Performance Evaluation of Multi Minimum Product Spanning Tree Index Structure on N-Feature Dimensions of Video Files

Meenakshi Srivastava¹ & Dr. S.K.Singh²
¹Amity Institute of Information Technology, Amity University, Uttar Pradesh, India
²Amity Institute of Information Technology, Amity University, Uttar Pradesh, India

Received: March 30, 2018 Accepted: May 09, 2018

ABSTRACT
In order to provide more relevant results, searching multimedia data such as video files requires analysis of content in the data. The content of the video files can be modelled with different features as well as on the basis of metadata. The selection of the features plays an important role in the construction of retrieval system. The information contained in video files can be analysed in a variety of ways depending upon various feature dimensions. Maximum of present research work on video information retrieval focuses on retrieval algorithms which are based on similarity search by extracting and matching the feature from the video like object motion, colour, shape, texture, loudness, power spectrum, bandwidth, and pitch etc. For extracting similarity based on multiple dimensions an efficient indexing algorithm is very much required. In our earlier research it has been shown that Multi Minimum Product Spanning Tree based index structure can easily map n number of features like shape, texture, colour etc of any multimedia object. In present paper performance of Multi Minimum Product Spanning Tree (MMPST) Index Structure has been evaluated on n features of multimedia objects like video files. Our results show that MMPST can easily provide the best similar records in minimum number of similarity check.

Keywords: Multi Minimum Product Spanning Tree; Indexing; Multimedia Information Retrieval; Content Based Retrieval.

1. Introduction
There are two basic components of a retrieval system first one is a similarity measurement module and the second one is a multidimensional indexing algorithm. Similarity measurement module is used to find the most similar objects, and the multidimensional indexing is used to accelerate the query performance in the search process. To measure the similarity, the general approach is to represent the data features as multidimensional points in metric and then calculate the distances between the corresponding multidimensional points. Selection of metrics has a direct impact on the performance of a retrieval system [1]. Euclidean distance is the most common metric used to measure the distance between two points in multi-dimensional space. However, for some applications, Euclidean distance is not compatible with the human perceived similarity [2]. Another limitation with similarity measurement module is that performance of retrieval decreases when high dimensionality data is compared to compute similarity. As a solution Principal Component Analysis (PCA) [3] is utilized to change the first information of high dimensionality into another facilitate framework with low dimensionality by discovering information with high segregating power. The new facilitate framework evacuates the excess information and the new arrangement of information may better represent the data. These second factor which affects the performance of retrieval system is the algorithm which used for indexing. Indexing algorithm maps the objects in the database with the preferable keywords or features so that searching for a specific multimedia object can be done in minimum time. Content based retrieval of multimedia objects such video files deals with semantic gap [4]. The main reason for semantic gap in video retrieval is that human eyes can easily compare/match two videos on different features and can easily state that object wise two videos are 50% similar and colour wise 70% similar etc. Whereas the maximum of existing methodologies of video retrieval group the objects with maximum similarity in a class or cluster and indexing is done on selected key features. Such type of retrieval achieves maximum similarity based on selected features only and fails to provide different level of similarity on various features. In [5] authors had proposed Muti Minimum Product (MMPST) based indexing approach. MMPST indexes the multimedia object on the basis of their level of similarity. The main advantage of MMPST is that no features can be mapped without adding cost on tree construction. In [6] MMPST has been constructed on two features of bio images. In present manuscript performance of MMPST has been evaluated on three features of video files. Result shows that increasing the number of feature does not have effect the retrieval speed.
2. State Of Art
Research in content based video recovery today is an exuberant restrained, growing in breadth [7]. Video Retrieval Based on Textual Queries [8] introduced an approach that empowers look based on the printed data display in the video. Locales of printed data are indented inside the edges of the video. Video is then commented on with the literary substance show in the pictures. Programmed Content-Based Retrieval and Semantic Classification of Video Content [9] displayed a learning system where development of an abnormal state video record is imagined through the union of its arrangement of natural highlights. This is done through the medium of help vector machines (SVM). The help of vector machines relate each arrangement of information focuses in the multidimensional include space to one of the classes amid preparing. In Content-Based TV Sports Video Retrieve Based on Audio-Visual Features and Text [8] creators propose content-based video recovery, which is a sort of recovery by its semantical substances. Since video information is made out of multimodal data streams, for example, visual, sound-related and printed streams, creators portray a technique of utilizing multimodal examination for programmed parsing sports video. Video Retrieval of Near- Duplicates utilizing k- Nearest Neighbour Retrieval of Spatio- Temporal Descriptors [9] depicts a novel procedure for actualizing video seek capacities, for example, recovery of close copy recordings and acknowledgment of activities in reconnaissance video. Recordings are partitioned into half-second clasps whose stacked edges create 3D space-time volumes of pixels[10]. Pixel areas with reliable shading and movement properties are removed from these 3D volumes by a limit free various levelled space time division procedure. Every area is then portrayed by a high-dimensional point whose parts speak to the position, introduction and, when conceivable, shade of the district. In the ordering stage for a video database, these focuses are doled out names that determine their video clasp of starting point. All the marked focuses for every one of the clasps are put away into a single twofold tree for productive k-closest neighbour recovery. The recovery stage utilizes video portions as questions. Work introduced in Fast Video Retrieval by mans of the Statistics of Motion Within the Regions-of-Interest [11] manages vital issue to rapidly recover semantic data from an immense interactive media database. In this work, authors proposed a measurement based calculation to recover the recordings that contain the asked for question movement from video database. The survey of literature shows that retrieval of video object based on different level of similarity in the feature is a goal which needs to be farfetched.

