ISSN:0975-9646

Amrita Maity et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012, 4404 - 4407

Performance Evolution of Heart Sound Information Retrieval System in Multicore Environment

Amrita Maity [#1], P. Sivakumar [#2], M Rajasekhara Babu [#3], Ch. Pradeep Reddy*
[#]*School of Computing Science and Engineering*
**School of Information Technology and Engineering*
VIT University, Vellore-14, T.N., India

***Abstract*— The most difficult and problematic task for cardiology specialists is the interpretetation of heart sounds. It requires special skill and clinical experience along with detailed and expensive tests. However, while detection of heart disease heart sound is proffered for the first step of diagnosis . Recently, many Computer aided auscultation has emerged as a cost-effective technique to analyze and interpret the heart sounds. In this study we propose a feasible technique for developing a heartbeat sound information retrieval system using sound wave matching useful towards automated heart disease detection. The feasibility of these text based retrieval approaches are shown from retrieval experiments with around 23 digital heart sound recordings which are repeatedly matched with the database heart sound and the corresponding disease is given. The sound information retrieval is carried out parallel for better performance.**

***Keywords*— Heart sounds, auscultation, MIDI, pitch tracking**

I INTRODUCTION

The heart disease detection becomes complicated as the heart sound is mixed with recording environment condition, auscultation body points etc.Moreover background noise mixed with digital heart sound recordings adds to it. The heart sounds are the sounds generated by blood turbulence and beating heart. These sounds facilitate the heart disease diagnosis with easy accessibility and reasonably low cost. Thus auscultation is an important diagnostic method for cardiovascular analysis.

Computer aided auscultation has come up as a cost effective technique to analyze and interpret the heart sounds. It equips physicians with an unbiased tool to make accurate diagnostic and treatment decisions [1, 2].As the heart sounds mixed by background noise, detection of heart diseases becomes complicated. Not only that, complication of such detections increases by a number of other factors including similarity among heart diseases, recording environment conditions, auscultation body points. There are several methods for automated detection and classification of heart diseases and heart sound analysis that have been proposed till yet. Some of them used Artificial Neural Network method for detection and classification of heart sound [3-5].Another advanced technique that it used for diagnosis the heart problem is Hidden Markov Model[6,7] that they suggest HMM for segmentation of heart sound recorded for clinical and classification purpose.

In this study an approach towards developing a heartbeat sound information retrieval system useful towards automated heart disease detection has been proposed. The audio format of heart sound recordings are preprocessed and converted into the MIDI format. Next the MIDI files are encoded to text strings using N-grams[8]. These text strings are then indexed and tested for retrieval using both database and Information Retrieval(IR) systems. The Longest common subsequence (LCS) string matching algorithm was used for identifying similarities from the database. With IR, full text indexing of the recordings was used and retrieved using known item searches from a search engine.

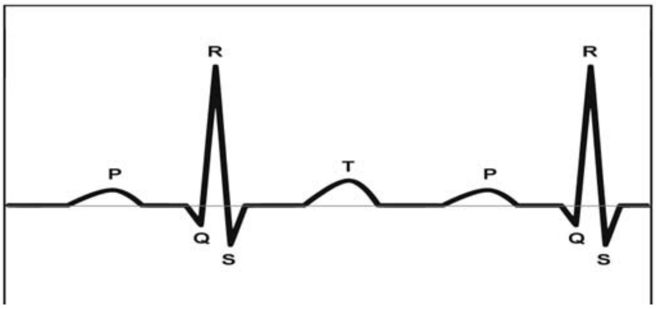

Figure 1: Standard heartbeat cardiograph diagram

II DESIGN

There are various retrieval experiments already been proposed. But in the existing system all the sound information retrieval has been done with the help of serial approach. In this paper the heart sound information retrieval has been done applying the parallel approach. At the end the performance is calculated comparing both the approaches.

