International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 08 | Aug-2016 www.irjet.net p-ISSN: 2395-0072

Content Based Video Retrieval

Jyoti Kashyap[1], Gaurav Punjabi[2], Kshitij Chhatwani[3]

[1]Guide, *Department of Electronics and Telecommunication, Thadomal Shahani Engineering College, Bandra, Mumbai, INDIA.*

[2,3] Student, *Department of Electronics and Telecommunication, Thadomal Shahani Engineering College, Bandra, Mumbai, INDIA.*

---***---

Abstract - *Searching for digital information on web, especially images, music, and video, is quickly gaining importance for the business and entertainment industry. Content-based video retrieval (CBVR) is a prominent research interest. The ability of a computer to automatically recognize objects in videos is so low that the existing technique for extracting semantic features from all kinds of videos is incapable of retrieving videos based on semantic feature. Here a new approach is proposed for video clip retrieval based on Earth Mover's Distance (EMD). Instead of imposing one-to-one matching constraint it allows many-to-many matching methodology and is capable of tolerating errors due to video partitioning and various video editing effects. We formulate clip-based retrieval as a graph matching problem in two stages. In the first stage, to allow the matching between a query and a long video, an online clip segmentation algorithm is employed to rapidly locate candidate clips for similarity measure. In the second stage, a weighted graph is constructed to model the similarity between two clips. EMD is proposed to compute the minimum cost of the weighted graph as the similarity between two clips.*

***Key Words*: — Content based Video Retrieval, CBVR, HSV, Hue Saturation Value, EMD, Earth Movers Distance, Histogram.**

1. INTRODUCTION

With the drastic growth of multimedia data in internet, TV stations, enterprises and personal digital archives, an effective yet efficient way of retrieving relevant multimedia information such as video clips is a highly challenging issue. Since the past decade, numerous researches have been conducted for content-based video retrieval. Nevertheless, most works are concentrated on retrieval by single shot, rather than retrieval by multiple shots (video clip). In this paper, we proposed a new approach based on Earth Mover's Distance (EMD) for similarity measure between two video clips [1].

2. SCOPE OF PROJECT

With vast increase in multimedia and technology CBVR has become popular. If user doesn't know the name of video CBVR system can be used to get the entire video with the help of a query clip. The retrieval accuracy of today's CBVR algorithms is still very limited. Apart from many other difficulties, the bottleneck is the gap between low-level features and semantic features of videos. Users prefer semantic information of videos. So the searching results seldom make users satisfied. The above disadvantages are covered in the work proposed here using hybrid method of relevance feedback and EMD algorithm [1].

3. WORKING OF THE PROJECT

The gap between low-level features and semantic features of videos is a major difficulty. Users prefer semantic information of videos. So the search results seldom satisfies user's objectives. The ability of computers to automatically recognize objects in videos is very less and hence the existing techniques are incapable of retrieving videos based on semantic feature. Nowadays, it is considered to be efficient to capture more semantic information by relevance feedback in video retrieval systems. Relevance feedback[3] originated from text-based information retrieval and is a powerful supervised active learning technique to improve the retrieval performance. In order to approach the query of a user, relevance feedback is viewed as the process of automatically altering an existing query by incorporating the relevance judgments that the user provide for the previous retrieval. In video retrieval, relevance feedback will first solicit the user's relevance judgments on the retrieved key frames returned by CBVR systems. Then, the system refines retrieval results by learning the query targets from the provided relevance information. The main idea is to make use of positive and negative examples obtained from a user to improve the system performance. For a given query, the system first retrieves a list of ranked key frames according to a predefined similarity metrics. Then, the user marks the retrieved key frames as relevant (positive examples) to the query or not (negative examples). The system will refine the query based on the feedback and retrieves a new list of key frames and presents to the user. To some extent, this can narrow the gap between low-level (visual) features and high-level (semantic) features[2]. The detailed similarity ranking is based on many-to many mapping by EMD. The major contributions of this approach are *similarity ranking.* We model two clips as a weighted graph with two vertex sets: Each vertex represents a shot and is stamped with a signature (or weight) to indicate its significance during matching. The signature symbolizes the duration of a shot. EMD is then employed to compute the minimum cost of the graph, by using the signatures to control the degree of matching under many-to-many shot mapping.

