Eng. & Tech. Journal, Vol.30 , No.14, 2012

A Proposed System for Sound Retrieval Using MAS and ANN

Dr. Abeer Tariq
Computer Science, University of Technology/Baghdad
Email: Abeer282003@yahoo.com
Ikhlas khalaf
Computer Science, University of Technology/Baghdad
Email: ekhlasiq@yahoo.com
Shatha habeeb
Computer Science, University of Technology/Baghdad
Email: shathahabeeb@yahoo.com

Received on: 31/5/2011 & Accepted on: 3/5/2012

ABSTRACT

As the use of sounds for computer interfaces, electronic equipment and multimedia contents, has increased, the role of sound design tools has become more important. In sound retrieval, picking one sound out from huge data is troublesome for users because of the difficulty of simultaneously listening to plural sounds and sometimes there are difficulties with speech and sound recognition. Consequently, an efficient retrieval method is required for sound databases.

This research proposes a system aim to deal with sound retrieval in both two cases: authenticity and normal. In the first case, authenticity, two algorithms has been develop one for building the authentication database and the second deal with user sound sample to retrieve the matched authenticated samples. In the second case normal we develop algorithm to deal with user sound sample to retrieve all the matched samples. Many techniques used in this proposed system such as Artificial Neural Network (ANN), Data Encryption Standard (DES), Multi Agent System (MAS) and Fourier transformation (FT). Using these combinations of advanced and adaptive techniques supports the system to be reliable, secure and parallel.

Keywords: sound retrieval, authentication, MAS, ANN, FT, DES.

نظام مقترح لاسترجاع الصوت بأستخدام نظام متعدد الوكلاء والشبكات العصبيه الاصطناعية

الخلاصة

مع زيادة اهمية الاصوات في حقل الحاسوب والمعدات الالكترونيه والمحتويات المتعددة الوسائط ،اصبح دور ادوات تصميم الصوت اكثر اهميه. عند استرجاع الصوت فان عملية الاختيار من بين كميه هائله من البيانات تكون عمليه شاقه للمستخدمين بسبب صعوبة الاستماع

لعدة اصوات معا وفي بعض الاحيان توجد هناك صعوبات في تمييز مختلف الاصوات. بناءا على ذلك نحتاج الى وسيله فعاله لاسترجاع قواعد البيانات السليمه.

هذا البحث يقترح نظام يهدف لتعامل مع نظام استرجاع الصوت في كلا الحالتين:الامنية (التخويل) وبدون امنية (الطبيعيه) في الطريقه الاولى البحث طور خوارزميتان الاولى تبني قاعدة البيانات والثانيه للتعامل مع عينة صوت المستخدم لاسترجاع عينة الصوت الاصليه المطابقه اما الحاله الثانيه الطبيعيه فالبحث طور خوارزميه للتعامل مع عينة المستخدم واسترجاع كل العينات المطابقه. وقد تم استخدام العديد من التطبيقات في النظام المقترح مثل الشبكات العصبية وطريقة التشفير المسماة تشفير البيانات القياسية ونظام متعدد الوكلاء والتحويل بطريقة فورير. وباستخدام هذه التقنيات المتكيفة والمتقدمة تم دعم النظام ليكون موثوق ومعول عليه ويتسم بالامنية ويعمل بشكل متوازي.

INTRODUCTION

Audio data is an integral part of many modern computer and multimedia applications. Numerous audio recordings are dealt with in audio and multimedia applications. The effectiveness of their deployment is greatly dependent on the ability to classify and retrieve the audio files in terms of their sound properties or content. Rapid increase in the amount of audio data demands for a computerized method which allows efficient and automated content-based classification and retrieval of audio database . For these reasons, commercial products of audio retrieval are emerging.

Sound retrieval, Sounds are time-varying signals in the real world and, indeed, all of their meaning is related to such time variability. Therefore, it is interesting to develop sound analysis techniques that allow to grasp at least some of the distinguished features of time-varying sounds, in order to ease the tasks of understanding, comparison, modification, and resynthesis. One of the most useful visual representations of audio signals is the sonogram, also called spectrogram that is a color- or grey-scale rendition of the magnitude of the STFT on a 2D plane where time and frequency are the orthogonal axes [1, 2, 3, 4, 5, 6,].