3. Indexing of Video Data and Retrieval
Video indexing is a procedure of labeling recordings and sorting out them in a powerful way for quick access and recovery. Mechanization of ordering can altogether decrease preparing cost while wiping out monotonous work [6]. The regular highlights utilized as a part of a large portion of the current video recovery frameworks are the highlights, for example, shading, surface, shape, movement, confront, sound, classification and so forth. Clearly more the quantity of features used to depict the information, better the retrieval accuracy. In any case, as the element vector measurement increments with expanding number of highlights, there is an exchange off between the retrieval accuracy and many-sided quality [12]. So it is fundamental to have insignificant highlights representing the information, minimally. In present work multiple features of a Video Frame like Texture, Color, Shape has been represented in matrix and Multi Minimum Product Spanning Tree based Index structure is constructed.

3.1. Feature Extraction
A brief introduction of various extracted features is given below.
3.1.1 Texture Features
Surface can be characterized as the visual examples that have properties of homogeneity that don’t come about because of the nearness of just a solitary shading or force. Gray co-occurrence matrix (GLC) is one of most rudimentary and critical techniques for surface include extraction and portrayal. Its unique thought is first proposed in Julesz (1975) [13]. Julesz found through his well known investigations on human visual impression of surface, that for a substantial class of surfaces no surface combine can be segregated in the event that they concur in their second-arrange insights [14].

3.1.2 Color Feature
Color is a standout amongst the most broadly utilized visual highlights in mixed media setting and picture/video recovery, specifically [15].

3.1.3 Shape Features
An arrangement of estimations is made which freely describe some part of the shape. An extensive accumulation of cases, portray the shape measurably. Measurably example acknowledgment approach has been common for a long time for shape acknowledgment. Straightforward district developing calculation is
to fragment a high contrast picture in locales. Automation of indexing can significantly reduce processing cost while eliminating tedious work [16].

3.2 Indexing on Multiple Features

Video ordering is a procedure of labelling recordings and sorting them in a viable way for quick access and recovery. Computerization of ordering can essentially diminish handling cost while disposing of monotonous work [4]. The conventional features used in most of the existing video retrieval systems are the features such as colour, texture, shape, motion, object, face, audio, genre etc. It is obvious that more the number of features used to represent the data, better the retrieval accuracy. It has been stated by many researches that when the retrieval system is designed on only limited low level features, semantic gap exists [3]. Generally a retrieval system designed on low level features, tends to find the most similar object rather than searching for an extent on multiple features. For enabling retrieval which is based on multiple features as well up to different extent of similarity on individual feature requires a robust indexing algorithm. Multi Minimum Product Spanning Tree can efficiently index the object on the basis of maximum similarity on various features.

4 Methodology and Implementation

Java programming environment is utilized for implementation. Eclipse, Jcreator, Oracle 9i are utilized for the advancement of the framework. Proposed framework is executed for the java virtual machine empowered windows based pc's with web association. Test recordings and casings were downloaded from www.archive.org[17]. We have considered distinctive classifications of pictures like e-learning. The texture, Colour and shape features were extracted from video file. Each extracted features are represented in vector. Principal Component analysis is performed on feature vector to reduce the dimensionality. Euclidean distance is computed to set the similarity relation on individual features. Depending upon the distance between two video files percentage of similarity is computes. These percentages of similarity are represented in matrices as shown in Fig. 1.

**Figure 1: Sample of Feature Similarity Matrix on Three Features**

Adjacency matrix shown in Fig 1 is used to create the graph. Process of creation of MPST (on one feature) is shown Fig. 2. Now at the next step the Minimum Spanning Tree (MST) will be created by evaluating the minimum products. Once the Individual MST on different feature matrix is created, Multi Minimum Product Spanning Tree (MMPST) is created on three MST using their product. MMPST identifies the maximum similarity using all three features as root node.

![Fig. 2: Process of Creation of MPST](image-url)
MMPST index maps the database object using two separate indexes, the first index is dense structure which contains the percentage of similarity between objects and the second one is sparse index which contains different level of possible similarity on individual features. Memory mapping of sparse and dense structure is shown in Fig3.