4. PROBLEM DEFINITION & SOLUTION

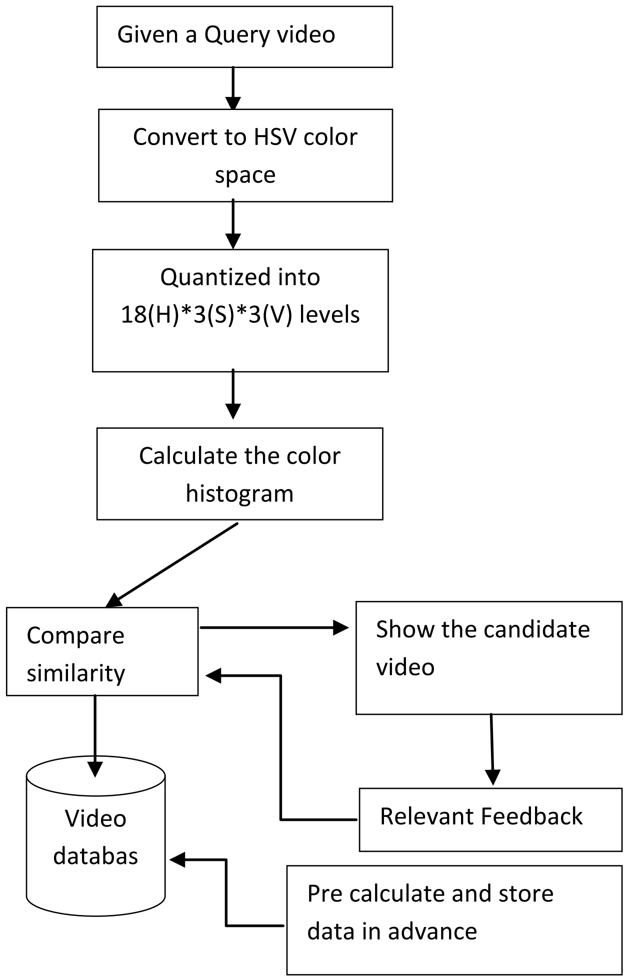
Searching for digital information on web, especially images, music, and video, is quickly gaining importance for business and entertainment industry. Content-based video retrieval (CBVR) is now becoming a prominent research interest [8].CBVR is the application of computer vision techniques to video retrieval problem, i.e. searching for digital videos in large databases. "Content-based" means that the search analyzes the contents of the video rather than the metadata such as keywords, tags, or descriptions associated with the video. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the video itself. Having humans manually annotate videos by entering keywords or metadata in a large database can be time consuming and may not capture the keywords desired to describe the video .In addition to that, manually entering the textual information about the video to be retrieved can be impractical for very large databases. Many standards have been developed to categorize videos, but still face scaling and wrong categorization issues. The evaluation of the effectiveness of video retrieved is subjective and has not been well-defined. In the same regard, CBVR systems have similar challenges in defining success .Most study on CBVR was based on extracting visual features such as color, texture, shape and motion, etc. However, despite all the research efforts, the retrieval accuracy of today's CBVR algorithms is still very limited.

Apart from many other difficulties, the bottleneck is the gap between low-level features and semantic features of videos. Users prefer semantic information of videos. So the searching results seldom make users satisfied. The ability of computers to automatically recognize objects in videos is so low that the existing technique for extracting semantic features from all kinds of videos is incapable of retrieving videos based on semantic feature.