The basic operation of the retrieval system is as follows. First, a suitable corpus of audio examples must be accumulated and parameterized into feature vectors. The corpus must contain examples of the kinds (classes) of audio to be discriminated between, e.g. speech and music, or male and female talkers. Next, a tree-based quantizer is constructed. This is a \supervised" operation and requires requires the training data to be labeled, i.e. each training example must be associated with a class. This is all the human input required. The tree automatically partitions the feature space into regions ("cells") which have maximally different class populations.

To generate an audio template for subsequent retrieval, parameterized data is quantized using the tree. An audio file can be characterized by finding into which cells the input data vectors are most likely to fall. A template is just an estimate of the vector counts for each cell; in other words a histogram. This template captures

the salient characteristics of the input audio, because sounds from different classes will have very different counts in the various histogram bins, while similar audio _les should have similar counts. To retrieve audio by similarity, a template is constructed for the query audio. Comparing the query template with corpus templates will yield a similarity measure for each audio _le in the corpus. These can be sorted by similarity and the results presented as a ranked list as in conventional text retrieval, then distance measures that may be used to estimate the similarity between templates, and thus between the corresponding audio files. The last step presents experimental retrieval results on a corpus of sounds and another of music. Quantitative measures of retrieval performance are shown, as well as comparative results from another retrieval method on the same corpus[7,8]. Figure(1) illustrate audio template construction.

BASIC TECHNIQUES USED IN PROPOSAL

There are three basic techniques used in our proposed system we will explain them according to using them in our proposal:

ANN Back Propagation

Artificial Neural network models in artificial intelligence are usually referred to as artificial neural networks (ANNs); these are essentially simple mathematical models defining a function or a distribution over or both and , but sometimes models are also intimately associated with a particular learning algorithm or learning rule [9].

In our proposed system we use this ANN for noise removal, by applying the two phases: propagation and weight update, figure (2) present the proposed architecture for ANN BP [9].

Phase 1: Propagation, Each propagation involves the following steps:

1. Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activations.
2. Back propagation of the propagation's output activations through the neural network using the training pattern's target in order to generate the deltas of all output and hidden neurons.

Phase 2: Weight update, For each weight-synapse:

1. Multiply its output delta and input activation to get the gradient of the weight.
2. Bring the weight in the opposite direction of the gradient by subtracting a ratio of it from the weight.

This ratio influences the speed and quality of learning; it is called the *learning rate*. The sign of the gradient of a weight indicates where the error is increasing, this is why the weight must be updated in the opposite direction. Repeat the phase 1 and 2 until the performance of the network is good enough for noise removal.

DES

Data Encryption Standard *(DES)* is a block cipher that uses shared secret encryption.. It is based on a symmetric-key algorithm that uses a 56-bit key. DES consequently came under intense academic scrutiny which motivated the modern understanding of block ciphers and their cryptanalysis. It takes a 64-bit block of plaintext as input and outputs a 64-bit block of ciphertext. Since it always operates

on blocks of equal size and it uses both permutations and substitutions in the algorithm, DES is both a block cipher and a product cipher [10].

As all Know DES encryption method very suitable for closed environment and since our secure part in proposal system is closed one so DES is the much reliable method for security in our authenticated database. Figure (3) display the DES [10] .

MAS

A multi-agent system (MAS*)* is a system composed of multiple interacting intelligent agents. Multi-agent systems can be used to solve problems which are difficult or impossible for an individual agent or monolithic system to solve. Examples of problems which are appropriate to multi-agent systems research include online trading, disaster response, and modeling social structures [11].

In the proposal, the aim of using MAS is to make the retrieval work in parallel on all database fields, see figure (4) and figure (5). We aim to make agents in a multi-agent system have several important characteristics these are **Autonomy**: the agents are at least partially autonomous for search and matching. **Local views**: the agents must have a full global view of the system, since the system is too complex for an agent to make practical use of sound knowledge. **Decentralization**: must be controlling agents to deal with overall database. Typically our multi-agent systems refer to software agents. Agents in a multi-agent system could equally well be adaptive procedures teams. A multi-agent system may contain combined adaptive procedure-agent teams. Multi-agent systems can manifest self-organization and complex behaviors even when the individual strategies of all their agents are simple [11].

