Finding Disease Similarity by Combining ECG with Heart Auscultation Sound

F Wang, T Syeda-Mahmood, D Beymer

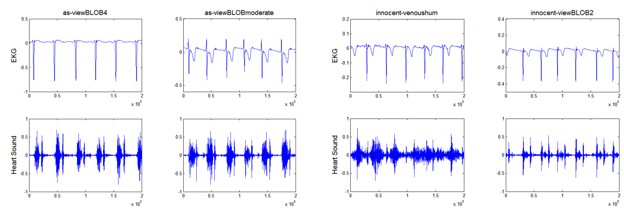
IBM Almaden Research Center, San Jose, CA, USA

Abstract

Heart auscultation and ECG are two very important and commonly used diagnostic aids in cardiovascular disease diagnosis. Physicians routinely perform diagnosis from simple heart auscultation and visual examination of ECG waveform shapes. It is common knowledge to physicians that patients with the same disease have similar-looking ECG shapes and comparable heart sounds. A key idea explored in this paper is to automatically capture such shape similarity in the ECG and audio signals, which are combined to find disease similarity. Specifically, we present a general method of capturing the perceptual shape similarity of the ECG and audio waveforms by modeling the morphological variations in the signals representing the same disease across patients. Differences in shape corresponding to the same disease are modeled as a constrained non-rigid translation. Patients with similar diseases are retrieved by recovering the non-rigid alignment transform using a variant of dynamic time warping. Results are presented that demonstrate the method on audio shape-based discrimination of various cardiovascular diseases.

1. Introduction

Multimedia data is widely used in medical diagnosis through auditory and visual examination by physicians. In particular, heart auscultations and ECGs are two very important and commonly used diagnostic aids in cardiovascular disease diagnosis. While physicians routinely perform diagnosis by simple heart auscultation and visual examination of ECG waveform shapes, these two modalities reveal different diagnostic information about the heart. The heartbeat, for example, can reveal abnormal sounds caused by valvular disease, pericarditis, and dysrhythmia. On the other hand, the ECG unveils abnormal electrical activity during the contraction of the myocardium. Therefore different patients may have similar heart sounds but very different ECGs (Fig. 1a), or similar ECGs but different heart sounds (Fig. 1b). Consequently, disease similarity between different patients can be better characterized by combining the two modalities, which in turn, will provide more accurate diagnosis of a particular heart disease.



(a) (b)

Figure 1. Illustration of (a) similar heart sounds with different ECGs; (b) similar ECGs with different heart sounds.

It is common knowledge to physicians that patients with the same disease have similar-looking ECG shapes and comparable heart sounds. A key idea explored in this paper is to automatically capture such shape similarity in the ECG and audio signals (the shape of the audio signal is represented through perceptually salient envelope curves), which are combined to find disease similarities. Specifically, we present a general method of capturing the perceptual shape similarity of ECG and audio waveforms by modeling the morphological variations in the signals representing the same disease across patients. Different morphological variations of the shape corresponding to the same disease are modeled as a constrained non-rigid translation. Retrieval of patients with similar disease is achieved by recovering the non-rigid alignment transform using a variant of shape-based dynamic time warping.

This paper makes several novel contributions. To our knowledge, this is the first attempt to combine ECG with heart auscultation sounds to find disease similarity. Secondly, it advances the state of the art in cardiology by introducing the notion of automated similarity search for physician decision support. Unlike previous work in computing ECG/audio similarity that focuses on spatio-temporal feature extraction and classification, our method of computing signal similarity is based on the recognition of shape patterns produced by ECG/audio signals.

2. Previous work

While we are not aware of disease similarity retrieval based on combining ECG and heart auscultation sound, there is considerable work on both ECG and audio analysis for retrieval and computer-aided diagnosis of heart diseases, which we now review.

Single ECG analysis and ECG classification are well-researched fields, and the popular techniques include neural network [1], machine learning methods, wavelet transforms [2] and genetic algorithms. The rule-based methods rely on the accuracy of the P-Q-R-S-T segment detection [3]. Errors in estimation of these feature values can cause major errors in disease-specific interpretation. The parametric modeling methods, on the other hand, are good at spotting major disease differences but can't take into account fine morphological variability due to heart rate (eg. ventricular vs. supra-ventricular tachycardia) and physiological differences. Related work in the time alignment of ECGs also exists. Dynamic time warping has been a popular technique in ECG frame classification [3], and more recently, in the recognition of heart beat patterns for synthetically generated signals [4]. In all such alignments, however, the amplitude of the signal was used rather than a detailed modeling of the shape. To overcome this problem, we presented a method of capturing the perceptual shape similarity of ECG waveforms in [5] by combining shape matching with dynamic time warping

