG. 8d show the results. It is seen that the result is much closer to the ground truth. It is not a perfect fit because the training data is rather different from the test data,

therefore, some small shape deformation in the test contour may not be realizable in the subspace.

In accordance with the present invention, the contours can be viewed using a visualization tool to facilitate the diagnosis. To make the doctor diagnose the heart diseases from the echocardiograms more easily, LV contours are displayed by color vectors in real time. The length of the vectors represents the displacement magnitude while the direction of the vectors represents the movement direction of the points on the LV wall contours. For exemplary purposes, the color orange is used to depict when the motion is contraction and the color dark blue is used to depict when the motion is dilation. FIG. 9 shows one of the representing results. The movements of the points on the LV contours are tracked. The movements are smoothed by Gaussian in both temporal and spacious domain. To make it easier seen by the doctors, the length of the vectors is magnified by three times.

By this visualization method, the doctor can easily see the motion of each segment of the endocardium with the magnitude and the directions. Combination of the global motion compensation with this visualization method can make the doctors easily see the contraction direction and magnitude in all the segments of the LV. An ischemic region or other abnormal regions can be identified directly and easily by human eyes.

Having described embodiments for method for determining feature sensitivity during a medical examination, it is noted that modifications and variations can be made by persons skilled in the art in light of the above teachings. It is therefore to be understood that changes may be made in the particular embodiments of the invention disclosed which are within the scope and spirit of the invention as defined by the appended claims. Having thus described the invention with the details and particularity required by the patent laws, what is claimed and desired protected by Letters Patent is set forth in the appended claims.

We claim:

1. A method for tracking a global shape of an object in motion, the method comprising the steps of:
  - defining one or more control points along the global shape;
  - tracking each of the one or more control points as the object is in motion;
  - representing uncertainty of a location of a control point in motion;and
  - exploiting the uncertainty to constrain the global shape using a prior shape model.
2. The method of claim 1 wherein the step of tracking each of the one or more control points employs a Bayesian kernel matching approach.
3. The method of claim 1 wherein the step of tracking each of the one or more control points employs an optical flow based approach.
4. The method of claim 1 further comprising the step of:
  - remedying error accumulation during tracking of the one or more control points by using a multi-template adaptive matching framework.
5. The method of claim 1 wherein the prior shape model is a subspace shape constraint model.
6. The method of claim 1 wherein a strongly adapted Principal Component Analysis (SA-PCA) model is used to constrain the shape.
7. The method of claim 1 wherein the step of exploiting the uncertainty further comprises using a covariance matrix to employ a subspace shape constraint model using a non-orthogonal projection or information fusion.

8. A system for tracking a global shape of an object in motion comprising:
- means for defining one or more control points along an initial contour of the global shape;
  - means for tracking each of the one or more control points as the object is in motion;
  - means for representing uncertainty of a location of a control point in motion;
  - means for exploiting the uncertainty to constrain the global shape using a prior shape model.
9. The system of claim 8 wherein the tracking means employs a Bayesian kernel matching approach.
10. The system of claim 8 wherein the tracking means employs an optical flow based approach.
11. The system of claim 8 further comprising:
- means for remedying error accumulation during tracking of the one or more control points by using a multi-template adaptive matching framework.
12. The system of claim 8 wherein a strongly adapted Principal Component Analysis (SA-PCA) model is used to constrain the shape of the contour.
13. The system of claim 8 wherein the step of exploiting the uncertainty further comprises using a covariance matrix to employ a subspace shape constraint model using a non-orthogonal projection or information fusion.
14. A method for visually tracking movement of a shape of an object comprising the steps of:

generating one or more first color vectors to represent contraction of control points along the contour of the shape;

generating one or more second color vectors to represent dilation of control points along the contour of the shape;

periodically displaying the first and second color vectors thereby displaying movement of the shape.

15. A system for visually tracking movement of a shape of an object comprising:

means for generating one or more first color vectors to represent contraction of control points along the contour of the shape;

means for generating one or more second control vectors to represent dilation of control points along the contour of the shape;

means for periodically displaying the first and second color vectors thereby displaying movement of the shape.

16. A method for tracking a global shape of an object in motion, the method comprising the steps of:

defining one or more control points along the global shape;

tracking each of the one or more control points as the object is in motion;

building multiple appearance models for each control point; and  
combining motion vectors produced by each model for tracking the shape.

17. The method of claim 16 further comprising the step of:  
using robust information fusion to compute an estimate of the location of each control point.

18. The method of claim 16 wherein the multiple appearance models are built based on uncertainty of a location of one or more of the control points.