PROPOSED DESIGN

The proposed system aim to enhance sound retrieval systems by dividing the sound retrieval methods into two parts: secure sound retrieval and normal sound retrieval. After displaying the both algorithms we will display the details of these two algorithms such as explaining the ANN used for noise removal, DES used for encryptions and finally MAS used for parallel matching process.

Secure Sound Retrieval

This part aim to build secure sound retrieval system for authorized customers by two steps, these steps are written as algorithms we called them algorithm1 and algorithm 2.

Algorithm 1 Building authentication
Input : complete keywords sound of authorized customers
Output: encrypted DB consist all the authorized sound

Process
Step $_0$: take complete keywords sound for authorized
Step $_1$:1.a- analyze the sound by taking the signal quantization
1.b- Convert it to samples
1.c- Take the samples and convert each one of them to sequence of numbers.
Step $_2$: convert information of sound analysis in each sample to binary model
Step $_3$: A - apply BP ANN on each binary sample to remove noise from the sound
B - encrypt the model using DES
Step $_4$: store the encrypted model in the DB
Step $_5$: end.

Algorithm 2 sound retrieval
Input : samples of authorized sound.
Output : matching the samples using MAS

Process.
Step $_0$: take the sample of sound
Step $_1$: 1.a- analyze the sound by taking the signal quantization
1.b- Convert it to samples
1.c- Take the samples and convert each one of them to sequence of numbers
Step $_2$: convert information of sound analysis to binary model.
Step $_3$: **A**- apply Bp ANN on each binary sample to remove noise from the sound
B- encrypt using DES
Step $_4$: matching fields of DB using Multi agent system to find the authentication.
Step $_5$: end.

Normal sound retrieval

This part aim to build normal sound retrieval system for normal customers by the following algorithm called algorithm3.

Algorithm 3 sounds matching.
Input : Samples of sound.
Output : Retrieve all the intended sounds

Process.
Step $_0$: take sound sample
Step $_1$: 1.a- analyze the sound by taking the signal quantization
1.b- Convert it to samples
1.c- Take the samples and convert each one of them to sequence of numbers..
Step $_2$: convert information of sound analysis to binary model.
Step $_3$: apply BP ANN on each binary sample to remove noise from the sound.
Step $_4$: while not end of DB Do
A- Match the model using MAS with all samples in DB.
B- Retrieve all the intended sounds.
end {while }
Step $_5$: end.

IMPLEMENTATION

In this section to explain the implementation of the proposed system we will focuses on display the basic points of sound retrieval for both cases authentication and normal. First step in both cases must be the sample of sound the user wants to retrieve, see figure (6).

The second step in both cases is the processes of converting the sound sample to Fourier transformation, which will be noised sample, see figure (7). Then remove noise of the sample of sound transformed in figure (6) by using ANN BP, see figure (8), finally encrypt the sample in the authentication case only, see figure (9). Then by using MAS in both cases the retrieval process will be done to extract all related sounds matching the sound sample recorded in figure (6), see figure (10) which will introduce the overall process of proposal.

DISCUSSION AND CONCLUSIONS

Before introducing our conclusions we will introduce modest comparisons results obtained from implementing the proposed system. This comparison included comparing the traditional algorithms of sound and the both proposed algorithms normal and authentication with considering the most famous four attributes these are speed, performance, accurate and security with 100% ratio.

Algorithms/attributes	Traditional Alg.	Normal Alg.	Authentication Alg.
Speed	**80%**	**95%**	**85%**
Performance	**85%**	**90%**	**90%**
Accurate	**90%**	**90%**	**90%**
Security	**70%**	**70%**	**90%**

While we build that theoretical system we reach to the following conclusions:

1. Sound retrieval still has incomplete frame.
2. Using ANN BP make the sound retrieval deal freely with noise.
3. Using MAS make the matching process much more speed since it work in parallel in passing the databases.
4. Using DES for encryption very suitable with sound retrieval DB, since it is deal with close environment.

