# Different Indexing Techniques

**Dr. B. B. Meshram Ganesh P. Gaikwad**
M.Tech Computer Engineering VJTI Mumbai

## Abstract

**This paper describes about Audio Indexing, Video Indexing, Content Based Image Indexing, and Content Based Multimedia Indexing i.e. Content-based indexing techniques. Indexing is concerned with compactly storing a large collection of terms and rapidly retrieving a set of candidate terms satisfying some property from a large collection of terms. Index is a structure or object in the database Indexes provide fast access to rows in tables meeting certain condition in a query. In this paper audio indexing, Audio Feature Extraction and Fingerprint Indexing are concerned with accurate and efficient indexing of fingerprint images.**

**Index Term:** Content-based image indexing and retrieval (CBIR), Audio Indexing, AFeX, Content-based video indexing,

## I. INTRODUCTION

Index is a data structure that organizes data records on disk to optimize certain kind of retrieval operation. Index allows us to efficiently retrieve all records that satisfy search condition on the search key field of the index. An index is a structure or object in the database indexes provide fast access to rows in tables meeting certain condition in a query. Indexing is concerned with compactly storing a large collection of terms and rapidly retrieving a set of candidate terms satisfying some property from a large collection of terms.

Indexes provide fast access to rows in tables meeting certain condition in a query. Indexing is Minimize I/O, improving performance, especially in random access.

It is a dynamic structure: Changes as the data in the table changes. A table can have several indexes, to satisfy several queries. Database indexes are similar to indexes in books.

In indexing the Multiversion indexing approaches the options for extending single version indexing schemes to handle multi version data. In Multiversion indexing there are four different approaches: Chaining (CH), Data Page Version Selection (DP), Primary Index Version Selection (PI), and All Index Version Selection (AI).

In recent years the fast growth of multimedia information, content based video analysis, indexing and retrieval have attracted increasing attention. Many applications have emerged in areas such as video-on-demand, distributed multimedia systems, digital video libraries, distance education, entertainment, surveillance and geographical information systems. Content-based video analysis aims at obtaining a structured organization of the original video content and understanding its embedded semantics like humans do.

## II. RELATED WORK

### A. CONTENT BASED IMAGE INDEXING AND RETRIEVAL SYSTEM

In recent year Digital image libraries and other multimedia databases have been largely expanded, so storage and retrieval of images in such libraries has become a real demand in industrial, medical, and other applications. So the Content-based image indexing and retrieval (CBIR) is considered as a solution.

There are generally two different ways to enhance the performance of CBIR system:

1. Retrieval and
2. Indexing algorithms

1) Retrieval Algorithm Enhancement: Initially all images of an imagebase are indexed and their indexes are stored in a feature database (featurebase). Then, at each step of the enhancement process, similar images to each query are retrieved using the current retrieval algorithm parameters. Finally, the retrieval algorithm parameters are modified such that the similarity between the CBIR and user retrieved images (A and B respectively) for all queries is maximized.

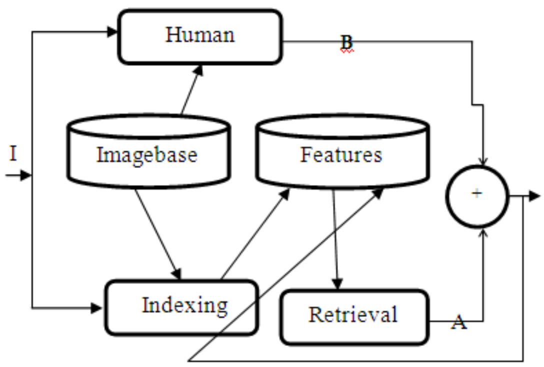

Fig. Block diagram of the indexing algorithm enhancement process in a CBIR system

If only one query image (**I**) is used in the enhancement process, the retrieval algorithm parameters will be optimized to evaluate the similarity between that query image and all the images in the database.

2) Indexing Algorithm Enhancement: In Indexing algorithm enhancing the indexing parameters only for one query image is not meaningful. The aim of enhancement is global optimization of the indexing algorithm parameters. Optimization of the indexing algorithm is a more difficult task compared to the retrieval algorithm enhancement. Because each time the indexing algorithm parameters are modified, all images of the reference imagebase should be indexed again.