19. The method of claim 18 wherein the robust information fusion is Variable-Bandwidth Density-based Fusion (VBDF).

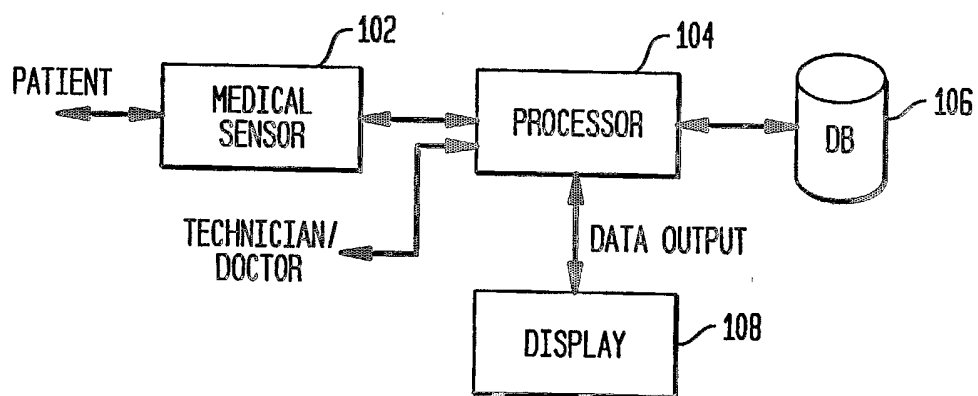
20. The method of claim 19 wherein robust information fusion comprises the steps of:

inputting a number of measurements represented by mean vectors and covariance matrices;

fusing together the measurements by computing variable bandwidth density estimates; and

detecting a most significant node of the variable bandwidth density estimate.

