Shape-based Similarity Retrieval of Doppler Images for Clinical Decision Support

T. Syeda-Mahmood[1], P. Turaga[2], D. Beymer[1], F. Wang[1],A. Amir[1], H. Greenspan[1],K. Pohl[1]

[1]IBM Almaden Research Center, 650 Harry Road

[2]University of Maryland, College Park

stf@almaden.ibm.com,pavan.turaga@gmail.com,beymer,wangfe,arnon,hayitg,pohl@almaden.ibm.com

http://www.almaden.ibm.com/cs/projects/aalim *

Abstract

Flow Doppler imaging has become an integral part of an echocardiographic exam. Automated interpretation of flow doppler imaging has so far been restricted to obtaining hemodynamic information from velocity-time profiles depicted in these images. In this paper we exploit the shape patterns in Doppler images to infer the similarity in valvular disease labels for purposes of automated clinical decision support. Specifically, we model the similarity in appearance of Doppler images from the same disease class as a constrained non-rigid translation transform of the velocity envelopes embedded in these images. The shape similarity between two Doppler images is then judged by recovering the alignment transform using a variant of dynamic shape warping. Results of similarity retrieval of doppler images for cardiac decision support on a large database of images are presented.

1. Introduction

With more and more patient records now containing multimodal imaging data, an exciting application of image and video retrieval is emerging in the area of clinical decision support. Cardiologists in particular, routinely use multiple imaging modalities including X-ray imaging, ultrasound imaging, and CT imaging for their decision making. However, their diagnosis methodology is still single sample-guided in that only the data from the given patient is used along with their a priori knowledge to make decisions. If content-based retrieval techniques could be used to retrieve similar case data and hence similar patients, it can enable enhanced decision making for physicians. For example, using similar case data, physicians can validate their current hypothesis. Further, by examining the associ-
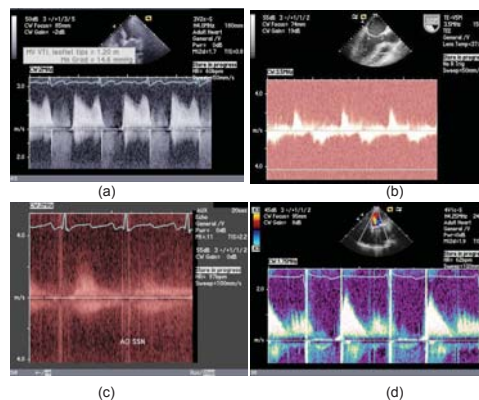

Figure 1. Illustration of CW Doppler images. (a) M pattern in flow Doppler profile typical of Mitral stenosis. (b) Baseline need not be distinctive. (c) Doppler profile has visibility problems. The region is a zoomed-in view. (d) Aliasing artifacts alter the pattern.

ated diseases with the similar patient cases retrieved, they can check for any overlooked possibilities or alternate interpretations. Finally, they can learn of statistical correlations (or co-morbidities) between diseases, treatment and outcomes, thus paving the way for a whole new way of practicing medicine.

While structured clinical information such as demographics, and vital signs is clearly relevant, the most challenging aspect of finding similar cases is determining similarity in unstructured modality data. What makes two X-ray images, or two echocardiogram videos similar is not their color or texture but the underlying disease they depict. Thus image and video retrieval methods would need to focus on disease-specific patterns for finding similar cases.

In this paper, we attempt to address this challenging application of clinical decision support using disease-specific shape-based retrieval of similar cases on flow Doppler imaging. Flow Doppler imaging, recorded during an

echocardiographic exam, is widely used by clinicians to detect diseases of the valves such as Aortic Regurgitation, Aortic Stenosis, Mitral regurgitation, and Mitral stenosis[6]. In the continuous wave Doppler (CW) mode scan [6], the shape of the Doppler signal tracings convey information about the functioning of the various valves and arterial structures. Figure 1a shows one such velocity profile with the shape representing mitral stenosis. Traditionally, clinicians have manually traced such velocity envelopes to extract measurements such as decay time, pressure gradient, pressure half-time, and velocity time integral which are then matched to normal and abnormal values based on clinical guidelines. Automated analysis of Doppler imaging has so far been predominantly focused on automatic tracing of the velocity-time profiles to provide more accurate measurements for echo-cardiographers. The identity of the valve being depicted in the Doppler images is assumed to be known in such cases.

In this paper we utilize Doppler imaging in a new way to enable automated decision support. Specifically, we observe that different valvular diseases appear as characteristic shape patterns in Doppler images. *By measuring the similarity in the shape pattern conveyed within the velocity region of two Doppler images, we can infer the similarity in their diagnosis labels.* Figure 2 illustrates the discriminability of diseases in CW Doppler imaging. Figure 2a-b show examples of velocity patterns from moderate and severe mitral stenosis. Similarly, Figure 2c-d show examples of aortic stenosis and regurgitation respectively. Finally, Figure 2e-f show examples of tricuspid regurgitation and pulmonary regurgitation. As can be seen, each of these patterns is discriminative. Further, members of the same disease class exhibit remarkable similarity in appearance as shown in Figure 2i-j which shows two instances of moderate pulmonary valve regurgitation. Thus it is plausible that the disease similarity can be inferred by developing a measure for capturing the visual similarity of the velocity region within Doppler images.