Automatic analysis of heart sounds has been investigated for detecting heart abnormalities. The predominant approach in heart sound analysis is based on feature extraction and classification, as is conventional for audio analysis. These features can be roughly classified into two categories, the spatio-temporal features such as the zero-crossing rate (ZCR), hidden Markov features etc., or frequency-domain features such as Mel-frequency Cepstral Coefficients (MFCC). In [6], a neural network was trained on heart sounds to classify several valve-related disorders. However, the identification of disorders was based on detection of s1 and s2 sounds and noting their separation. For many diseases, such as pericardial rub, the periods s1 and s2 are hardly distinct and hence are not reliable for audio similarity detection.

3. Methods

We now present the details of our proposed joint ECG and heart sound analysis algorithm. Our approach is based on a key observation that patients with the same disease have similar-looking ECG shapes and visual appearance of the audio signals, which is captured through perceptually salient envelope curves. Therefore different morphological variations of the shape corresponding to the same disease can be modeled as a constrained non-rigid translation transform. Matching of both ECGs and heart sounds then involve recovering the non-rigid alignment transform using a variant of shape-based dynamic time warping. Matching scores from the ECG and heart sound signals are then combined to infer disease similarity between patients. Each step of our algorithm is illustrated in Fig. 2. We begin with the preprocessing module, which generates features needed for signal matching.

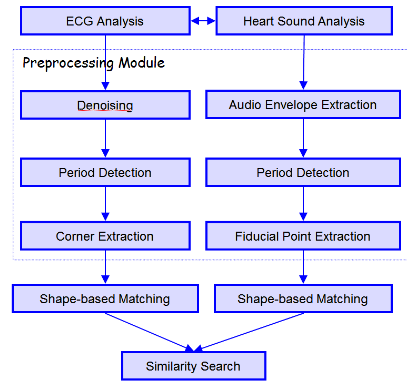


Figure 2. Illustration of the processing steps of our ECG and heart sound based disease similarity search algorithm.

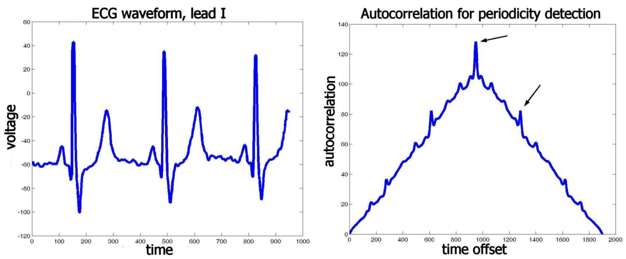
3.1. Preprocessing Module

ECG and heart sounds are aqcuired differently, so compared with the ECG signal, heart sounds are highly non-stationary, i.e. their frequency properties vary with time. Therefore the two modalities require different treatment during the preprocessing step.

3.1.1. ECG preprocessing

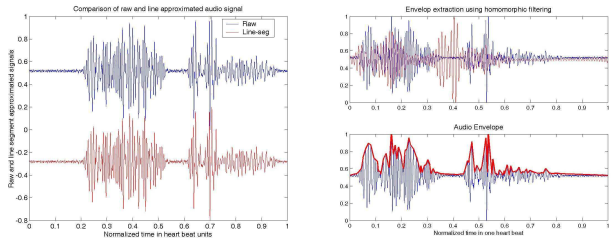
The main steps of the ECG preprocessing were introduced in [5]. For convenience, we reproduce the key steps here. To extract a single heart beat duration, we normalize a signal $f(t)$ in amplitude to $[0,1]$, and compute the autocorrelation function. As shown in Fig. 3b, the peaks in the autocorrelation function correspond to the various periodicity patterns found in the signal. We note the most common inter-peak duration as representative of a heart beat duration and extract a segment of recovered duration from the ECG signal. This segment becomes the basis of our shape-based alignment scheme. This normalization of the time axis is performed to ensure that all signals being compared are one heart beat long and have their time values range from 0 to 1.0.