FIG. 1



2/10  
FIG. 2

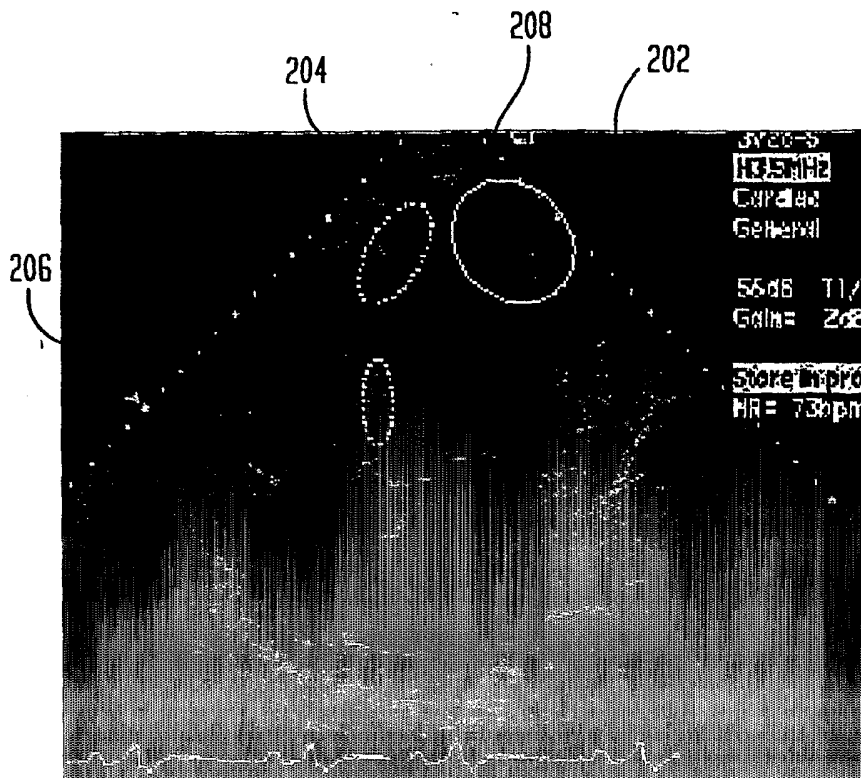


FIG. 3

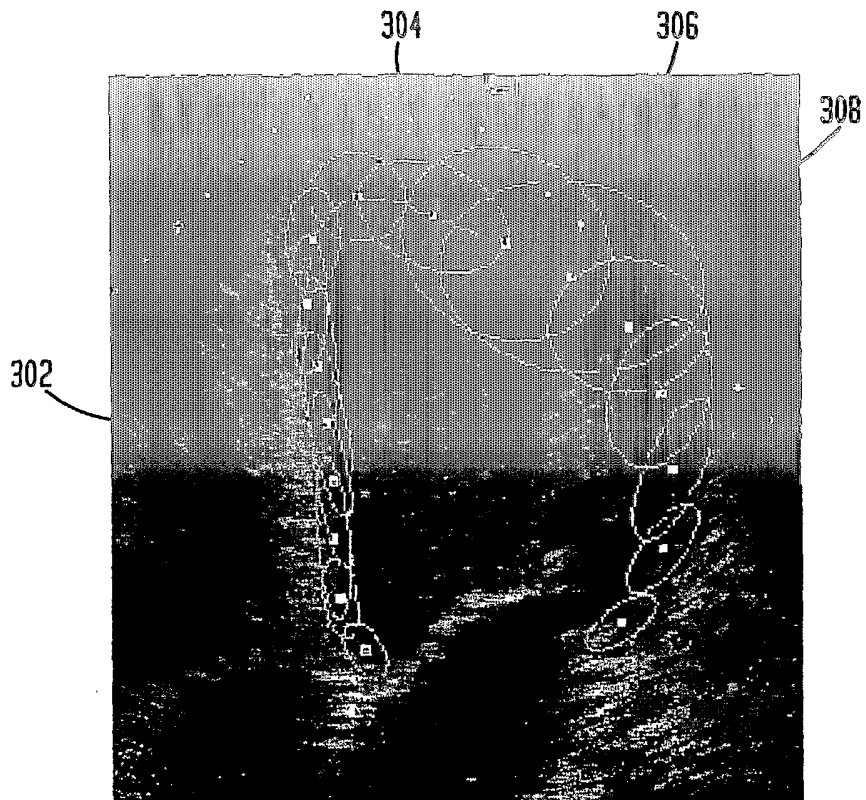
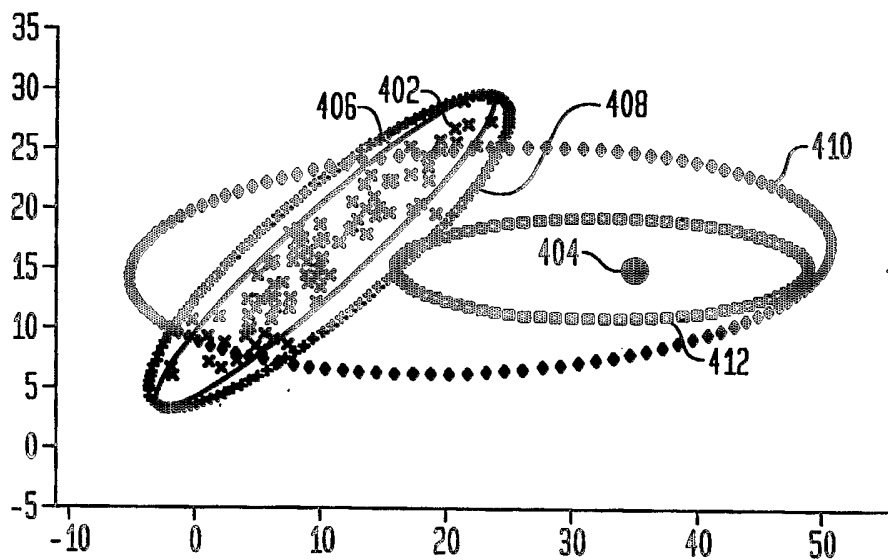
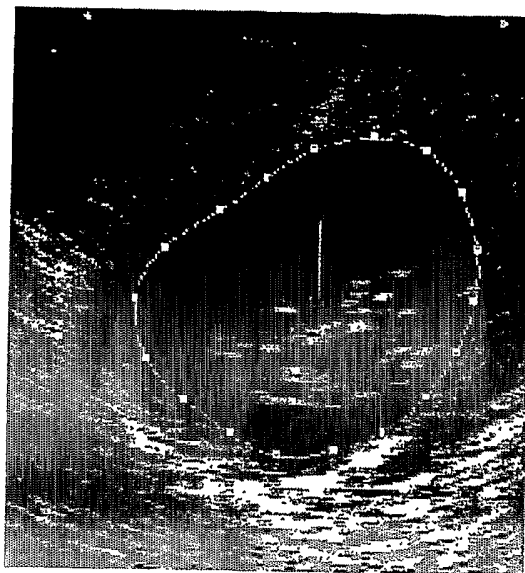