Problem Solution

We propose to design an integrated query based system that involves providing the CBVR system with an example video over which the entire search is based. This query technique removes the difficulties that can arise when trying to describe videos with words. The given query video is converted to CIE-lab color space and HSV color space [1]. Lab color space is a color opponent space with dimensions L for lightness and a and b for the color opponent dimensions based on nonlinearly compressed CIE XYZ color space coordinates. HSV (hue-saturation-value) is one of the most common cylindrical-coordinate representations of points in an RGB color model. Then color histogram is computed which is further used in EMD algorithm .The color histogram is a statistic that can be viewed as an approximation of an underlying continuous distribution of colors values. The EMD algorithm is then used to compare similarity between the sample video and the videos from the database. The database comprises of a large collection of videos with pre-calculated and stored data in advance. This will result in the first set of the retrieved candidate videos from the database. The retrieved result will be available on to the user interface enabling them to give their feedback about the result. Since we wish to propose a system that understands the level of user satisfaction, we employ a relevance feedback mechanism [9], wherein relevance feedback is expected from the users indicating their level of satisfaction. Relevance feedback [3] is used to filter out the videos which do not match up with the user's satisfaction.

Block Diagram

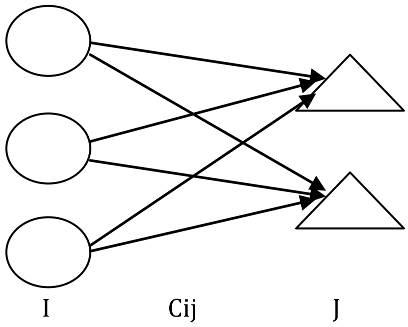


Relevance Feedback

Relevance feedback [2] is a feature of some information retrieval systems. The idea behind relevance feedback is to take the results that are initially returned from a given query and to use information about whether or not those results are relevant to perform a new query. There are three types of feedback: explicit feedback, implicit feedback, and blind or "pseudo" feedback.

EMD (Earth Movers Distance)

To overcome the drawback of traditional text-based image retrieval systems which require considerable amount of human effort, content-based image retrieval (CBIR)[7] was introduced in the early 1990s. Most current CBIR systems such as QBIC system, Photo book system, Simplicity system, etc., often use the information given in one pass at each iteration and refined query is treated just like a starting query. However, extensive experiments on CBIR systems show that in many cases low level image features cannot describe the high level semantic concepts in the users mind. In order to improve the retrieval accuracy of CBIR systems, we investigated several feature extraction algorithms, compared with the proposed application of EMD, hereby relevance feedback is applied on EMD in our approach. The main strategy of relevance feedback [9] is distance re-weighting. By introducing the users as a part of the retrieval loop, they are permitted to mark the images the system returns and find out which one is relevant or non- relevant. Then the system goes to the next retrieval procedure determined by their feedback information. The results, in general, will be considered to satisfy the user's needs through 2 or 3 times relevance feedback. A weighted graph is constructed to model the similarity between two clips, and then EMD is employed to compute the minimum cost of the weighted graph as the similarity value between two clips. EMD is based on the well-known transposition problem.



I: set of suppliers and J: set of consumers;
Cij: cost of shipping a unit of supply from i to j.

The problem is to find the optimal set of flows fij such that

Minimize $\sum_{i\in I}\sum_{i\in I} cij\, fij$

Such that

fij>=0(no reverse shipping)

$\sum_{i\in I} fij = yj$ *(Satisfy each consumer's need)*

$\sum_{j\in J} yj \leq xi$ *(Feasibility)*

EMD of Color Histogram

h=[h(1),h(2),.....,h(M)] , g=[g(1),g(2),...,g(N)]

$$\sum_{j} g(j) \leq \sum_{i} h(i)$$

Assume

EMD (h,g)=

$$\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} cij\ fij}{\sum_{i=1}^{M}\sum_{j=1}^{N} fij}$$

where cij is distance between color i in color space h and color j in color space g, fij is move fij units of mass from i in h to j in g.

The advantages of EMD are obvious that first, the cost of moving "earth" reflects the notion of nearness properly, without the quantization problems of most current measures. Moreover, the EMD support partial matching in a very natural way and it allows flexible structures.