The fiducial points extracted from time series are corners. A simple line segment approximation that does a recursive partitioning of the time series curve is used. The shape information at each corner is modeled using $(t_i, \vec{f}(t_i), \theta(t_i), \phi(t_i))$ where $\theta(t_i)$ is the included angle in the corner at t_i, and $\phi(t_i)$ is the orientation of the bisector. Using the angle of the corner ensures that wider QRS complexes are not matched to narrow QRS complex. The angular bisector ensures that polarity reversals such as inverted T waves or change in ST elevation can be captured.



(a) (b)

Figure 3. Illustration of pre-processing steps in shape matching algorithm. (a) ECG waveform for lead I. (b) Autocorrelation-based period detection (the distance between the two arrows).



(a) (b)

Figure 4. Illustration of the processing steps of audio signal. (a) line segment approximation. (b) envelop extraction using homomorphic filtering and described method.

3.1.2. Heart sound signal preprocessing

Our approach to fiducial point extraction approximates the audio signal by its envelope curve and chooses curvature extrema points as fiducials. Various algorithms are available for envelope extraction from signals, including homomorphic filtering. While these algorithms are less sensitive to noise-related fluctuations, they frequently extract the low frequency component of the signal rather than a faithful approximation of the perceptual envelope.

To recover the audio envelope, we adopt a new approach where we first form a line segment approximation to the audio signal, which is similar to the line segment approximation step in the ECG preprocessing. Each consecutive pair of line segments then defines a corner feature C_i. We define the audio envelope (AE) as the set of points which is a local maximum, with its signal values above the baseline. Similarly, all peaks below the baseline can be recorded as a minima envelope curve. Due to the inherent symmetry of the heart sounds, we can restrict the matching to maxima leaving the minima for detailed verification later.

Figure 4b, lower row, illustrates the audio envelope extracted from the audio signal depicted in Figure 4a for Atrial Septal Defect. In comparison, the result of homomorphic filtering-based envelope extraction is shown in the upper row of Figure 4b. As can be seen, the boundary shape of the signal is well-captured in the audio envelope in comparison to the filtered envelope. Since this method of envelope extraction is sensitive to noise, we perform a noise filtering of the audio signals before signal matching. We use a wavelet transform to filter the noise, which, unlike linear low pass filters, allows features in the original signal to remain sharp.

Accurate detection of the period of the audio signal is a difficult problem because of the non-regularity of the heart sound. It is further complicated by recording noise in digital stethoscopes. However, since the ECG and the heart auscultation sound should have comparable period cycles, we can use the extracted period from ECG to screen a bunch of candidate periods, which are computed as the peaks of autocorrelation function of the audio envelopes.

3.2. Shape-based matching

Given a query ECG or audio signal, the matching audio signals are found by a two-step process in which they are first brought into rough alignment and then matched using the dynamic time warping algorithm introduced in [5]. The rough alignment is necessary since the periodicity detection algorithm isolates one heart beat interval from the original signal, and there may be an initial translation bias depending on the starting point. As a result, the signals may need to be circularly shifted to perform an initial registration. As the translation required here is usually much larger than allowed during DTW alignment, it must be extracted separately.

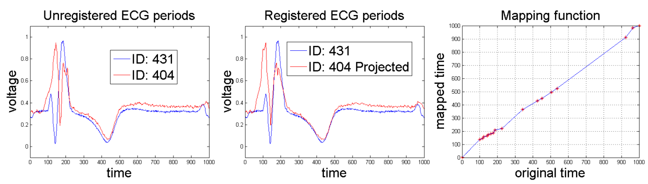
Once the pair of signals is initially registered, DTW alignment is performed to match the fiducial features as described in [5]. Final verification and ranking is performed on the retrieved matches by projecting the query signal onto the retrieved signal and ranking based on the resulting mean square error between the projected query and the retrieved signal.

This pair-wise matching of single heart beat intervals is repeated over multiple such heart beat segments over the available data and the average residual error is used to rank the matches. Finally, ECG and audio signals are combined by adding the matching scores from the DTW alignment to infer disease similarity between patients. Future work will explore other modality fusion methods.

4. Results

We now present the evaluation of our shape matching algorithm. Due to privacy concerns, there are only a limited number of data sets available for public experimentation. We used ECGs and sounds from a number of medical schools, including Boston Children Hospital, Univ. of Washington Medical School, UCLA Medical School, UCSD, and reference CDs provided by digital stethoscope makers (Littman). Thus, the ground truth labels for diseases were known during the evaluation. Currently, the collection has over 170 examples for various kinds of diseases, Mitral regurgitation, Mitral Stenosis, septal defects, Cardiomyopathy, etc. Each disease was represented by 3-10 patient samples.