Using fully automated processing to discover similarities in Doppler images is, however, a challenging task. First, reliable pre-processing is needed to separate the Doppler frames from the rest of the frames in an echocardiogram video recording and the relevant region containing velocity profiles has to be isolated within these frames. Next, the detection of similarity should account for variations in heart rate and signal intensity. Finally, the similarity measure should be robust to individual inter-patient variations in the shape profile within the same disease class as shown in Figure 2i-j, while still being able to discriminate between mild, moderate and severe cases of the disease, as shown in Figure 2a-b and g-h. Here Figure 2a-b, g-h show moderate and severe cases of mitral regurgitation and mild and moderate cases of mitral stenosis, respectively.

The rest of the paper describes our approach to finding similar Doppler images that addresses the above issues. It makes several novel contributions. To our knowledge, this is the first work to address disease recognition and retrieval from echo Doppler patterns. Although clinicians are implicitly aware and often 'eye ball' such patterns, systematic correlation between diseases and shape patterns in Doppler velocity flows has not yet been documented well, even in medical text books. Secondly, the method of non-rigid shape matching of flow velocity envelopes is easily applicable for other time series where internal variations between members of a class can be modeled under a similar transform. By using this shape modeling approach, built-in verification is achieved as in object recognition due to the recovery of the registration parameters during shape similarity-based retrieval. Finally, unlike most medical image retrieval methods focused on theoretical evaluation on benchmark datasets that stop at retrieving images, we take to the next step by evaluating the utility of such retrieval for clinical decision support.

1.1. Background

In Doppler imaging, ultrasound waves of a known frequency are transmitted and the amplitude and frequency of the received signal is recorded. Due to the Doppler effect, the motion of blood and tissue within the heart's chambers induces frequency shifts between the transmitted and received signals [6]. Since blood moves at much higher velocities than the surrounding heart tissue, it induces higher frequency shifts than tissue. Thus, high-pass filtering of the received signal eliminates the response from the surrounding tissue and provides information exclusively of blood-flow. While there are different modes of Doppler imaging, including 2D, color, PW,etc. [6], the CW Doppler has become popular due to its high temporal and velocity resolution as it avoids aliasing through continuous scanning. In the CW mode scan, the shape of the Doppler signal tracings convey information about the functioning of the various valves and arterial structures.

Figure 1 shows sample Doppler images taken during an echocardiogaphic exam of patients and illustrates the challenge these images pose both from the point of image processing and pattern recognition. The Doppler velocity flow pattern is captured in a box-like region in each of the examples in this figure. As can be seen, the location and size of the containing region can change based on machine and nature of exam (eg. zoom into a region). The Doppler flow pattern is sometimes clearly apparent as in Figure 1a with the distinctive M-shape for mitral stenosis, while in other cases, it is difficult to even see the pattern as in Figure 1c. Velocity flow spreads in both directions from a baseline, which is also difficult to detect as seen in Figure 1b where the salient horizontal line is not necessarily the baseline. Fi-
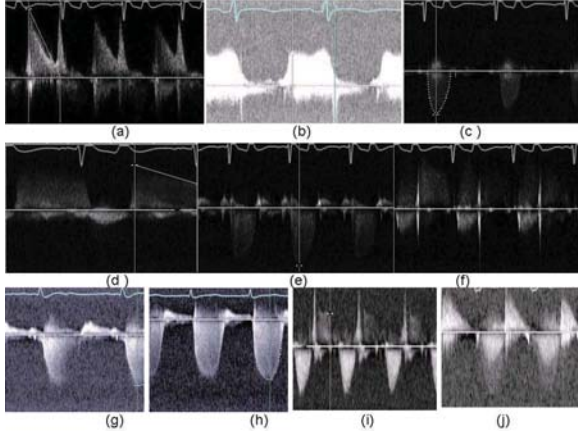
Figure 2. Illustration of shape patterns conveyed by Doppler images for various cardiac diseases.(a)-(b) Mild and moderate Mitral Stenosis. (c)Aortic Stenosis. (d) Aortic regurgitation. (e)Tricuspid regurgitation. (f) Pulmonary regurgitation. (g)-(h) Moderate and severe Mitral regurgitation. (i)-(j) Moderate pulmonary regurgitation.

nally, aliasing effects may still be present in severe cases of disease as seen in Figure 1d where the vertical streaks in velocity are artifacts that need to be filtered.

2. Related Work

The medical image retrieval community has been active since the early days of content-based retrieval [14, 5, 8] with ImageCLEF [2, 10] now offering reference X-ray collections for benchmarking medical retrieval algorithms. Very little of this work, however, has moved to clinical practice for actual decision support and has focused mainly on image classification[2]. For clinical decision support, similarity ranking needs to advance to the next step of integrating with associated disease, treatment and outcome data with patients (in electronic records) in order to validate diagnosis, learn about alternatives and co-morbidities. Treating this as a pure image classification is not sufficient since in reality patients have multiple diseases whose combined effect on the diagnostic image appearance can be quite complex and would need considerably large number of samples for training.