<table>
<thead>
<tr>
<th>Sparse Index on Object</th>
<th>Data File</th>
<th>Dense Index on F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>V1, 50, 75,80</td>
<td>10</td>
</tr>
<tr>
<td>V2</td>
<td>V2, 40, 90, 10</td>
<td>20</td>
</tr>
<tr>
<td>V3</td>
<td>V3, 60, 40, 60</td>
<td>60</td>
</tr>
<tr>
<td>V4</td>
<td>V4, 50,60,75</td>
<td>70</td>
</tr>
<tr>
<td>V5</td>
<td>V5, 30, 10, 70</td>
<td>75</td>
</tr>
<tr>
<td>V6</td>
<td>V6, 40, 50, 100</td>
<td>80</td>
</tr>
<tr>
<td>V7</td>
<td>V7, 40, 50, 100</td>
<td>100</td>
</tr>
<tr>
<td>V9</td>
<td>V9, 65, 45, 20</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3 Mapping of Sparse Index, Dense Index and Data File

Fig. 3 represents a snapshot memory mapping between Sparse Index of Objects, Data File and Dense Index created over feature 3(F3).

4. Results and Discussion

We have performed a set of query. Short video clips of 30 seconds were utilized. The threshold value T and no frames were varied. The first four runs were executed keeping the low threshold, and no of frames captured were 50. The threshold value was increased for next four run and no of captured frames were set to 70. For last four runs the threshold value was kept to maximum and no of frames to be captured was set to 100. Same set of experiments were reported for all features. Table 1 gives a summary of precision rate achieved at various level of threshold and no frames for different features.

| Table 1 Precision at 50, 70, 100 frames.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shape Feature</td>
<td>Colour Feature</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Precision at 50 frames</td>
<td>0.333</td>
<td>0.513</td>
</tr>
<tr>
<td>Precision at 70 frames</td>
<td>0.321</td>
<td>0.528</td>
</tr>
<tr>
<td>Precision at 100 frames</td>
<td>0.880</td>
<td>0.782</td>
</tr>
</tbody>
</table>

Table 1 displays that better precision rate was recorded at maximum no of captured frames and highest Threshold value.

After achieving satisfactory result in video retrieval through varied threshold values, the relation between two video file were set in the form of percentage of similarity. The similarity values in the form of percentage were stored in different matrix and the steps for creation of MMPST were followed.

Now a set queries were executed on MMPST, and the retrieval result were compared with the retrieval speed of Inverted Index. A set of queries like “Show all videos which are 70% colour wise similar and 90% texture wise similar” had been executed in time efficient manner. Comparisons of query evaluation between MMPST and Inverted Index are given in Table 2. The results clearly shows that retrieval time taken by MMPST index is lesser that the time taken by Inverted Index.
Video Retrieval is expansive territory that incorporates features from various fields of computer science including AI, machine learning, Database Management etc. There have been huge quantities of calculations attached in these fields to perform different video recovery assignments. Maximum of present research work done in the field of video retrieval focus on overall similarity on selected features, rather than different levels of similarity based on selected features. In present paper video retrieval is improved by incorporating different features of a video file as well as the performance of Multi Minimum Product Spanning Tree is evaluated through various queries based on multiple features of video. Experimental outcomes demonstrate that combination of extricated features enhances video indexing and recovery. There can be expansive territory of use of proposed algorithm in criminal database recovery, biomedical data, and social media data analysis etc.

References

23. www.archive.org

Table 2 Result of Query execution on MMPST and Inverted Index

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
<th>MMPST</th>
<th>Inverted Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total Records</td>
<td>Time in msec.</td>
</tr>
<tr>
<td>Q2</td>
<td>Texture Similarity =90% Color Similarity= 80%</td>
<td>8</td>
<td>54</td>
</tr>
<tr>
<td>Q10</td>
<td>Shape Similarity =80% Color Similarity=90%</td>
<td>5</td>
<td>48</td>
</tr>
<tr>
<td>Q15</td>
<td>Texture Similarity =70% Shape Similarity=80%</td>
<td>6</td>
<td>34</td>
</tr>
</tbody>
</table>

Conclusion

Video Retrieval is expansive territory that incorporates features from various fields of computer science including AI, machine learning, Database Management etc. There have been huge quantities of calculations attached in these fields to perform different video recovery assignments. Maximum of present research work done in the field of video retrieval focus on overall similarity on selected features, rather than different levels of similarity on selected features. In present paper video retrieval is improved by incorporating different features of a video file as well as the performance of Multi Minimum Product Spanning Tree is evaluated through various queries based on multiple features of video. Experimental outcomes demonstrate that combination of extricated features enhances video indexing and recovery. There can be expansive territory of use of proposed algorithm in criminal database recovery, biomedical data, and social media data analysis etc.