REFERENCES

[1] Rabiner, L. and B.-H. Juang. Fundamentals of speech recognition. Prentice All, first edition, 1993.

[2] Reynolds, D. A. T. F. Quatieri, and R. B. Dunn. Speaker verification using adapted gaussian mixture models. Digital Signal Processing, 10(1–3), 2000.

[3] Slaney, M. I. Center, and C. San Jose. Semantic-audio retrieval. In ICASSP, volume 4, 2002.

[4] Turnbull, D. L. Barrington, D. Torres, and G. Lanckriet. Towards musical query by semantic description using the cal500 data set. In SIGIR '07:30th annual international ACM SIGIR conference on Research and development in information retrieval, pages 439–446, New York, NY, USA, 2007. ACM.

[5] Turnbull, D. L. Barrington, D. Torres, and G. Lanckriet. Semantic annotation and retrieval of music and sound effects. In IEEE Transactions on Audio, Speech and Language Processing, 2008.

[6] Vapnik, V. The nature of statistical learning theory. SpringerVerlag, 1995.

[7] von Ahn, L. and L. Dabbish. Labeling images with a computer game. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 319–326. ACM Press New York, NY, USA, 2004.

[8] Wan, P. and L. Lu. Content-based audio retrieval: a comparative study of various features and similarity measures. Proceedings of SPIE, 6015:60151H, 2005.

[9] Jiawei Han, Micheline Kamber, *Data Mining: Concepts and Techniques*,2005, www.cs.sfu.ca/~han/DM_Book.html

[10] Katos, V., A Randomness Test for Block Ciphers. *Applied Mathematics*, Elsevier Publications, 2005.

