Similarity Search and Identification of Online Videos with Graph Transformation Techniques

Mr. Susheel George Joseph

M.C.A, M.Tech, M.Phil(CS)

(Assistant Professor, Department of M.C.A, Kristu Jyoti College of Management and Technology,
Changanassery, Kerala. susheelgj@gmail.com)

_____

***Abstract* - This project presents a matching approach for the video subsequence identification problem. The subsequence of any original video can be copied and by the other users without their permission. A query processing technique is used to retrieve similar frames, i.e., the mapping relationship between the input query and database video. The query processing is done for matching the video subsequence even if the temporal order made in the duplicate copies i.e. the single frame may be inserted or the order of the frames may be changed. The second part is implement the Ordinal signature in the video files, where the different copies with same content should have the similar signature either it may be any format. If the similar copies are found with different keys then it is assumed to be duplicate copy.**

***Keyword* - Video Signature, Ideal Video Similarity, Dynamic Query Ordering, Ordinal Signature**
_____

I. INTRODUCTION AND SCOPE

At present, it is often undesirable to manually check whether a video is part of a long stream by browsing its entire length; thus, a reliable solution of automatically finding similar content is imperative using content based searching.

In this paper I use query processing technique for temporal order database and find out the similar video in the stream, if there exists some transformation distortion, and insertion that would be a duplicate copies. Next Ordinal signature method is used for assigning the keys to the videos with same content and find out the duplicates with different keys but same content.

Recognition for copyright enforcement

Video content owners would like to be aware of any use of their material, in any media or representation. For example, the producers of certain movie scenes may want to identify whether or where their original films have been reused by others, even with some kind of remixing for multimedia authoring.

TV commercial detection

Some companies would like to track their TV commercials when they are aired on different channels during a certain time period for statistic purpose. They can verify whether their commercials have been actually broadcasted as contracted, and it is also valuable to monitor how their competitors conduct advertisements to apprehend their marketing strategies.

Video subsequence identification is that, while retrieval task conventionally returns similar clips from a large collection of videos which have been either chopped up into similar lengths or cut at content boundaries, subsequence identification task aims at finding if there exists any subsequence of a long database video that shares similar content to a query clip.

With the growing demand for visual information of rich content, effective and efficient manipulations of large video databases are increasingly desired. Many investigations have been made on content-based video retrieval. However, despite the importance, video subsequence identification, which is to find the similar content to a short query clip from a long video sequence, has not been well addressed. This paper presents a graph transformation and matching approach to this problem, with extension to identify the occurrence of potentially different ordering or length due to content editing.

The mapping relationship between the query and database video is first represented by a bipartite graph. The densely matched parts along the long sequence are then extracted, followed by a filter-and-refine search strategy to prune some irrelevant subsequences. During the filtering stage, Maximum Size Matching is deployed for each subgraph constructed by the query and candidate subsequence to obtain a smaller set of candidates. During the refinement stage, Sub-Maximum Similarity Matching is to identify the subsequence with the highest aggregate score from all candidates, according to a robust video similarity model

Formulation of the problem and analysis

At present, it is often undesirable to manually check whether a video is part of a long stream by browsing its entire length; thus, a reliable solution of automatically finding similar content is imperative using content based searching.

Disadvantage

– Require more Processing Time
– Require Boundary Editing

In this paper I use query processing technique for temporal order database and find out the similar video in the stream, if there exists some transformation distortion, and insertion that would be a duplicate copies. Next Ordinal signature method is used for assigning the keys to the videos with same content and find out the duplicates with different keys but same content

Advantages
- Avoid Duplicates
- Temporal Order Matching
- Fast Retrieval
- More secure