Much of the work on Doppler image analysis has been on automatically tracing the velocity time profiles. Early work attempted to extract the E-wave of the Mitral valve profile using neural networks[11]. The velocity curve tracing was attempted using edge detection algorithms for the brachial artery Doppler images in [15]. Later work applied contour tracing for Mitral and tricuspid valves in[13] to extract clinical parameters such as peak velocity, velocity-time integral, etc. In a recent work [7], the tracing of mitral valve inflow Doppler spectra was combined with segmentation of mitral inflow structures. Since different measurements are

made for different valves, these approaches assume that the valve identity is known. Further, the approach is specific to a disease or cardiac structure under study as described in the work of [4, 9]. Similarly, [3] focus on diseases of the Mitral valve by extracting the pressure-gradient from the Doppler images and [13] focus on extraction of velocity envelopes in cases of atrial fibrillation. A more recent work proposed a machine learning algorithm for automatically tracing the envelopes of specific spectra such as the mitral valve inflow Doppler spectra[7]. In contrast, we present a comprehensive approach to extracting velocity profiles independent of the disease while performing end-to-end processing of actual echo studies of patients.

3. Feature extraction from Doppler images

The input in our case is a full echo study available as an echocardiogram video. From this, the Doppler frames need to be separated from other frames in an echocardiogram video that depict moving heart regions, and textual measurement-only frames. To isolate the Doppler frames, we build rectangular templates capturing the Doppler region through a training process using echo frames from various echo machines (Siemens Sequoia, Siemens Cypress, etc.) that capture the expected position and size of the Doppler regions in an echo video frame. By applying the templates, all Doppler frames in an echo video were isolated. To distinguish CW Doppler frames from PW and other Doppler, we applied an optical character recognition (OCR) engine (Tesseract) in a band above the rectangular regions isolated, to look for appropriate keywords such as 'CW' or 'PW'. This template matching step also gives the Doppler containing region as shown in Figure 3b for the raw frame in Figure 3a. The selected region may contain measurement panels (eg. pop-up box on upper left in Figure 3a) that are adjacent or overlapping the Doppler signal containing region.

3.1. Extracting velocity envelopes

To highlight the velocity profiles within the selected region, we apply a foreground separation from background thresholding algorithm[12]. This classic algorithm called Otsu's thresholding, calculates the optimum threshold separating the foreground and background classes so that their combined spread (intra-class variance) is minimal using the shape of the histogram of pixel intensities. The algorithm however leaves salt and pepper noise in the resulting images as shown in Figure 3c for the selected region in Figure 3b. By retaining large perceivable regions (greater than 50 pixel area in this case) ensures that much of the velocity signal is captured as shown in Figure 3d. A morphological close step is applied to fill up the small holes in the bright regions to yield a clean velocity signal containing region as shown in Figure 3e for the image in Figure 3d. We then trace the

boundary pixels of all white regions to get contours such as those shown in Figure 3f. As can be seen the velocity profile is contained in these boundaries, although the adjacent ECG trace from above as well as calibration axes are highlighted as well. To extract the final velocity envelope, we retain strong boundaries that are on either side of the baseline. This is the horizontal axis line shown in all images of Figure 3. As this line can be sometimes occluded by the measurement bars (eg. vertical bars in Figure 3b), measurement screens, or artifacts, simple image processing is not sufficient to detect the baseline. Our approach exploits the velocity unit marker on the left denoted by 'm/s' which is always positioned at the axis line. For this we build templates for the region containing the m/s symbol and use template matching to isolate its location in the Doppler frame. The y coordinate of this region is then used as an estimate of the baseline. Finally, to remove the effect of spiking artifacts that may be embedded within the velocity signal (eg. Figure 1d), we employ a temporal median filter of a window size of 3 pixels on the extracted envelopes.
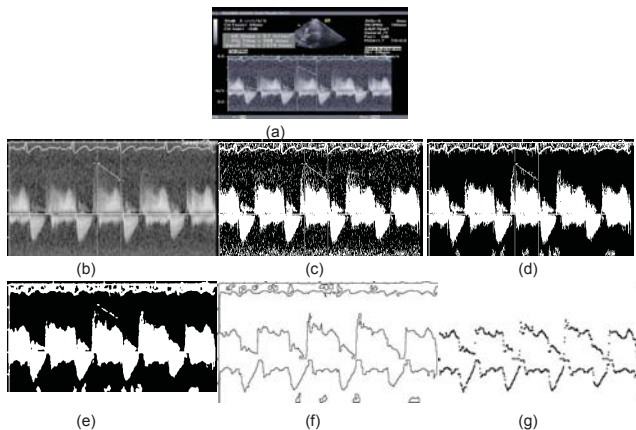

Figure 3. Illustration of velocity envelope extraction from Doppler images. (a) Original Doppler image frame from an echocardiogram video sequence. (b) Doppler velocity containing regions retained in the original image (c) Image of (b) thresholded to separate the foreground.(d) Largest region retained. (e) morphological close operation to fill up holes in largest region. (f) Boundary pixels of the largest region. (g) velocity upper and lower envelopes.