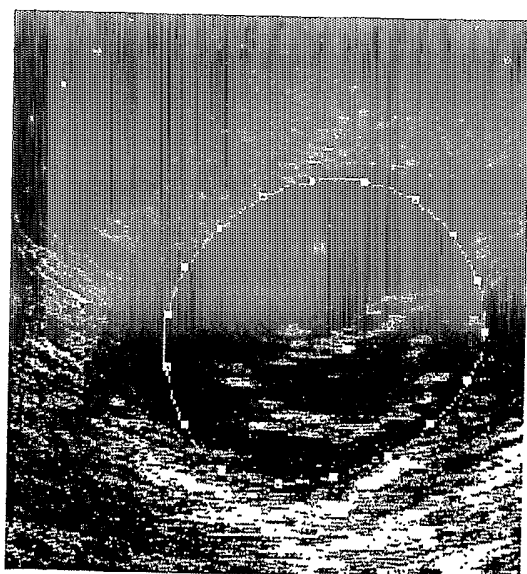
FIG. 4



**FIG. 5A**



**FIG. 5B**



**FIG. 5C**

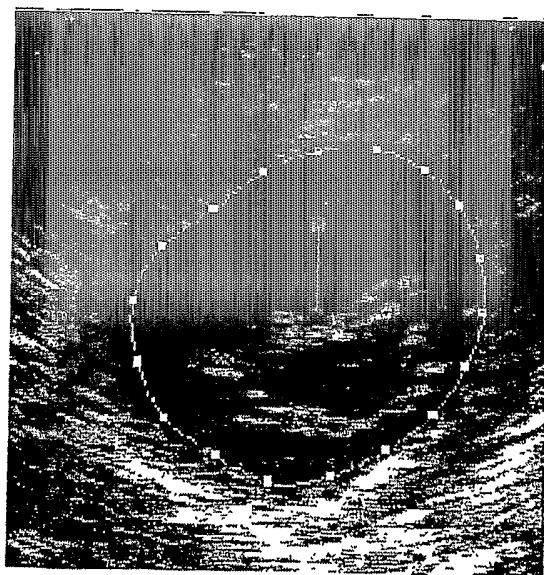


FIG. 6A

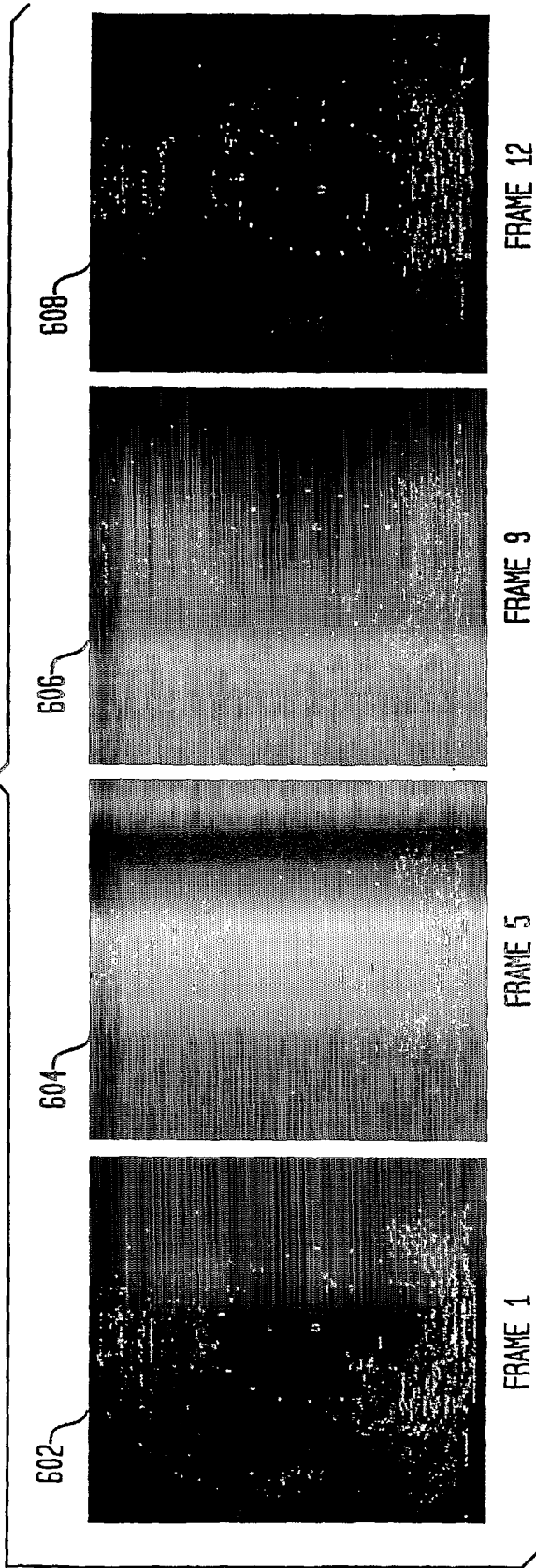


FIG. 6B

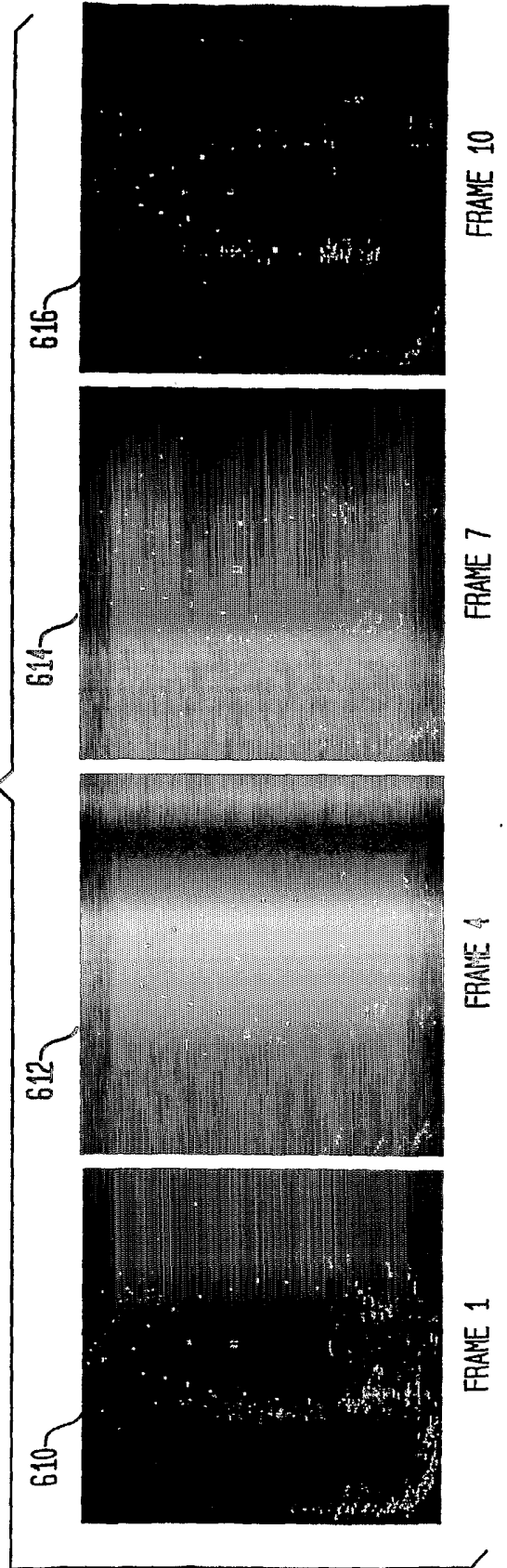