Design Relevance Feedback Strategy Based On EMD

The EMD can be computed efficiently, especially if it is used for image retrieval systems where a quick response is required. However the relevance feedback strategy based on EMD can enhance the precision of the retrieval result by moving the query vector and updating the weighting factors simultaneously [7]. Hereby the image database is defined as IMD={Pk | k=1,...,K}, where K is the total number of the database. Generally, the users solely estimate the top number of images the system return. We also assume that the relevance sample set marked by users is Pr, while the non-relevance sample set is Pn .The sum of Pr and Pn is N0. The signature of each image by EMD has a unique factor called emdFactor, representing the relevance feedback's influence on the distance of EMD algorithm. The higher value of emdFactor, the more important it can match the query image. Apparently, emdFactor can be obtained by the user's feedback influence. If the image belongs to Pr, the emdFactor will become larger, otherwise, if it belongs to Pn, the emdFactor will be taken minus operate. There is a function controls the change of emdFactor which is based on similar strategies using in PageRank. The value of relative similarity is equal to the former one multiply emdFactor, and thus a new retrieval result will be sorted according to the new EMD distances.

6. RESULTS

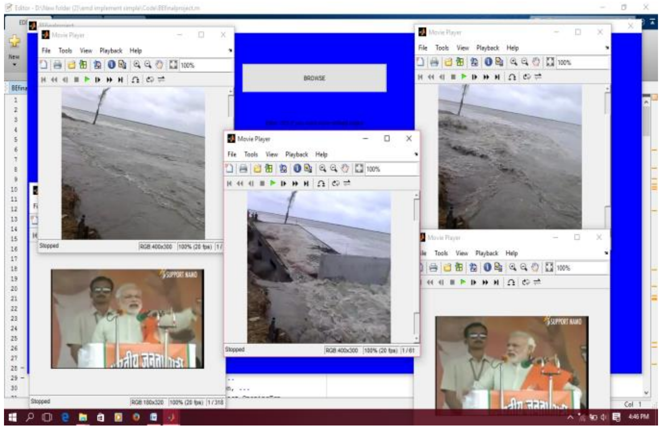


Fig-1 Histogram Output

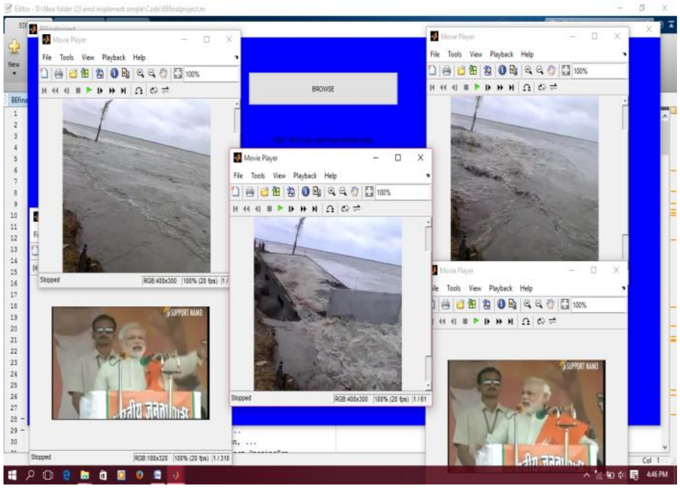


Fig-2 EMD Output

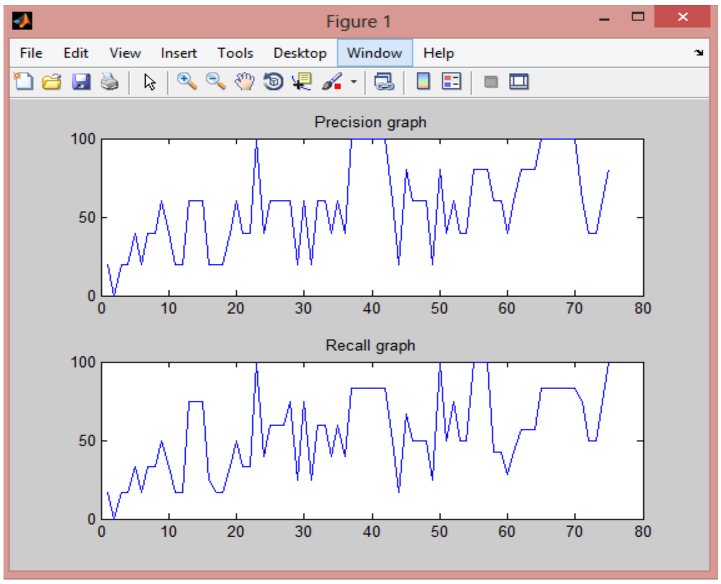


Chart-1 Precision and Recall graph for HSV Method

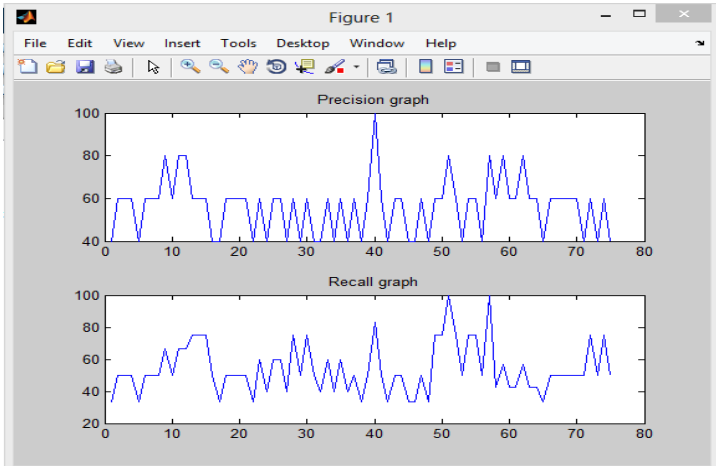


Chart-2 Precision and Recall graph for EMD Method

The time taken by HSV is 3-4 Sec and EMD is 10-11 Sec. The comparative result shows that the success rate for HSV is 80% and EMD is 100%.

7. CONCLUSION