II. ARCHITECTURAL DESIGN

Architecture diagram shows the relationship between different components of system. This diagram is very important to understand the overall concept of system.
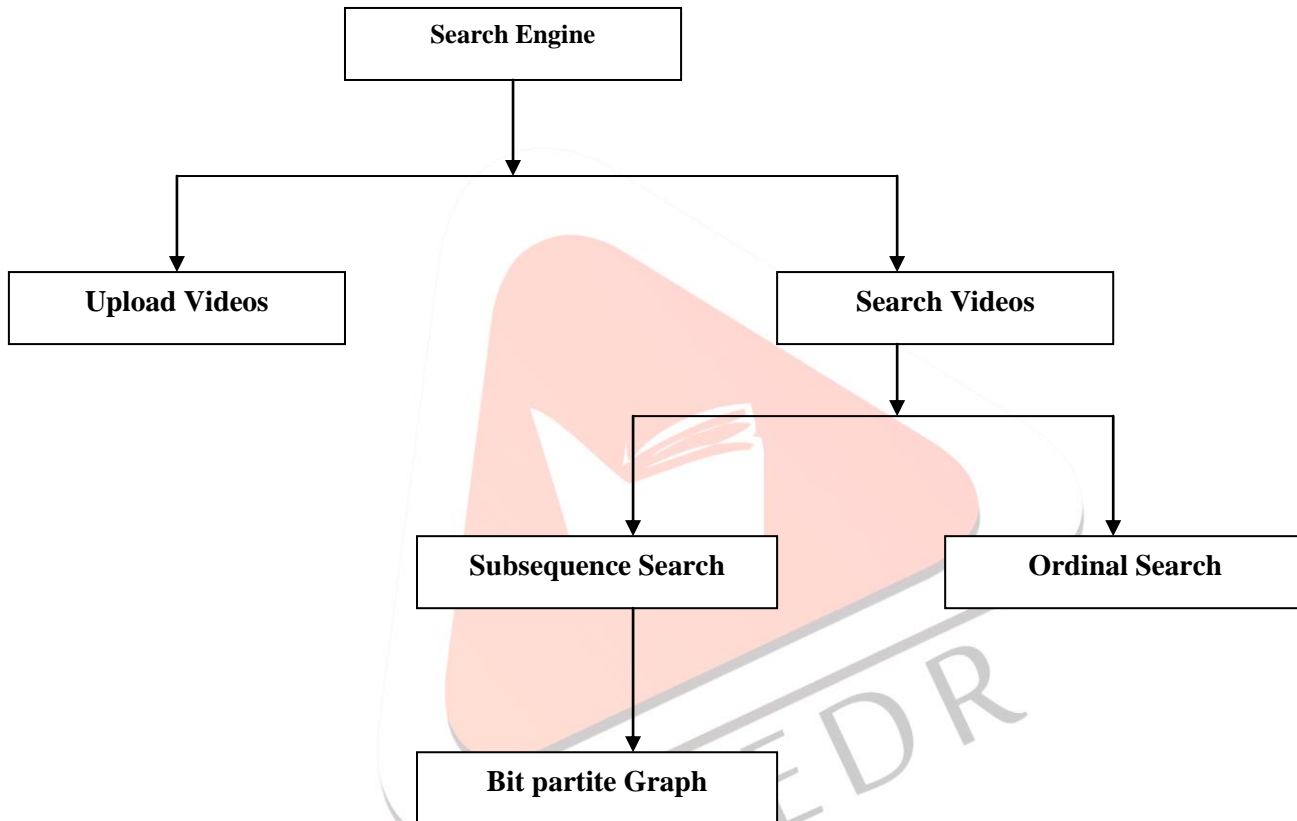

Fig 1

Most computer systems are developed to satisfy a known user requirement. This means that the first event in the life cycle of a System is usually the task of studying whether it is feasible to a Computerize a system under consideration or not. Once the decision is made, a report is forwarded and is known as Feasibility Report.