A. Existing system

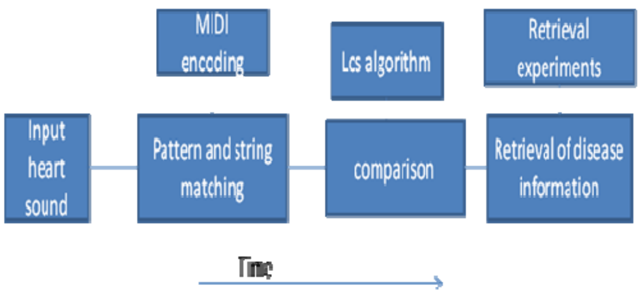

Figure 2: serial approach of heart sound information retrieval

B. Proposed system

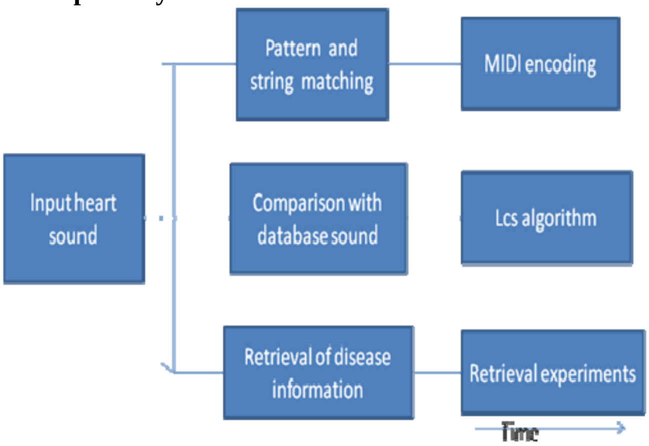

Figure 3: parallel approach of heart sound information retrieval

In serial approach the retrieval experiments are carried out one after another. Hence the efficiency is less. But in parallel approach retrieval experiments are carried out simultaneously as a result performance is better with respect to the serial approach.

III. IMPLEMENTATION

Retrieval approaches---
Two different approaches has been described in this section to describe the feasibility of heart sound information retrieval system.

A. Information Retrieval

The method of finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from large collections that are stored on computers called IR [9] where documents relevant to a user query can be retrieved quickly. The N-gram approach has been used to obtain musical words from MIDI encoding [8]. In adopting this technique, the heartbeat
Sounds would be transcribed to musical notation. The summary of steps taken in obtaining monophonic musical sequences from polyphonic sequences is as follows:

1. The piece has been divided into overlapping windows of n different adjacent Onset times using a gliding window approach
2. All possible combinations of melodic strings are obtained from each window.

The details of encoding these n-grams into musical words are available in Doraisamy et.al [8].

B. String matching

To test the retrieval of the database with another text-based approach, string matching was used. To measure the similarity of entered query and the existing records of heart sounds within the database, a Text-based system has been introduced. In a full text search algorithm, the Longest Common Subsequence (LCS), has been used that examines all words in every stored text file is used to match search words supplied by the heartbeat sound. In this approach LCS examine the similarity of query with existing heart records in the data base.

IV DATA COLLECTION

This section describes the preprocessing of the audio sound files and MIDI encoding.

A. Data preprocessing

Background noises along with many undesirable sounds mixed with the heartbeat sound creates lot of problems in the process of recording heart sounds. Such sounds are generally produced by the stethoscope di r wit tie t' Iu sounds and many other breathing are other internal sources of noise. To solve this problem the collection of heart sounds are preprocessed prior to encoding.
Audio format is converted to wave format, amplifying and normalizing Heart sound records and then applying noise reduction effect on the data sets. Figure 3 shows the normal record sound before and after the preprocessing.