FIG. 7A

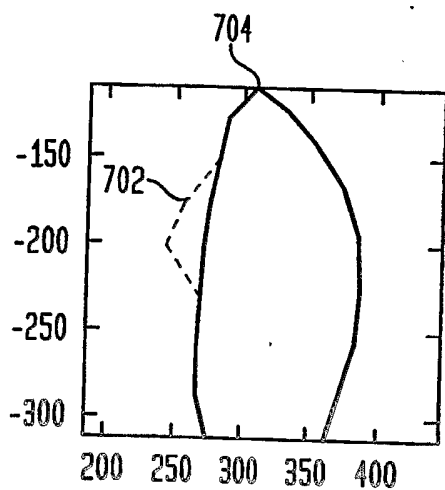


FIG. 7B

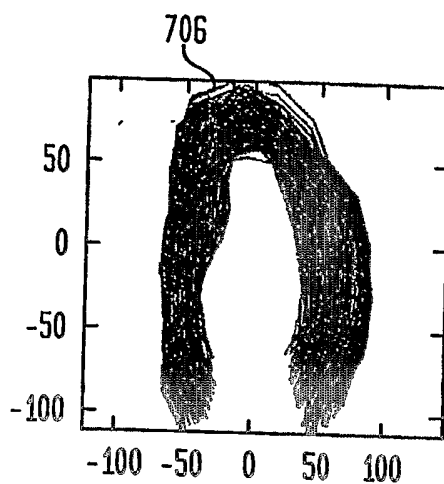


FIG. 7C

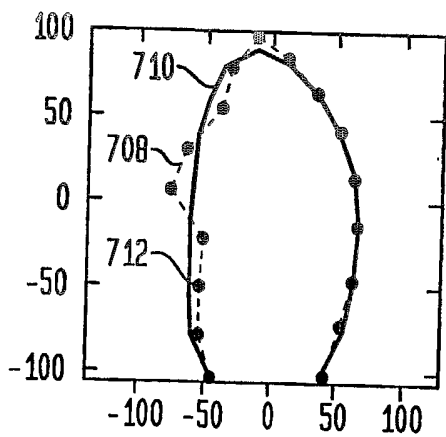
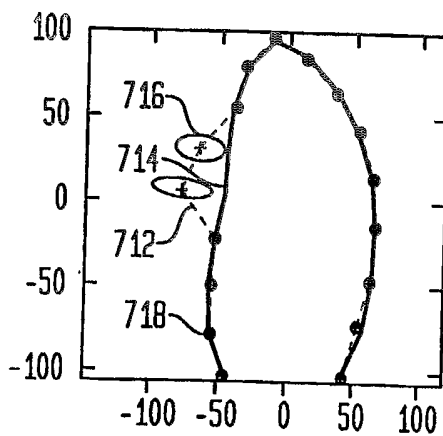
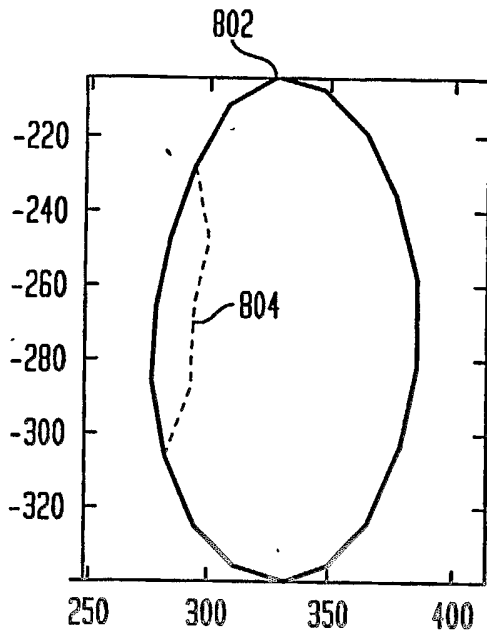