In order to have an effective CBVR system, we studied various similarity comparison algorithms and reached to the conclusion of using EMD algorithm for our system owing to its superior performance in comparison with other algorithms. In order to have a system that meets users' satisfaction we have made an effective Relevance feedback mechanism. We have also designed an effective and user friendly interface that enables them to provide feedback with ease. The interface also enables the user to witness the upgrade of the retrieved result because of their feedback, which gives them an idea about the improving accuracy of the system. Thus, we have designed a system that not only provides best possible results but also gives fulfils the user's requirements successfully.

REFERENCES

1. Yuxin Peng and Chong-Wah Ngo ,EMD-Based Video Clip Retrieval by Many-to-Many Matching , 2005-2006.
2. Fangshi Wang, Meng Xiao, Chenlong Wang, De Xu, Semantic Video Retrieval Based on Relevance Feedback with Long-Term Memory Learning, 2006-2007.
3. Xia Wu, Dinan Fu ,Apply Hybrid Method of Relevance Feedback and EMD Algorithm in a Color Feature Extraction CBVR System, 2008.
4. Jaume Verges-Llahi and Alberto Sanfeliu, Evaluation of Distances between Color Image Segmentations, 2000.
5. Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas ,A Metric for Distribution with Applications to Image Databases,2002.
6. Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas ,The Earth Mover's Distance as a Metric for Image Retrieval,2002.
7. Simardeep kaur and Dr.Vijaykumar Banga , CBIR: Survey and Comparison between RGB and HSV model,2013.
8. Navdeep Kaur and Mandeep Singh, Content Based Video Retrieval Using Color, Texture And Time Based Activity,2014.
9. Zhong Jin, Irwin King, and Xuequn Li, Content-Based Image Retrieval by Relevance Feedback,2000. Rui Y and Huang T S, Relevance feedback technique in video retrieval,1995 .