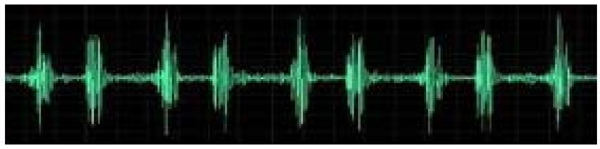

Figure 4: heart sound before preprocessing

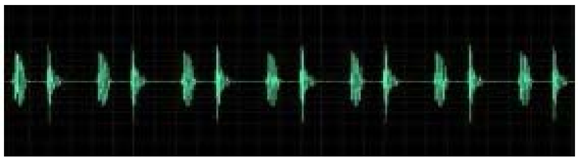

Figure 5: after preprocessing

The number of dataset is 23 and they are classified in eight categories as shown in table 1.In addition to it 6 records of normal heart sound used in the experiment were captured by means of a electronic stetoscope.Other heart sounds already captured in the form of comprehensive data set as benchmarks to evaluate the proposed system. All the rerecords that collected included ten cycles of heart sounds and approximately 10 sec long.

Name of category	Number of records
Normal sounds	5
Split second sound	2
Ejection sound	7
Clicks	3
Gallop Rhythm	2
Systolic Murmurs	4
Diastolic Murmurs	2
Continues Murmurs	8

Table 1: Properties of subject heart sounds

B. MIDI encoding

MIDI is known as a popular format in computer –based music production such as composition, transcription etc.With wave to MIDI converters; it is very much possible to generate MIDI data of PCM based sounds rather easily. It extracts acoustical characteristics such as pitch, volume, and duration for converting them into a sequence of notes for producing music scores. The quality of MIDI transcription would also depend on the algorithm capabilities which are utilized.

Since heart sound frequency is between 25-50Hz therefore encoding component has filtered heart sound with this frequency and ignored another sounds. In the first step we need to select a tool for MIDI encoder for changing audio files to MIDI format. For this purpose 4 MIDI converters were compared that are shown in Figure 6

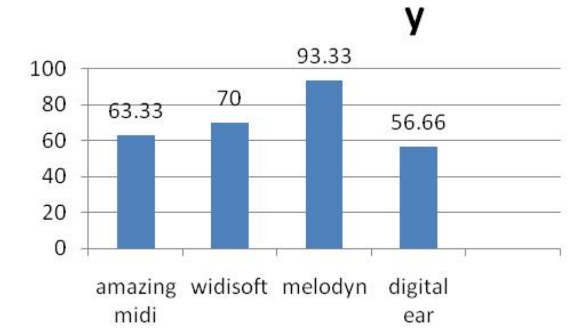

Figure 6: MIDI encoding correct ratio percentage

The MIDI correct ratio percentages are calculated by these formulae.

$$\text{correct}(\%) = \frac{\text{The number of correctly encoded pitches}}{\text{Total number of played pithes}} \quad (1)$$

V.Result Analysis

This section presents the evaluation process for the proposed system in this paper. A comprehensive experimental analysis using simulation is carried in order to show the usefulness of the proposed system. Finally the results are discussed.

A. Retrieval accuracy

Several tests have been performed to evaluate the number of cycles necessary to obtain highest retrieval accuracy. We started our test with a single cycle query and calculate the accuracy with the below metric.

$$\text{Retrieval Acuracy} = \frac{\text{total Number of cuurent diagnosis}}{\text{Total number of query}} \quad (2)$$

Then we repeated the test until each query contains 14 cycles. As shown in the fig the best performance is achieved when the heart sound is retrieved parallel. That means when the matching is done with more than one sound query in the database it yields the higher performance.

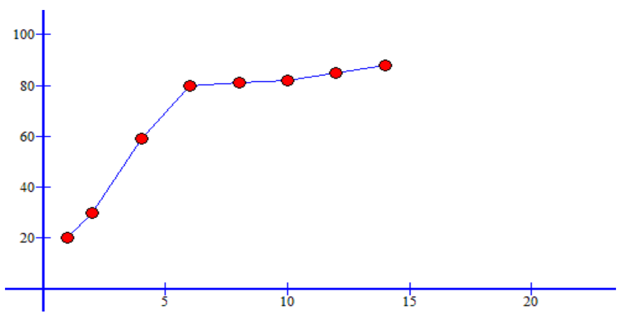

Figure 7: Effect of number of cycles on system performance

B. Retrieval Experiments

The experiment operation process is described as follows. First the heartbeat sounds are used the feed the proposed system. In this experiment we collectively used 33 heartbeat sounds(23 kept in database and 10 heart sounds using for queries. In the process of saving heart sound records in the database in order to easily distinguish relevant results we assign relevant category by using relation between category table and Auscultation Area table of the database to show the category of the ranked disease to the name of the disease. For instance, Systolic murmur is the relevant category for Aortic Stenos is problem.