[11] Tom Mitchell, *Machine Learning*, McGraw Hill, 2006

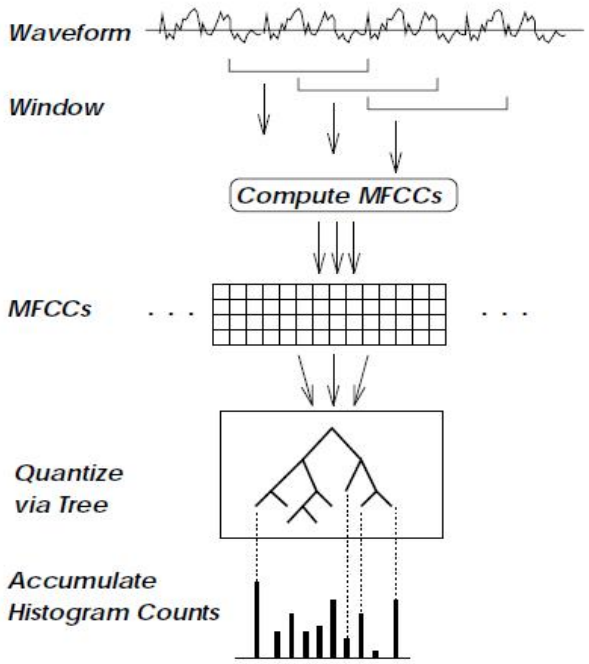

Figure (1) Audio template construction

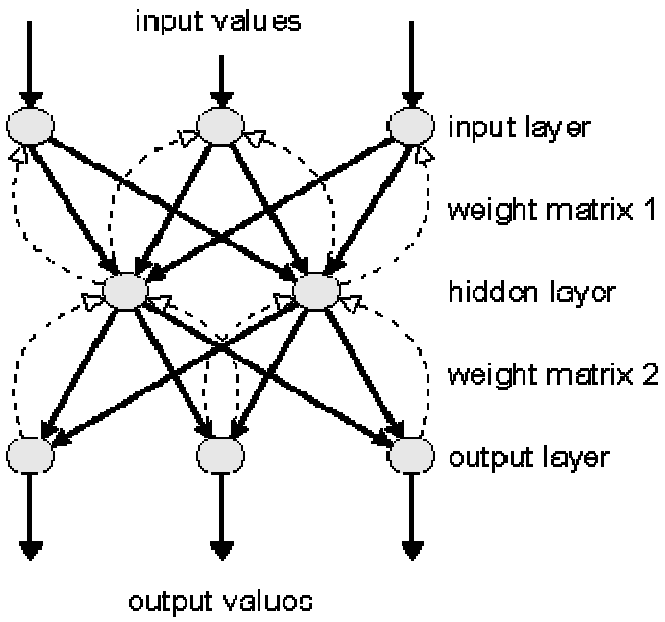

Figure (2): ANN BP

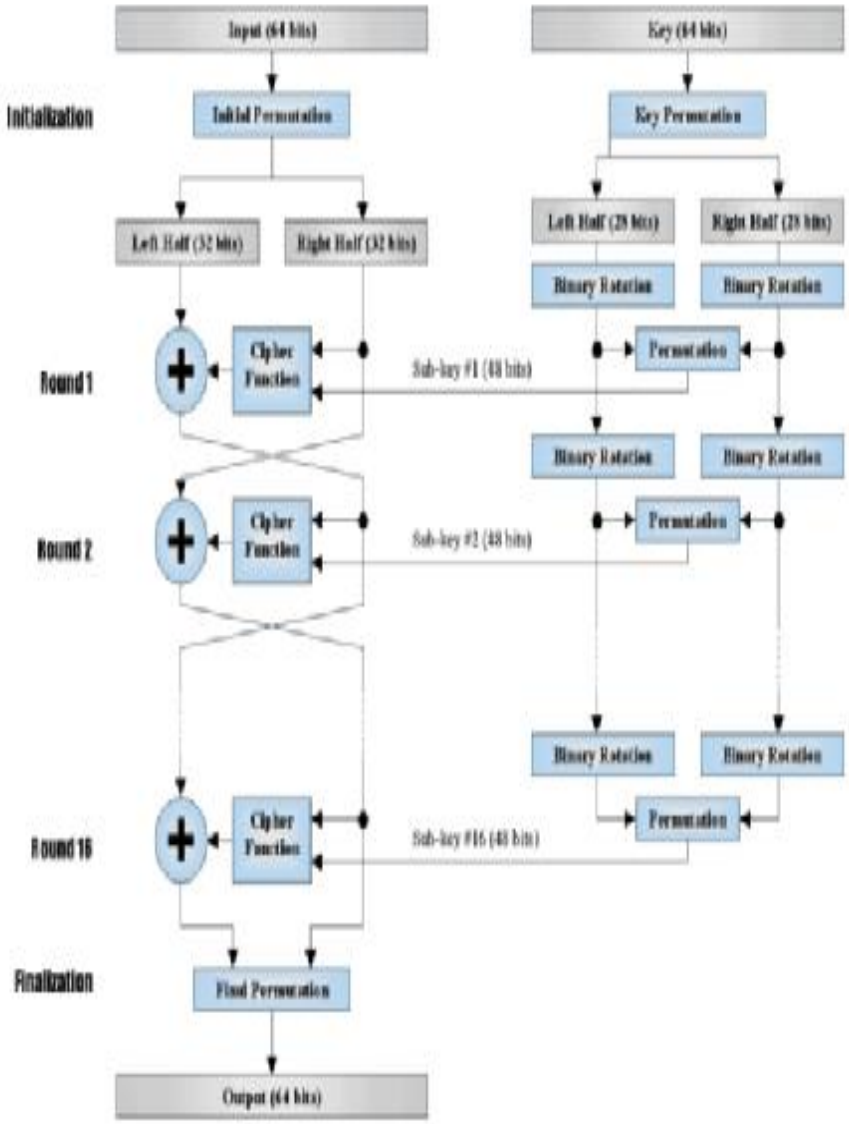

Figure (3): DES encryption method.