## B. VIDEO INDEXING

In the recent year there is fast growth of multimedia information, content based video analysis and indexing and retrieval have attracted increasing attention. So the many applications have emerged in areas such as video-on-demand, distributed multimedia systems, digital video libraries, distance education, entertainment, surveillance and geographical information systems. Content-based video analysis aims at obtaining a structured organization of the original video content and understanding its embedded semantics like humans do. Content-based video indexing is the task of tagging semantic video units obtained from content analysis to enable convenient and efficient content retrieval.

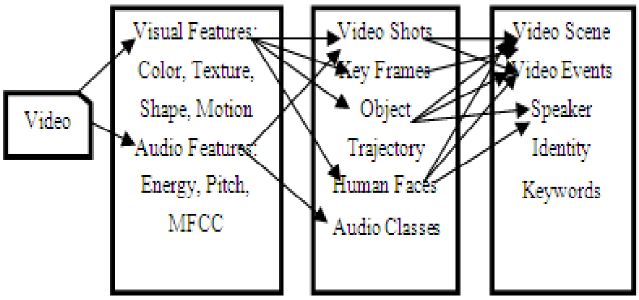


Fig.Generic three-level video indexing structure

First extract some low- to mid-level audiovisual features, then partially derive or understand the video semantics by analyzing and integrating these features. Fig. 1 shows a hierarchical video indexing structure, where many popularly used features such as color, texture, shape, motion, shots, keyframes, object trajectories, human faces, as well as classified audio classes, constitute the low- to mid-level indexing features. Semantic gap still exists between the real video content and the video contexts derived from these features. This work proposes to extract two types of video indexing features, namely, video events and speaker identity, at the semantic level based on the integration of audio and visual knowledge.

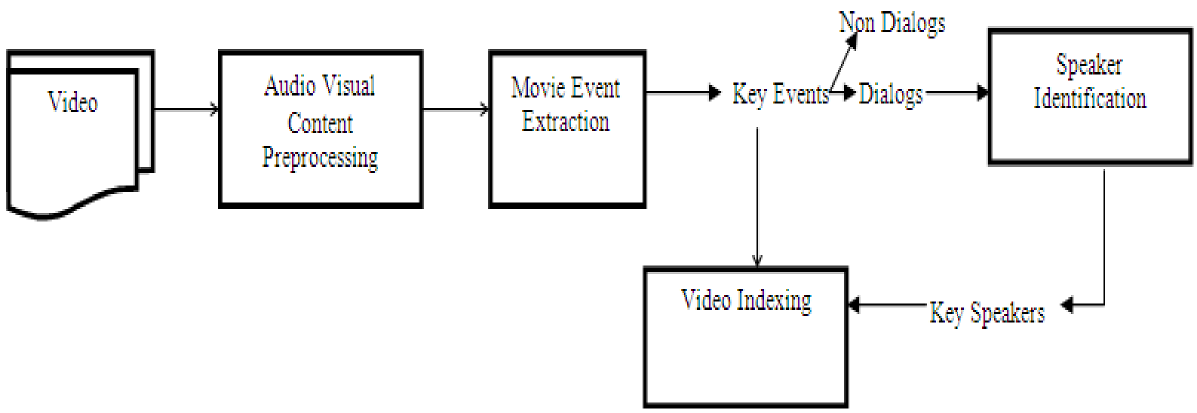


Fig. Overview of System Framework

**C. AUDIO INDEXING**

**1. Indexing Framework in MUVIS**

In order to provide both visual and aural indexing schemes MUVIS supports both visual and aural feature extraction frameworks in such a way that feature extraction modules can be developed independently and integrated into MUVIS to develop their own feature extraction modules and integrate System exclusively (during the run-time). This is the basis of the framework structure, which therefore allows third partied into MUVIS without knowing the details of the MUVIS applications.