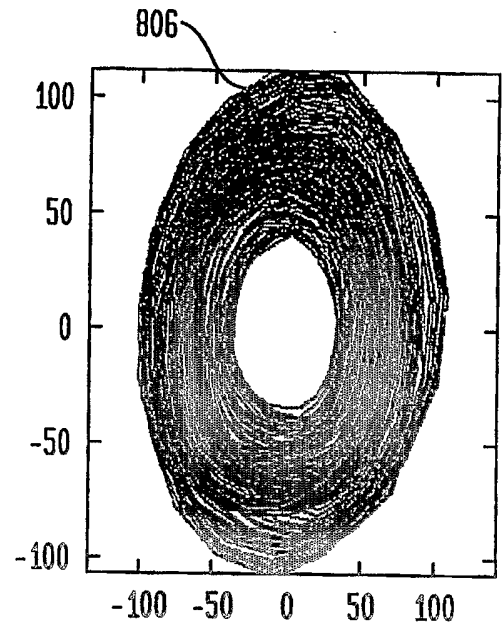
FIG. 7D



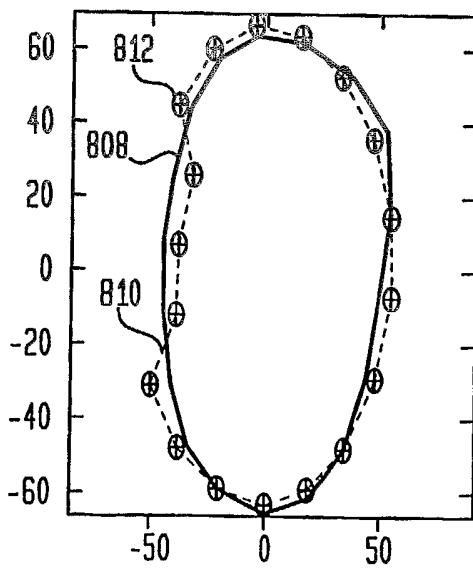
**FIG. 8A**



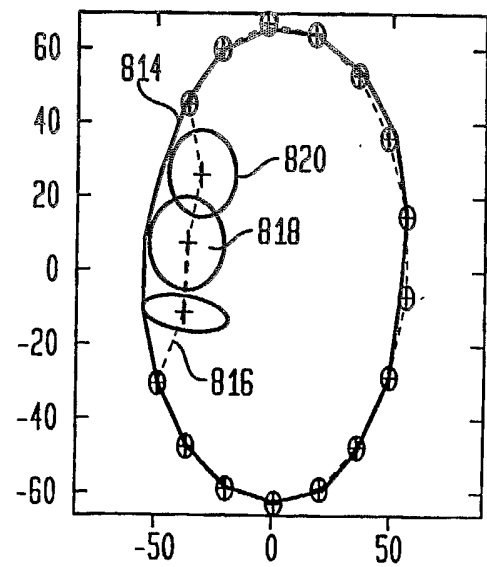