Detecting heart cycles: To make the subsequent matching invariant to heart rate, we segment the velocity envelopes into individual heart beat cycles. Reliable estimation of heart rate from these imagery is difficult particularly due to measurement overlays and noise in Doppler region. We developed a robust technique to detect the heart rate using several confirmation sources. The first estimate of heart rate comes from measurement supplied by the echo machine itself, overlayed on the echo frame, as, for example, HR=54 bpm, in the right column of text in Fig. 4a. Template matching first isolates the text "HR=", and then OCR is used
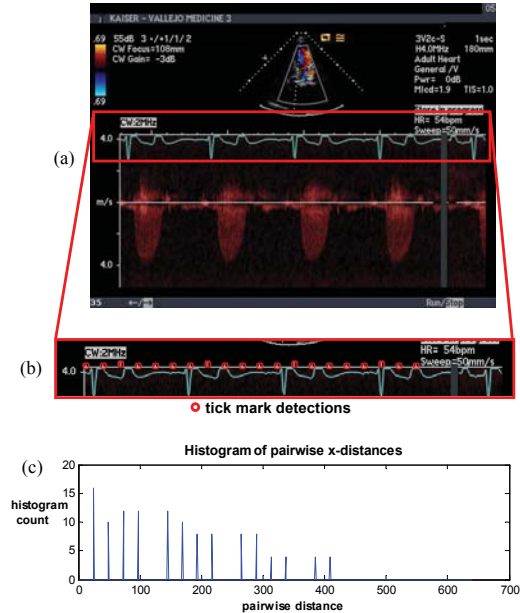

Figure 4. To estimate signal periodicity, we combine the measured heart rate using OCR with a parsing of the tick marks of the Doppler plot's horizontal axis to map time to pixels. For original image (a), tick marks are detected (b), and a histogram (c) of the x-distance between all pairs of detections is computed. The inter-tick distance is the first signficant peak in the histogram.

on a bounding box just to the right to extract the heart rate in beats per minute. To convert the heart rate from beats per minute to actual pixel widths in the Doppler region, we detect the calibration markers on the top horizontal axis of the Doppler region (see Figure 4a). The inter-marker spacing is always 200 ms, so that estimating this distance d in pixels is sufficient for the transformation of time to pixel width. Here again, we employ a template matching approach to detect the calibration markers, and form a histogram of the difference in x-coordinates between all pairs of detections. The histogram is dominated by d and integer multiples of d, so we process the histogram to find the first significant peak. Fig. 4 shows a raw image (a), the tick detections in red circles (b), and the x-distance histogram (c).

As a second approach, we used the velocity signal region contained within the Doppler velocity envelope found in Section 3.1 directly for autocorrelation-based periodicity estimation. The Doppler envelopes are then segmented into distinct heart cycles using the pixel width information. To normalize for intensity and heart rate differences, all unit cycle envelopes are resampled to a fixed sample width. To avoid aliasing, the sample width should be larger than the longest time period that can be found in these images. Since the Doppler image size is typically 580 pixel wide, and there is at least one cycle of velocity flow captured, we take a uniform sampling of 500 samples.

3.2. Feature description of velocity envelopes

Important fiducial points are extracted from the upper and lower envelopes. For this, a simple line segment approximation of the envelope curves is achieved through a recursive partitioning of the curve. A threshold on minimum length = 5 pixels, and amplitude deviation of 0.01 of the normalized profile was found sufficient to remove much of the noise during tracing. The fiducial points are then chosen as corners in the line segment approximation where the curvature changes significantly (2 degrees-178 degrees). Each corner is then described using its parametric(t,f(t)) position, included angle θ, and the orientation of the bisector ϕ as $s(t) = < t, f(t), \theta, \phi >$. Using the angle of the corner ensures that the sharpness of the envelopes is retained as a feature. This will help discriminate between mild, moderate and severe cases of the disease. The orientation of the bisector, on the other hand, ensures that the errors due to polarity reversals in the envelope shape are avoided, thus improving the distinction between some regurgitation and stenosis patterns (where direction of the lobe is important).

4. Matching Doppler Envelopes

We now address the problem of matching Doppler images. The goal of the matching is to recognize the inherent shape pattern characterizing diseases (or combinations) by modeling the overall perceptual similarity in appearance within members of the same disease class. Such a matching should not only be invariant to heart rate, signal intensity, imaging artifacts, but also patient-specific details that cause the onset of different events within a heart cycle to shift in time in a non-rigid fashion. Further, even though the individual velocity envelopes have been segmented into single heart cycles, the start of the cycles are not necessarily synchronized. Thus the envelopes have to be circularly rotated to be brought into correspondence prior to matching. When the candidate envelope is a matching envelope, a simple 1D correlation can recover this time shift. The rotation of the signal, however, must be circular to preserve the shape.