**2. Audio Indexing and AFeX Framework**

Audio indexing is accomplished in several stages: Classification and Segmentation, Audio Framing within segments with certain

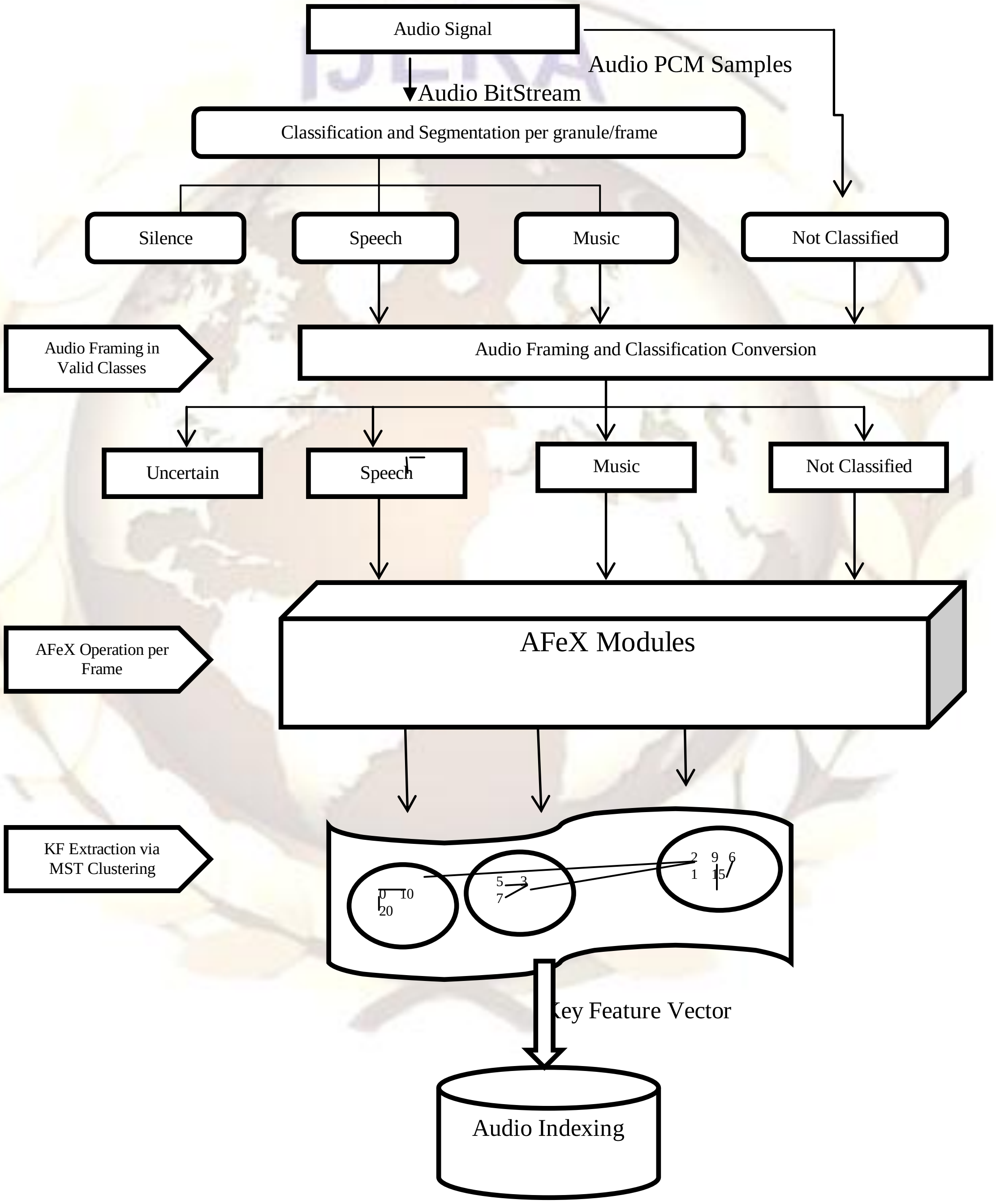


Fig. Audio Indexing Operation Flowchart

types, Audio Feature Extraction (AFeX), Key-Framing via Minimum Spanning Tree (MST) Clustering and finally the indexing over the extracted Key-Frames (KFs).