**FIG. 8B**



**FIG. 8C**



**FIG. 8D**





8/10

FIG. 9A

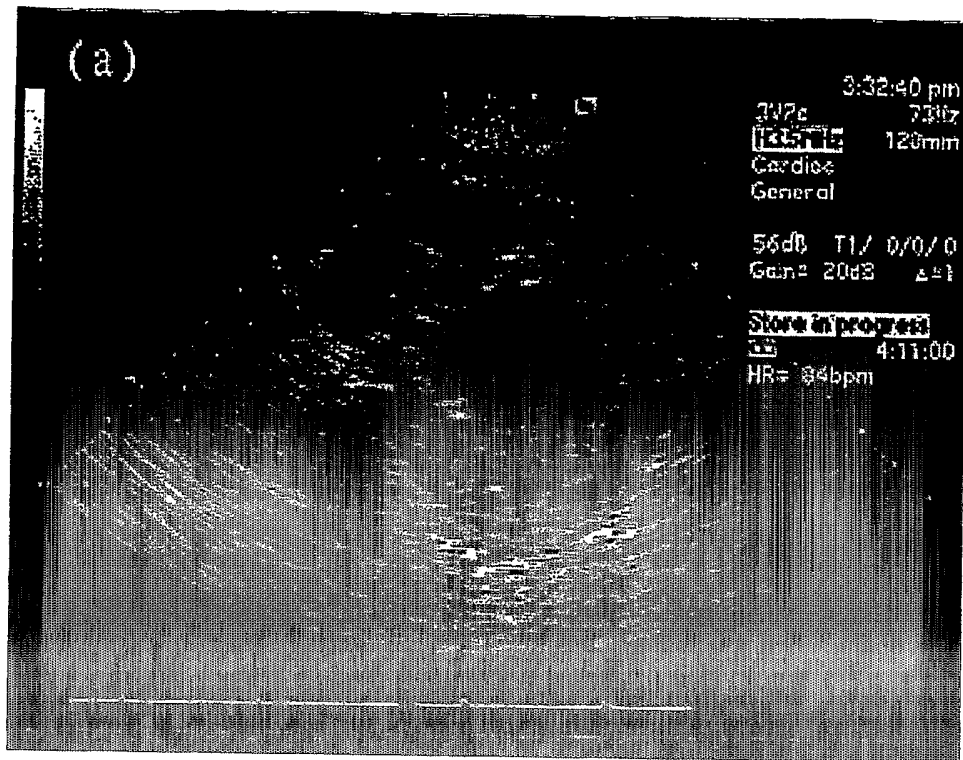


FIG. 9B