Assuming such rotation has taken place, we model the intra-class shape variations that maintain the overall similarity in perceptual appearance as a constrained non-rigid translation transform of the parametric representation of the envelope curves. For our purpose, a disease class X is a 3-tuple (disease type, valve type, severity level). Let the upper and lower envelope curves of a single period Doppler image segment be represented by $G(t) = < g_l(t), g_h(t) >$ corresponding to disease class X. Consider another (single period segment) Doppler image $F(t) = < f_l(t), f_h(t) >$ that is a potential match to $G(t)$ corresponding to the same disease. We model the relation between $G(t)$ and $F(t)$ as a constrained non-rigid transform $[a, b, \Gamma]$ such that

$$|F'(t) - G(t)| \leq \delta \qquad (1)$$

where the absolute symbol represents the distance metric that measures the difference between $F'(t)$ and $G(t)$, the simplest being the Euclidean norm and with

$$F'(t) = AF(\Phi(t)) \quad \text{with} \quad \Phi(t) = bt + \Gamma(t) \qquad (2)$$

The term $A = [a_1 \ a_2]^T$ is a scaling transform to align the two Doppler envelopes, b is a term (common for both envelopes) to reflect the time scale differences due to heart rates, and $\Gamma(t)$ is to model the non-rigid time shifts of the envelope due to differences in systolic and diastolic components within individual heart cycles among patients within the same disease class.

The parameters A and b can be solved by normalizing in amplitude and time. That is, if we transform $F(t)$ and $G(t)$ such that and

$$\hat{F}(t) = \frac{F(t) - F_{min}(t)}{F_{max}(t) - F_{min}(t)} \quad \text{and} \qquad (3)$$

$$\hat{G}(t) = \frac{G(t) - G_{min}(t)}{G_{max}(t) - G_{min}(t)}$$

then $a_1 = a_2 = 1$. Here G_{min} and G_{max} are the vectors of minima in each envelope. To eliminate solving for b, we can normalize the time axis, so that all time instants lie in the range $[0, 1]$. Since the single period segments were already scaled to a fixed pixel width prior to shape matching, we already have $b = 1$ in our case.

Recovering Γ

We recover the non-uniform time translation Γ using a variant of dynamic time warping (DTW). Instead of using time and derivative of curve as a constraint for DTW, we exploit detailed shape information of fiducial features such as angle of turn and orientation of bisector. This helps anchor the warping to important fiducial points along the curve. The non-rigid deformation of time for all other intermediate points can be recovered through time interpolation. This makes DTW not only efficient (as there are fewer fiducial points) but also more accurate as detailed shape information is taken into account for matching.

Let there be K features extracted from $\overrightarrow{F}(t)$ as $F_K = \{(t_1, \overrightarrow{f_1}(t_1), (t_2, \overrightarrow{f_2}(t_2), ..., (t_K, \overrightarrow{f_K}(t_K))\}$ at time $\{t_1, t_2, ..., t_K\}$ respectively. Similarly, let there be M fiducial points extracted from $G(t)$ as $G_M = \{(t'_1, \overrightarrow{g_1}(t'_1), (t'_2, \overrightarrow{g_2}(t'_2), ..., (t'_M, \overrightarrow{g_M}(t'_M))\}$ at time $\{t'_1, t'_2, ..., t'_M\}$ respectively. If we can find a set of N matching fiducial points $C_\Gamma = \{t_i, t'_j\}$, then the non-uniform translation transform Γ can be defined as:

$$\Gamma(t') = \begin{cases} t_i & \text{if } t = t'_j \text{ and } (t_i, t'_j) \in C_\Gamma \\ t_r + \frac{t_s - t_r}{t'_l - t'_k}(t' - t'_k) & \text{where } (t_r, t'_k), (t_s, t'_l) \in C_\Gamma \end{cases}$$
$$(4)$$

and t'_k is the highest of $\{t'_j\} \leq t'$ and is the lowest of $\{t'_j\} \geq t'$ that have a valid mapping in C_Γ. Other interpolation

methods besides linear (eg. spline) are also possible. Using Equations 1 and 5, the shape approximation error between the two Doppler signals is then given by:

$$|F'(t) - G(t)| = |\overrightarrow{F}(t' = \Gamma(t')) - G(t')| \qquad (5)$$

For each $G(t)$, we would like to select Γ such that it minimizes the approximation error in (6) while maximizing the size of match C_Γ. This additional step allows verification of candidate matches returned by dynamic shape warping.

Finding the best matching Doppler image to a given image can then be formulated as finding the 2D envelope G(t) such that

$$G_{best} = argmin_G |\overrightarrow{F}(\Gamma(t')) - \overrightarrow{G}(t')| \qquad (6)$$

while choosing the best Γ for each respective candidate match $G(t)$.