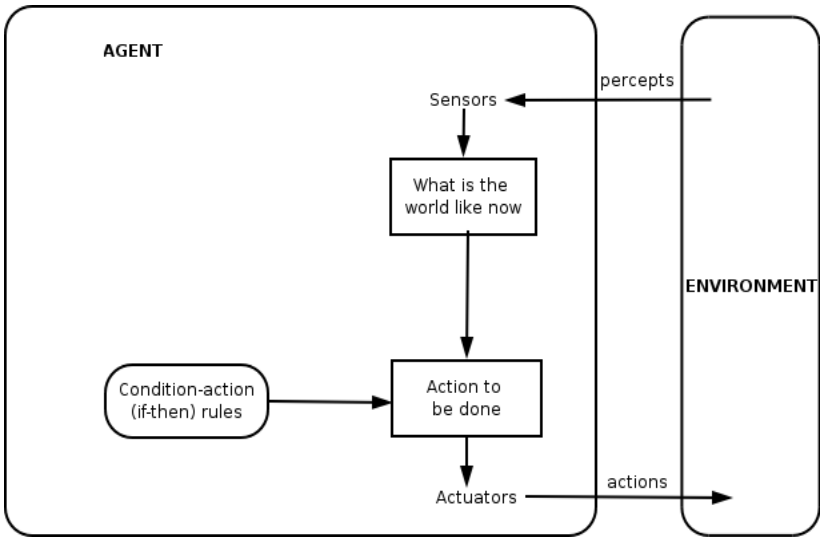

Figure (4): Simple reflex agent

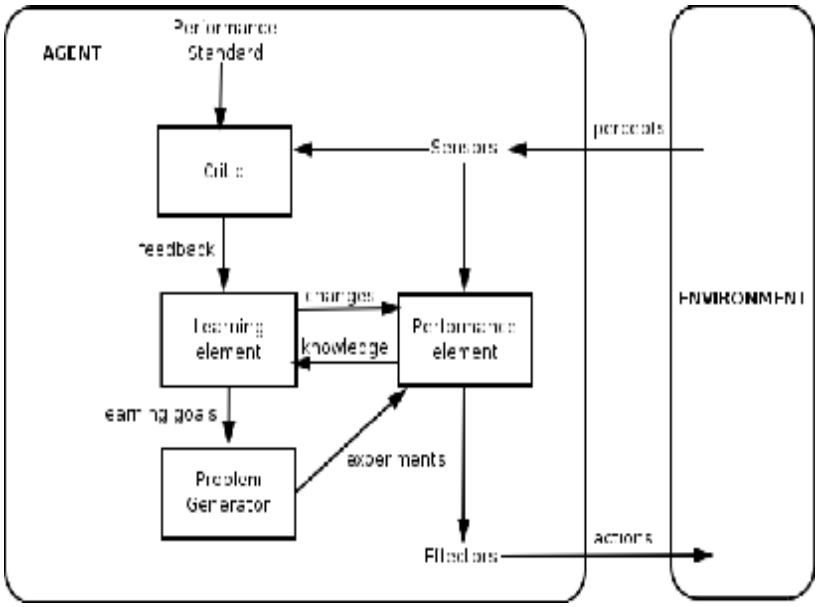

Figure (5): learning agent

Figure (6): recording the sample of sound wanted to be retrieved.