System Design

The Design Phase is the next phase of the development life cycle of a system and it involves the actual creation and design of a system. This involves putting together the different pieces that will create the system.
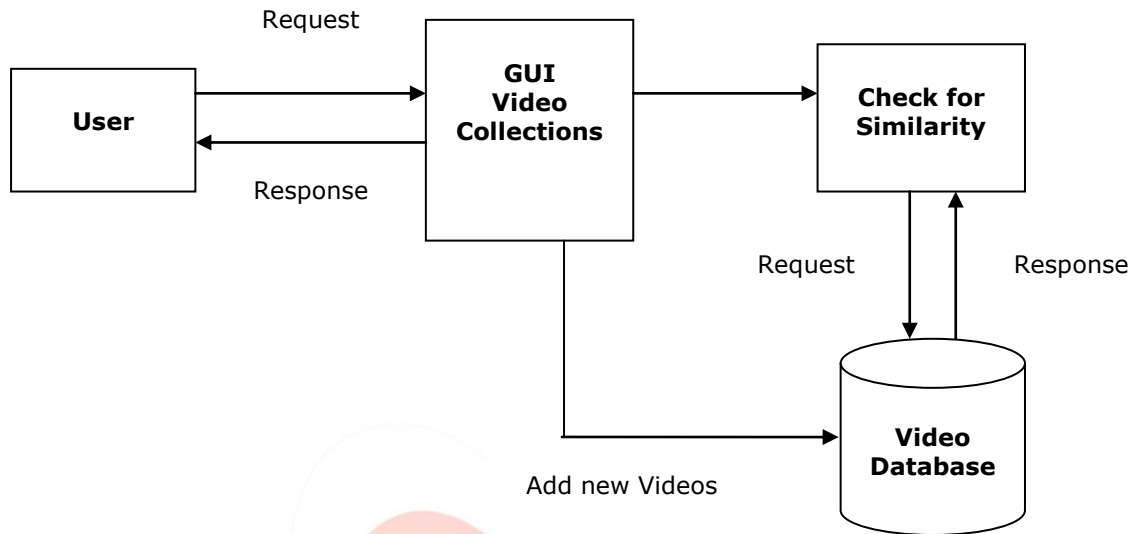
Data Flow Diagram

Fig 2

III. ALGORITHM (K-NN)

k-nearest neighbors algorithm (k-NN) is a method for classifying objects based on closest training examples in the feature space. k-NN is a type of instance based learning where the function is only approximated locally and all computation is deferred until classification. The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples.

In the classification phase, k is a user-defined constant, and an unlabelled vector (a query or test point) is classified by assigning the label which is most frequent among the k training samples nearest to that query point.

Modules
− Creations of search Engine
− Pattern Matching Process
− Plotting the Graph
− Ordinal Signature

1. **Creations of search Engine -** Search engine is created to upload the video files and search the similarity of the videos by query processing.GUI environment is created where the input query Q is given for pattern matching with database s.
2. **Pattern Matching Process -** In the pattern matching process the input query Q is trained with set of video inputs in the database S for the similarity search also for the temporal order.
3. **Plotting the Graph -** Statuses of each of the matching Frames are plotted in the respective graphs.
4. **Ordinal Signature -** Signature will be assigned to the single video with different prints and need to find out whether the signature is changed in any of the copy and found that as the duplicate copy.

REFERENCES

[1] R. Mohan, "Video Sequence Matching," Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP '98), pp. 3697-3700, 1998.
[2] Avideh Zakhor and Sen-ching Samson Cheung, Fellow, IEEE "Efficient Video Similarity Measurement With Video Signature" IEEE TRANSACTIONS, VOL. 13, NO. 1, JANUARY 2003.
[3] Hiroshi Murase, Kunio Kashino, Member, IEEE and Takayuki Kurozumi "A Quick Search Method for Audio and Video Signals Based on Histogra Pruning" Senior Member, IEEE, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 5, NO. 3, SEPTEMBER 2003
[4] H. Wang, A. Divakaran, A. Vetro, S.-F. Chang, and H. Sun, "Survey of Compressed-Domain Features Used in Audio-Visual Indexing and Analysis," J. Visual Comm. and Image Representation, vol. 14, no. 2, pp. 150-183, 2003.
[5] C.-Y. Chiu, C.-H. Li, H.-A. Wang, C.-S. Chen, and L.-F. Chien, "A Time Warping Based Approach for Video Copy Detection," Proc. 18th Int'l Conf. Pattern Recognition (ICPR '06), vol. 3, pp. 228-231, 2006.

[6] J. Shao, Z. Huang, H.T. Shen, X. zhou, E.-P. Lim, and Y. Li, "Batch Nearest Neighbor Search for Video Retrieval," IEEE Trans. aMultimedia, vol. 10, no. 3, pp. 409-420, 2008.

.