Recovering correspondence of fiducial points

We recover the correspondence of fiducial points F_K, G_M using shape-based dynamic time warping. For this, we form a dynamic programming matrix H where the element $H(i, j)$ is the cost of matching up to the ith and j^{th} element in the respective multi-d curves as

$$H(i,j) = \begin{cases} H_{i-1,j-1} + d(\overrightarrow{F}(t_i), \overrightarrow{G}(t'_j)) \\ H_{i-1,j} + d(\overrightarrow{F}(t_i), 0) \\ H_{i,j-1} + d(0, \overrightarrow{G}(t'_j)) \end{cases} \qquad (7)$$

with initialization as $H_{0,0} = 0$ and $H_{0,j} = \infty$ and $H_{i,0} = \infty$ for all $0 < i \leq K$, and $0 < j \leq M$. The shape constraints for matching are incorporated in the term $d(.)$. Also, the first term represents the cost of matching the feature point $\overrightarrow{F}(t_i)$ to feature point $\overrightarrow{G}(t'_j)$ which is low if the features are similar. The second term represents the choice where no match is assigned to feature $\overrightarrow{F}(t_i)$. The cost function $d(\overrightarrow{f}(t_i), \overrightarrow{g}(t'_j))$ is then given as the Euclidean distance between the two fiducial points using the 4 parameters as

$$d(\overrightarrow{F}(t_i), \overrightarrow{G}(t'_j)) =$$

$$\begin{cases} \sqrt{(t_i - t'_j)^2 + |\overrightarrow{F}(t_i) - \overrightarrow{G}(t'_j)| +} & \text{if } |t_i - t'_j| \leq \lambda_1 \text{ and} \\ \overline{(\theta(t_i) - \theta(t'_j))^2 + (\varphi(t_i) - \varphi(t'_j))^2} & |(\overrightarrow{F}(t_i) - \overrightarrow{G}(t'_j))| \leq \lambda_2 \\ & |\theta(t_i) - \theta(t'_j)| \leq \lambda_3 \\ & |\varphi(t_i) - \varphi(t'_j)| \leq \lambda_4 \\ \infty & \text{otherwise} \end{cases}$$

The thresholds $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ are determined through a prior learning phase in which the expected variations per disease class is noted.

The overall shape-based similarity search algorithm works as follows. During the database creation stage, all cardiac echo studies are processed to separate Doppler
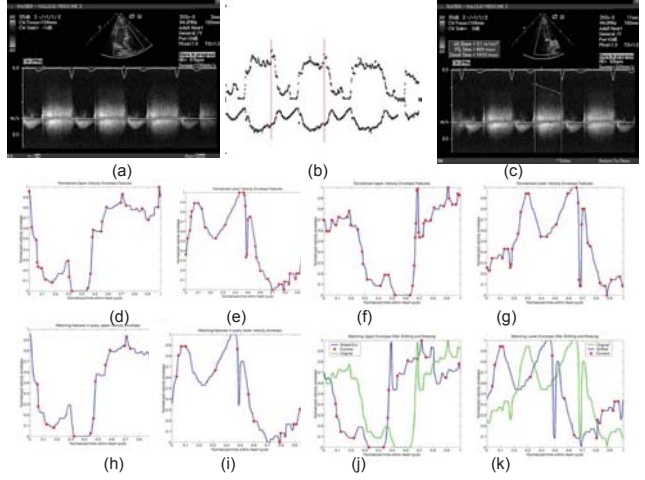

(a) (b) (c)
(d) (e) (f) (g)
(h) (i) (j) (k)

Figure 5. Please see text for details.

frames. The frames are pre-processed as described in Section 3 to extract envelope curves and their corner shape features. Given any new query Doppler image, a ranked set of matching Doppler images are obtained using the matching metric described in Section 4 above, and those that exceed a threshold are retained. From the set of Doppler images retrieved, a histogram of their associated disease tuples (disease type, valve type, severity type) is separately constructed. The peaks in the histogram correspond to label values that have the most support from the matching images, thus increasing their probability of being the correct label values for the query image. Similar statistical distributions can be found for other information associated with the patients, such as medications and outcomes for further enhanced decision support.

5. Results

We now report on experiments that use the above approach to do shape-based similarity retrieval of Doppler images.