### 3. Audio Classification and Segmentation

In order to achieve a better content analysis the first step is to perform classification and segmentation over the entire audio clip. In MUVIS system this operation is performed over AAC and MP3 bitstreams directly from the bitstream information. The method is generic and robust to several capture and encoding parameters.

The output from the classification and segmentation is the segments with certain duration in time and a certain class type such as speech, music and silence. During the feature extraction operation, the feature vectors are extracted from each individual segment with a different class type and stored and retrieved separately. This makes more sense for content-based retrieval and brings the advantage to perform the similarity comparison between the frames within the segments with a matching class type.

### 4. Audio Framing within Valid Audio Segments

There are three valid audio segments: speech, music and NotClassified. Segmentation and classification are performed per granule/frame basis, such as per MP3 granule or AAC frame, a conversion is needed to achieve a generic audio framing for indexing purposes. The entire audio clip is first divided into a user or model-defined audio frames, each of which will have a classification type as a result of the previous step.

The removal of audio frames in Uncertain and Silence types from AFeX operation has two advantages: first there is no need for AFeX operation on silent frames since no content information can be retrieved from them. Similarly the Uncertain frames are mixed and hence most of the times they are transition frames (i.e. music to speech, speech to silence, etc.).

### 5. Audio Feature Extraction (AFeX) Framework

Once audio framing is completed, feature extraction is applied onto certain frames with a valid class types (music, speech or NotClassified) for indexing. FeX framework for visual feature extraction has been presented in section 3.1. A similar framework (so called AFeX) has been developed to accomplish audio feature extraction operations. Therefore, AFeX framework mainly supports dynamic audio feature extraction module integration for audio clips and such a framework shall form a common basis to compare, merge and experiment among several exclusive AFeX modules to develop new algorithms and to improve efficiency.

### 6. Key-Framing via MST Clustering

The number of audio frames is proportional with the duration of the audio clip and once AFeX operation is performed, this may potentially result in a massive number of feature vectors, many of which are probably redundant due to the fact that the sounds within an audio clip are immensely repetitive and most of the time entirely alike. In order to achieve an efficient audio-based retrieval within an acceptable time, only the feature vectors of the frames from different sounds should be stored for indexing purposes. This is indeed a similar situation with the visual feature extraction scheme where only the visual feature vectors of the Key-Frames (KFs) are stored for indexing.

### D. FINGERPRINT INDEXING

Fingerprints are undoubtedly one of the most studied biometric traits and the most widely used in civil and forensic recognition systems. Fingerprint comparison algorithms are fast and quite accurate; identifying an unknown fingerprint over a large database poses challenging problems in terms of both accuracy and efficiency. To reduce the total number of comparisons, fingerprint identification systems typically use prefiltering techniques. These techniques can be grouped into two main categories: 1) Exclusive classification and 2) Fingerprint indexing. Fingerprint indexing technique based on the Minutia Cylinder-Code (MCC) representation.

Exclusive classification techniques split the database into a fixed number of classes; during the identification phase, the searched fingerprint is compared only to fingerprints belonging to the same class [2]. The main drawback is that the number of classes is usually small and fingerprints are unevenly distributed among them: When a single fingerprint has to be searched in a large database, exclusive classification is often not able to sufficiently narrow down the search.

## III. PROPOSED WORK

Online audio indexing system, which creates a searchable index of speech content embedded in digitized audio files. This system is based on our recently used offline audio segmentation techniques.

The data arrives continuously; the system first finds boundaries of the acoustically homogenous segments. Next, each of these segments is classified as speech, music or mixture classes, where mixtures are defined as regions where speech and other nonspeech sounds are present simultaneously and noticeably. The speech segments are then clustered together to provide consistent speaker labels. The speech and mixture segments are converted to text via an ASR system. The resulting words are time-stamped together with other metadata information (speaker identity, speech confidence score) in an XML file.