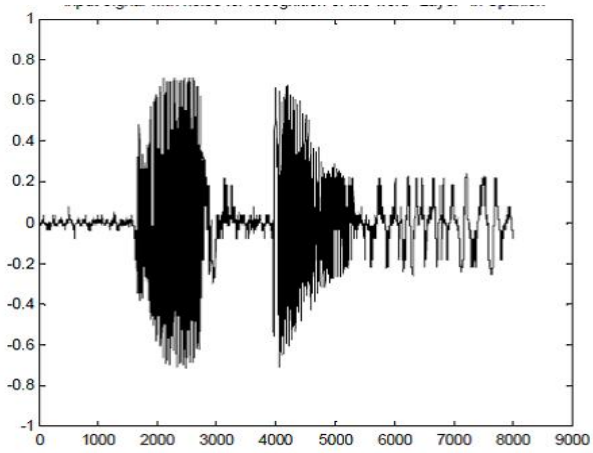

Figure (7): sample of sound after Fourier transformation with noise

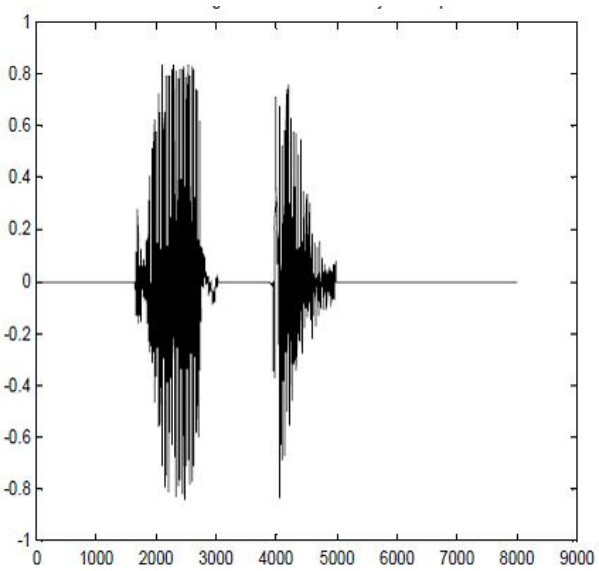

Figure (8): sample of sound after noise removal using BP ANN.

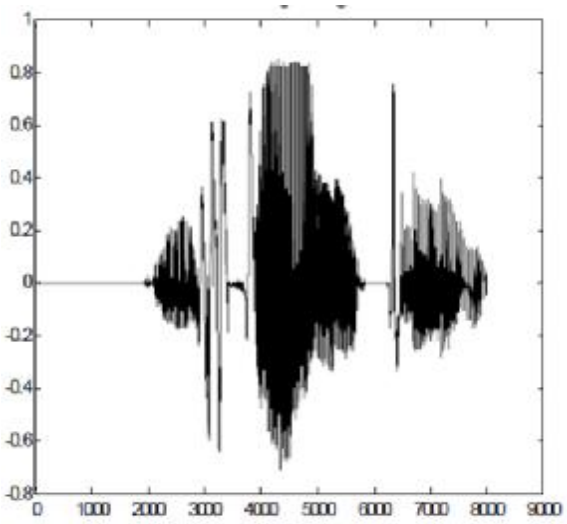

Figure (9): encrypted sample of sound after noise removal

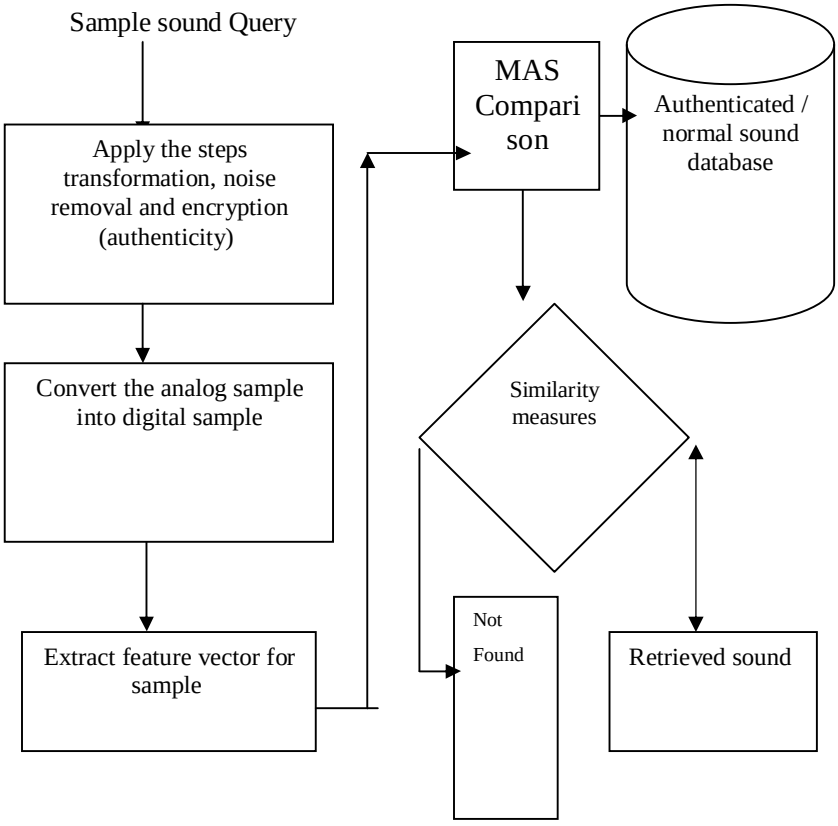

Figure (10): retrieve comprehensive database according to the query.