5.1. Experimental setup

A set of 2300 cardiac echo videos were collected from 1940 cardiac patients from a large hospital network in our area. Each echo study on the average had 20 CW Doppler frames giving rise to a collection of 34000 Doppler images. Of these about 200 CW Doppler frames were used to train the dynamic shape warping parameters of the matching metric. Each echo study was also associated with a clinical report that documents the findings. Specific diagnosis terms corresponding to ICD9[1] codes representing over 300 cardiac diseases were automatically isolated from the reports using text mining techniques (keyword spotting in summary area of the textual report), to serve as ground truth labels for the patient. Due to the large number of Doppler
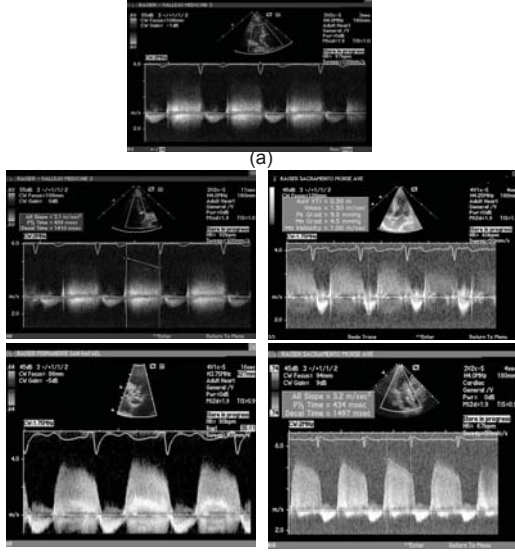
Figure 6. Illustration of shape-based disease similarity retrieval using Doppler images.(a) query Doppler image depicting moderate Aortic regurgitation. Doppler images retrieved in order from left to right, top to bottom. All images retrieved depict Aortic regurgitation, although they vary from mild to moderate Aortic regurgitation.

images, obtaining the correspondence between the disease label of the patient and the specific Doppler image depicting the condition (eg. moderate mitral stenosis) was challenging. To expedite the ground truth labeling,we first assembled all Doppler images from patients with a report label of a particular disease class (eg. moderate mitral stenosis)into one folder. By browsing through this collection, the images were manually divided into similar shape groups. A few representatives from each shape group were then examined by trained personnel (4 experts) who could identify the valve as well as the diseased state of the valve based on the measurements extracted from the valve. The images were then labeled with the disease (e.g regurgitation vs stenosis), the valve (mitral, aortic), and the severity (moderate, severe, etc.). In our experiments, we considered three disease variants, namely, normal, stenosis, regurgitation and three valve combinations, namely, Mitral, Aortic, and Tricuspid valves, and finally, 3 severity variants of mild, moderate, and severe. This gave rise to 2x3x3+1x3(normal)=21 possible combinations of valvular diseases.

Similarity retrieval results

We first illustrate the shape matching between two Doppler envelopes. Figure 5a shows a query Doppler image. Figure 5b shows its velocity envelope. After isolating single heart beat and amplitude normalization given by Equation 4, the upper and lower envelopes appear as shown in Figure 5d-e. The corner shape features used for matching are also indicated by red circles in these figures. A can-

didate matching image is shown in Figure 5c with its corresponding single heart cycle envelopes in Figure 5f-g. As can be seen, there are both missing and spurious features, as well as a global time shift. Using correlation to recover the global shift, and using dynamic shape warping for the non-rigid shape correspondence, we find the matching subset of features as indicated in Figure 5h-k for query and matching image respectively. As can be seen, important fiducial point similarity has been preserved through the non-rigid warping.

Figure 6 show results of similarity retrieval for a query Doppler image depicting Aortic regurgitation. All the matches retrieved show aortic regurgitation although the severity levels vary between moderate to mild (the third match) and the patterns and heart rate are different in these cases.

Disease similarity detection performance

We first evaluated the disease similarity detection using the conventional measures of precision and recall. To evaluate disease similarity detection, we noted the fraction of queries in which the top K predicted disease tuples D_T (includes disease,valve, and severity) derived from top ranking matches included the manually assigned label. We also measured the precision and recall for similarity retrieval as follows.

$$\text{Precision} = \frac{\text{\# matching disease tuples retrieved}}{\text{Total \# of tuples in retrieved images}} \quad (8)$$

$$\text{Recall} = \frac{\text{\# of matching disease tuples retrieved}}{\text{Total \# of tuples for the patient}} \quad (9)$$

The resulting Precision-Recall plots averaging over queries of the same disease class are shown for 6 diseases in Figure 7. As can be seen, the precision and recall are higher for Mitral Stenosis and Aortic Regurgitation due to their more discriminatory patterns. Thus these experiments confirmed the validity of the approach of using the shape patterns conveyed by Doppler images to infer the valvular disease labels for decision support.

It should be noted that the lower precision indicated in the precision-recall curve of Figure 7 for high recall is not necessarily a problem since due to co-morbidity associations (i.e. likely to co-occur in a population), other disease labels could be relevant. To evaluate this, we ran all images as queries against our image database and retained their top 20 matches. We found that 92.4% of the queries had all of their disease labels recovered with a maximum rank of 5. That is, the correct disease labels were within the top 5 matches. Among the non-matches, nearly 50% of them were valid associations as judged by clinicians. With this, the utility of shape similarity retrieval for clinical decision support was demonstrated.