We illustrate shape-based disease similarity detection using an example. Fig. 5a shows single periods of ECG waveforms extracted as described in Section 2. The waveforms are similar except for a non-rigid translation transform. This is computed using the DTW algorithm to give an alignment as shown in Fig. 5c. The resulting alignment of the waveforms is shown in Fig. 5b. The improvement in shape matching due to non-rigid DTW alignment can be clearly seen by comparing the simple overlaid shapes in Fig. 5b. As can be seen from Fig. 5c, the alignment is close to the diagonal illustrating a good match.



(a) (b) (c)

Figure 5. Illustration of ECG shape matching. (a) Two ECG waveforms that are shape-wise similar except for a non-rigid translation transform. (b) Result of alignment. (c) DTW-alignment function.

The alignment of the two audio signals corresponding to the ECGs are illustrated in Fig. 6. Fig.6a and 6b show the audio envelopes of the query and the matching database audio signals. Figure 6c shows the result of initial alignment by recovering the shift as explained in Section 3.1.2. Fig. 6d and e show the matching fiducial points found using DTW alignment. The alignment transform itself is in-

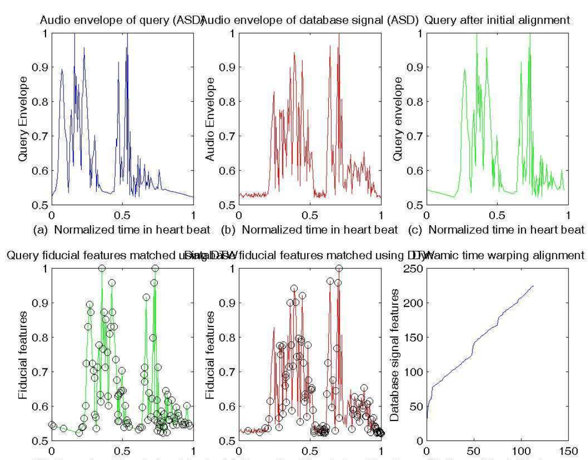


Figure 6. Illustration of audio shape matching.

We evaluated the precision and recall of our method by using all available samples per disease as queries and retrieving matches from the database for various choice of thresholds. The results are compared with the precision-recall curve using each single modality. The precision and recall values were averaged over the queries tested for the respective classes. Figure 7 shows the performance of the two methods on the entire ECG and audio database. Both precision and recall are higher using the combination of ECG and audio as compared with individual modalities.

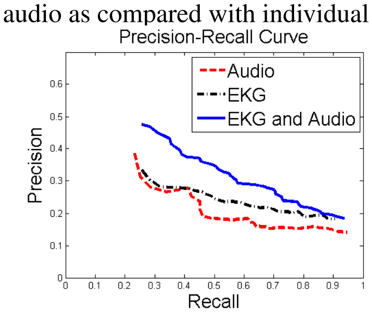


Figure 7. Illustration of precision and recall

5. Discussion and conclusions

In this paper, we have presented a novel algorithm for shape-based retrieval of audio and applied it to the problem of retrieval of similar heart records. Starting from the key idea that similarity in ECG and audio shape implies similarity in disease type, we demonstrated an algorithm for disease similarity search based on shape similarity of ECG and heart auscultation sounds.

References

[1] Izeboudjen N, Farah A. A new neural network system for arrhythmia's classification. In Neural Computation. 1998; 208–212.

[2] de Chazal P, Celler BG, Reilly RB. Using wavelet coefficients for the classification of the electrocardiogram. IEEE Conf on EMBS 2000;1:64–67.

[3] Alexakis C, Nyongesa H, Saatchi R, Harris N, Davies C, Emery C, Ireland R, Heller S. Feature extraction and classification of electrocardiogram (ecg) signals related to hypoglycaemia. Computers in Cardiology 2003;30:537–540.

[4] Tuzcu V, Nas S. Dynamic time warping as a novel tool in pattern recognition of ecg changes in heart rhythm disturbances. IEEE Conf on SMC 2005;1:182–186.

[5] Syeda-Mahmood T, Beymer D, Wang F. Shape-based matching of ecg recordings. IEEE Conf on EMBC 2007;.

[6] Alim O, Hamdy N, El-Hanjouri M. Heart diseases diagnosis using heart sounds. Proc National Radio Science Conference 2002;636 – 640.

Address for correspondence:

Fei Wang
650 Harry Road, San Jose, CA 95120
wangfe@us.ibm.com