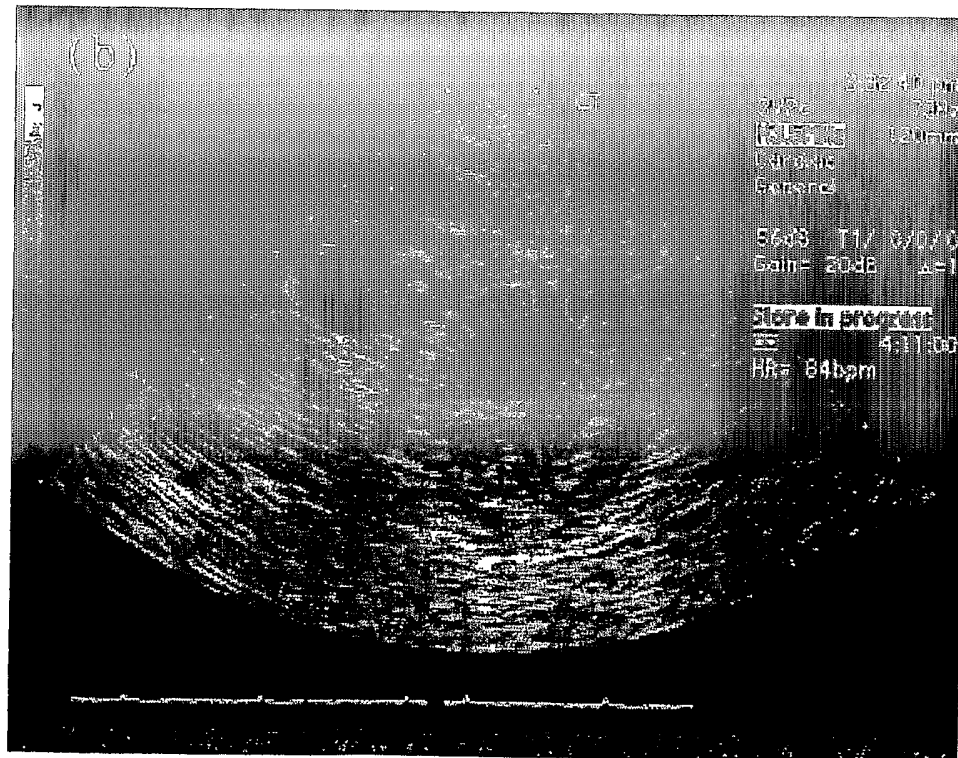


FIG. 9C

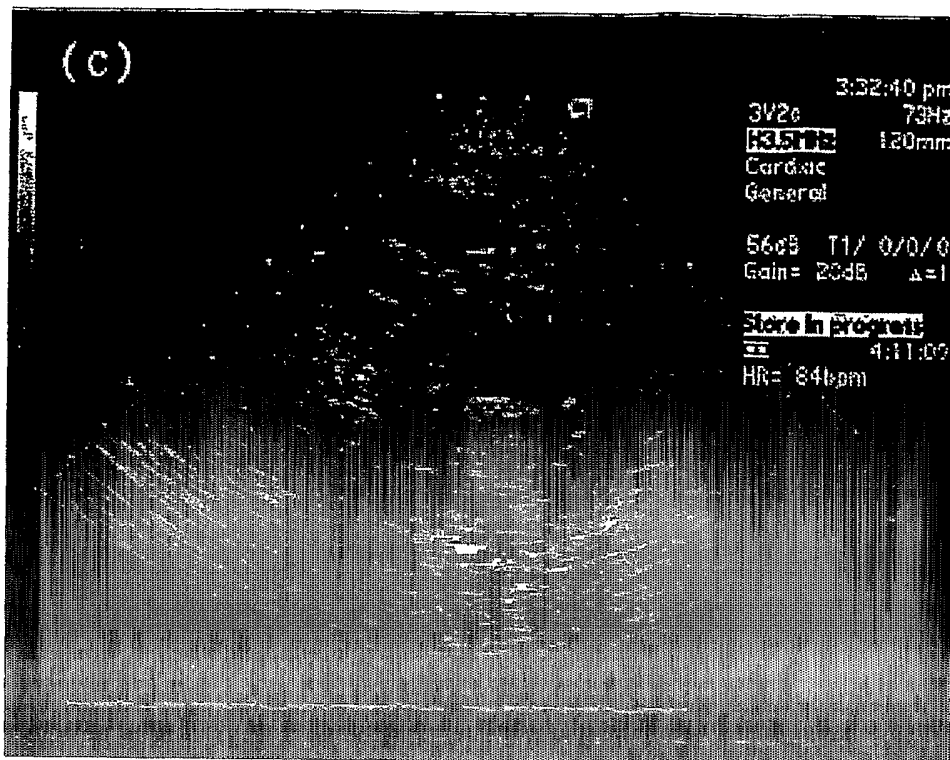


FIG. 9D

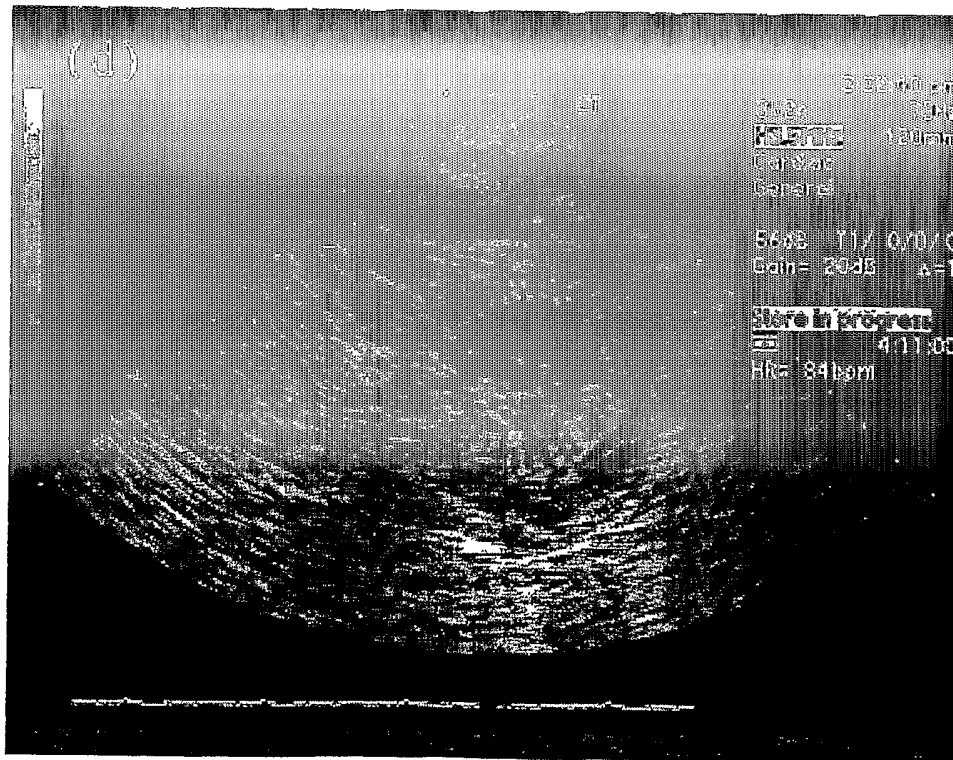


FIG. 10