Valve identification:

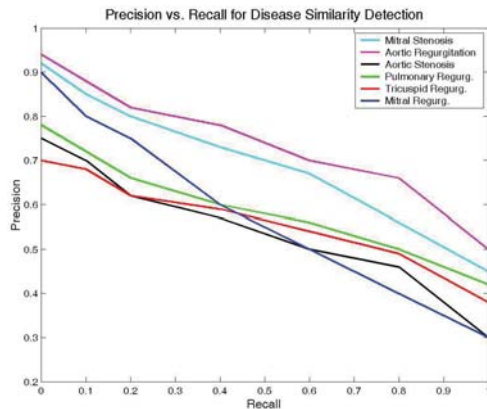| Valve	Stenosis	regurgitation
Mitral	0.92	0.89
Tricuspid	0.5624	0.93
Aortic	0.91	0.94

Figure 7. Table (top) of valve identification accuracy, and illustration (bottom) of precision versus recall for disease similarity retrieval using Doppler images.

Since our matching method explicitly recovers nonrigid deformation between matching shapes, it is expected to perform better than conventional machine learning methods, even for the task of classification. To evaluate this, we compared our approach to two other machine learning methods for the task of valve label classification. For this, we retained only the valve information from the disease labels of matching images, and assigned the most common valve label to the query image. The valve prediction accuracy was recorded as:

$$\text{accuracy}_{valve} = \frac{\text{Number queries w/matching } valve \text{ labels}}{\text{Total number of queries}}$$

(10)

The accuracy of valve prediction for various diseases are shown in the table in Fig. 7. As can be seen, mitral stenosis is easily recognizable due to its characteristics M pattern. The lower recognition of tricuspid stenosis is due to the small number of cases in our collection. In general, the similarity in the appearance of their regurgitation profiles for mitral, and tricuspid valves can also cause confusion.

Using the normalized envelope curves as input to a conventional machine learning classifier gave worse performance. In particular, we used 239 envelope curves drawn from various classes (disease-valve-severity) as training datasets and 466 cases for testing. A neural network algorithm (feedforward with backpropagation) found the average accuracy to be only 10.2%! Using support vector machines only improved the classification to 39.27%. Thus this experiment showed the value of our method for high precision disease-specific classification as well.

6. Conclusion

In this paper, we have explored the utility of shape based matching of Doppler images for disease similarity retrieval. The approach is generalizable to many disease classes and cardiac structures without any training. Experiments show promising results on various cardiac diseases for decision support.

References

[1] ICD-9 Diagnosis Codes. http://icd9data.com/.

[2] ImageCLEF. http://www.imageclef.org/.

[3] R. Barea, L. Boquete, M. Mazo, E. Lopez, O. Garcia, E. Milan, and R. B. Navarro. Automatic Measurement of Blood Flow using Artificial Vision. *IEEE International Conference on Devices, Circuits and Systems*, 1998.

[4] E. G. Caiani, V. Magagnin, C. Champlon, L. Delfino, M. Llambro, and M. Turiel. Quantification of coronary flow velocity reserve by means of semi-automated analysis of coronary flow Doppler Images. *Proc. EMBC*, 2004.

[5] J. G. Dy, C. E. Brodley, A. Kak, L. S. Broderick, and A. M. Aisen. Unsupervised Feature Selection Applied to Content-Based Retrieval of Lung Images. *IEEE Trans. PAMI*, 25(3):373–378, 2003.

[6] H. Feigenbaum, W. F. Armstrong, and T. Ryan. *Echocardiography, Sixth Edition*. Lippincott Williams & Wilkins, 2005.

[7] J. J. JinHyeong Park, S. Kevin Zhou and D. Comaniciu. Automatic Mitral Valve Inflow Measurements from Doppler Echocardiography. *Proc. MICCAI*, pages 983–990, 2008.

[8] Y. Liu and F. Dellaert. Classification driven medical image retrieval. In *Proc. of the IU Workshop*, 1998.

[9] V. Magagnin, E. G. Caiani, L. Delfino, C. Champlon, S. Cerutti, and M. Turiel. Semi-Automated Analysis of Coronary Flow Doppler Images: Validation with Manual Tracings. *Proc. EMBC*, 2006.

[10] H. Mller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of content-based image retrieval systems in medical applications–clinical benefits and future directions. *Intl. J. of Medical Informatics*, 73(1):1 – 23, 2004.

[11] S. Nudelman, A. L. Manson, A. F. Hall, and S. J. Kovacs Jr. Comparison of Diastolic Filling Models and Their Fit to Transmitral Doppler Contours. *Ultrasound in Medicine and Biology*, 21(8):989–999, 1995.

[12] N. Otsu. A threshold selection method from gray-scale histogram. *IEEE Trans. Syst., Man, Cybern.*, 9(1):62–66, 1979.

[13] O. Shechner, M. Sheinovitz, M. Feinberg, and H. Greenspan. Image Analysis of Doppler Echocardiography for patients with Atrial Fibrillation. *ISBI*, 2004.

[14] H. D. Tagare, C. C. Jaffe, and J. Duncan. Medical Image Databases: A Content-based Retrieval Approach. *J. of the American Medical Informatics Assoc.*, 4:184–198, 1997.

[15] J. Tschirren, R. M. Lauer, and M. Sonka. Automated Analysis of Doppler Ultrasound Velocity Flow Diagrams. *IEEE Trans. on Medical Imaging*, 20(12):1422–1425, 2001.