According to this classification the sample queries were made to database randomly to measure the similarity between the made queries and the existing records in the database. The number of queries made in the experiment is 10.However the performance metrics were calculated for the whole retrieved heart sounds and their returned results were taken into account.

In order to search the given query in the database, developed LCS algorithm is used for searching and retrieving database. This algorithm has been based on the text matching mechanism.

In order to compare the similarity of two strings LCS algorithm has been used. A similarity more than 70% is relevant to our method. If the result categories are similar with query categories they will be considered as correct retrieved heart diseases.

VI Conclusion

The feasibility of diagnosis by matching text strings obtained from heart sound transcriptions has been shown. As heart disease is a fatal disease so it needs to be detected as soon as possible. So this parallel approach of heart sound information retrieval can be very efficient. The above method will be very much useful in rural areas where still now the treatment is not that advanced. Thus heart sound information retrieval system using parallel approach has been a beneficial technique. However, more experiments and tests would be required in order for this to be used as large-scale heartbeat sound information retrieval system.

Category	LCS
Normal sound	89%
Split second sound	17%
Ejection sound	25%
Clicks	19%
Gallop Rhythm	38%
Systolic Murmurs	36%
Diastolic Murmurs	78%
Continues Murmurs	21%

Table 2: precision values.

Reference

[1] Total T, Polar H, Gluer I. Software Development For the Analysis of Heartbeat Sounds with LabVIEWinDiagnosis of Cardiovascular Disease. Journal of Medical System; 2008; 32:409-21

[2] Delgado-Trejos E, Quiceno-Manrique AF, Godino-Llorente JI, Blanco-Velasco M, Castellans-Dominguez G. Digital auscultation analysis for heart murmur detection. Ann Biomed Eng; 2009; 37(2):337-53.

[3] Azra'ai R.A, bin Taib M.N, Tahir N.M. Signal Processing and Communication Systems, ICSPCS2008; 2:15-17

[4] Babaei S, Geranmayeh A, Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals, Computers in Biology and Medicine 2009;39: 8-15
[5] Bhatikar S.R, DeGroff C, Mahajan R. L. A Classifier Based on Artificial Neural Network Approach for Cardiac Auscultation in Pediatrics. Artif. Intel. Med. 2005; 39: 251□260
[6] Chung Y, Classification of Continuous Heart Sound Signals Using the Ergodic Hidden Markov Model, J. Martí et al. (Eds.),_ 3rd Iberian conference on Pattern Recognition and Image Analysis._Girona,Spain:Springer;2007:5670
[7] Chung Y, A Classification Approach for the Heart Sound Signals Using Hidden Markov Models, 2006.4109/2006, 375-83
[8] Doraisamy S, Rüger R, Robust Polyphonic Music Retrieval with N-grams. Journal of Intelligent Information Systems 2003; 21:53-70
[9] Christopher D, Prabhakar R & Hinrich S.Introduction to Information Retrieval. Cambridge University Press, 2009; 154-6
[10] Itou N, Nishimoto K, A. Voice-to-MIDI System for Singing Melodies with Lyrics, ACE 2007. Salzburg, Austria: ACM, 2007; 183-89
[11] J.-M. Cauvin, C. le Guillou, B. Solaiman, M. Robaszkiewicz, P. le Beux, and C. Roux, "Computer-assisted diagnosis system in digestive endoscopy," *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 4, pp. 256–262, Dec. 2003.
[12] S. Daniel, Hirschberg. Algorithms for the Longest Common Subsequence Problem. Journal of ACM (JACM) 1977 ; 24:664-75