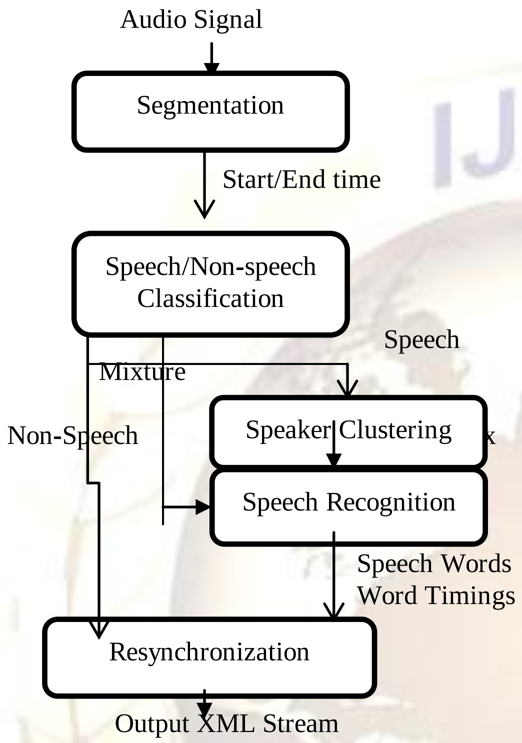


Fig. Online Audio Indexing System

## III. CONCLUSION

An ideal content-based video analysis and indexing system should offer flexible and efficient tools for video browsing and retrieval. The design of such a system demands effective ways to extract semantic information from various media sources such as audio, visual and textual, such that multi-level video abstraction can be efficiently performed. This work presented a content- based movie analysis and indexing scheme which aims at extracting semantically meaningful movie events and identifying target speakers in movie dialogs.

MUVIS framework is designed to bring a unified and global solution to content-based multimedia indexing and retrieval problem. The unified solution is basically achieved by designing the whole system of applications, which are handling the media during its life-time starting from capturing till indexing in such a way that the media can be indexed as efficiently as possible. The framework covers a wide range of media types and formats for handling, indexing and retrieving. It supports browsing, hierarchic video representation and summarization. Most important of all, MUVIS framework supports integration of the aural and visual feature extraction algorithms explicitly. This brings a significant advantage for third parties to develop and test several feature extraction modules that are independent from the MUVIS applications.

## REFERENCES

[1] Silberschatz, Korth and Sudarshan Database System Concepts, 5th Ed.

[2] Paul M. Bober and Michael J. Carey Indexing for Multiversion Locking: Alternatives and Performance Evaluation IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 9, NO. 1, JANUARY-FEBRUARY 1997.

[3] Hung-Min Sun_, Bying-He Ku_, Chung-Chi Wu_ and Ting-Yu Lin_Department of Computer Science National Tsing Hua University, Hsinchu, Taiwan Improving VQ Index Compression with Modified Relative Index Table 2010 First ACIS International Symposium on Cryptography, and Network Security, Data Mining and Knowledge Discovery, E-Commerce and Its Applications, and Embedded Systems

[4] Mahdi Saadatmand-Tarzjan and Hamid Abrishami Moghaddam A Novel Evolutionary Approach for Optimizing Content-Based Image Indexing Algorithms IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 37, NO. 1, FEBRUARY 2007

[5] Dana Codreanu Université Paul Sabatier IRIT UMR 5505 Université de Toulouse, France Sebastien Laborie LIUPPA Université de Pau et des Pays de l'Adour Anglet, France Modelling Multimedia Indexing Algorithms Description for Implicit and Explicit Indexation 2011 Fifth FTRA International Conference on Multimedia and Ubiquitous Engineering

[6] Ying Li, Member, IEEE, Shrikanth Narayanan, Senior Member, IEEE, and C.-C. Jay Kuo, Fellow, IEEEContent-Based Movie Analysis and Indexing Based on Audiovisual Cues IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR

VIDEO TECHNOLOGY, VOL. 14, NO. 8, AUGUST 2004 1073.

[7] Raffaele Cappelli, Member, IEEE, Matteo Ferrara, and Davide Maltoni, Member, IEEE Fingerprint Indexing Based on Minutia Cylinder-Code IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 33, NO. 5, MAY 2011.

[8] Moncef Gabbouj and Serkan Kiranyaz Institute of Signal Processing Tampere University of Technology Tampere, Finland AUDIO-VISUAL CONTENT-BASED MULTIMEDIA INDEXING AND RETRIEVAL – THE MUVIS FRAMEWORK.