International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 9 Issue 1 – JULY 2014.

Combined Tracking Technologies Using Data Mining With Markov- Chain Process

R.Priyanga[#1] and M.Geetha[*2]

[1]*PG Scholar, Muthayammal College of Arts&science*

[2] *Associate Professor, Muthayammal College of Arts&science*

Abstract- The growth in available online video material over the Internet is generally increasing day by day. To enhance the content based video retrieval, we propose a technique which incorporates Zero-Shots, Similarity Measure and Active Buckets. As multimedia data become ubiquitous in our daily lives, information retrieval systems have to adapt their retrieval performance to different situations in order to efficiently satisfy the users information needs anytime and anywhere. By incorporating three techniques the retrieved video will be the most relevant video for the query video. Explicit relevance feedback is used to satisfy the user if he is not satisfied with the retrieved video.

Keywords: Feature extraction, Key frames, Content based video retrieval, Similarity search, MSI.

I. INTRODUCTION

Information retrieval systems are ubiquitous in our daily lives. They support us to store, manage and access voluminous information based on the systems' underlying data. Thus, the primary aim of such information retrieval systems is to supply user's with relevant and useful information in order to satisfy their information needs. To this end, the system should perform the information retrieval process of a user-specific query in an effective and also efficient way.

Prominent examples of information retrieval systems are search engines in the World Wide Web. They enable us to do a search for interesting web pages, popular video clips, famous research papers, gorgeous images, and so on. Almost all multimedia information inside and outside the World Wide Web are made accessible via appropriate retrieval systems. In order to fulfill retrieval and browsing tasks in the user's sense, information retrieval systems use different kinds of models to represent their accessible data through additional semantic and syntactic information. This information reflects the contents of the stored data and represent the core of each information retrieval process.

To perform the content-based information retrieval process in an efficient way, we propose a technique which first converts the given input video into frames and by using KeyFrame low and high level features are extracted. By using combinational based similarity features relevant videos are clustered. Based upon the metadata availability and visually similar criteria relevant videos are stored in the bucket upon the level of similarity. Finally most relevant video is retrieved by the user.

Video Segmentation:

Usually, a video is created by taking a set of shots and composing them together using specified composition operators. A shot is usually a piece of video taken with a fixed set of cameras, each of which has a constant relative velocity. In general, a shot may have many associated attributes such as the duration of the shot, the type(s) of camera(s) used, and so on.

A shot composition operator is an operation that takes two shot, S1 and S2, and a duration t as input and merges the two shots into a composite shot within time t. Thus, for example, suppose we wish to compose together two shots S1 and S2, and suppose these two shots have duration t1 and t2 respectively. If f is a shot composition operator, then f(S1, S2, t) creates a segment of video of length (t1+t2+t). S1 is first shown and then undergoes a continuous transformationover time interval t, leading to the presentation of S2 next. f(S1, S2, t) then is a continuous sequence of video. In general, a video as a whole may be represented as

fn(…f2(f1(S1,S2,t),S3,t2)…,Sn+1,tn)

Video segmentation, involves identifying the frame(s) where a transition takes place from one shot to another. In cases where this change occurs between two frames, it is called a cut or a break. Converting given input into scenes, scenes into shots and shots into frames. Frame is a single picture or still shot, that shown as a part of larger video or movie. Based upon the length of the input video, number of frames varied. Normally 24 frames per second.

Feature Extraction:

To extract features according to video structural analysis results is the base of video indexing and retrieval. We focus on the visual features suitable for video indexing and retrieval. These mainly include features of key frames, objects, and motions.

The key frames of a video reflect the characteristics of the video to some extent. Traditional image retrieval techniques can be applied to key frames to achieve video retrieval. The static key frame features useful for video indexing and retrieval are mainly classified as color-based, texture-based, and shape-based.

1) Color-Based Features:

Color-based features include color histograms, color moments, color correlograms, a mixture of Gaussian models, etc. The exaction of color-based features depends on color spaces such as RGB, HSV, YCbCr and normalized r-g, YUV, and HVC. The choice of color space depends on the applications. Color features can be extracted from the entire image or from image blocks into which the entire image is partitioned. Color-based features are the most effective image features for video indexing and retrieval. In particular, color histogram and color moments are simple but efficient descriptors.

The merits of color-based features are that they reflect human visual perception, they are easy to extract, and their extraction has low computational complexity. The limitation of color-based features is that they do not directly describe texture, shape, etc., and are, thus, ineffective for the applications in which texture or shape is important.

2) Texture-Based Features:

Texture-based features are object surface-owned intrinsic visual features that are independent of color or intensity and reflect homogenous phenomena in images. They contain crucial information about the organization of object surfaces, as well as their correlations with the surrounding environment. Texture features in common use include Tamura features, simultaneous autoregressive models, orientation features, wavelet transformation-based texture features, co-occurrence matrices, etc.

The merit of texture-based features is that they can be effectively applied to applications in which texture information is salient in videos. However, these features are unavailable in non texture video images.

3) Shape-Based Features:

Shape-based features that describe object shapes in the image can be extracted from object contours or regions. A common approach is to detect edges in images and then describe the distribution of the edges using a histogram, edge histogram descriptor (EHD) to capture the spatial distribution of edges for the video search task. Shape-based features are effective for applications in which shape information is salient in videos. However, they are much more difficult to extract than color- or texture-based features.

B. Object Features

Object features include the dominant color, texture, size, etc., of the image regions corresponding to the objects. These features can be used to retrieve videos likely to contain similar Objects. Faces are useful objects in many video retrieval systems. The limitation of object-based features is that identification of objects in videos is difficult and time-consuming. Current algorithms focus on identifying specific types of objects, such as faces, rather than various objects in various scenes.

C. Motion Features

Motion is the essential characteristic distinguishing dynamic videos from still images. Motion information represents the visual content with temporal variation. Motion features are closer to semantic concepts than static key frame features and object features. Video motion includes background motion caused by camera motion and foreground motion caused by moving objects. Thus, motion-based features for video retrieval can be divided into two categories: camera-based and object-based. For camera-based features, different camera motions, such as “zooming in or out,” “panning left or right,” and “tilting up or down,” are estimated and used for video indexing. Video retrieval using only camera-based features has the limitation that they cannot describe motions of key objects. Object-based motion features have attracted much more interest in recent work.

II. SIMILARITY MEASURE

Video similarity measures play an important role in content based video retrieval. Methods to measure video similarities can be classified into feature matching, text matching, ontology based matching, and combination-based matching. We are using combination based matching which constitutes both feature and ontology based matching.

Algorithm for Hybrid Cluster
Input: Set of web pages from search engine Ws
Output: Set of key terms Ts
Step1: Read all web pages given for training Ws

Step2: Read stop word list Sw
Step3: For each web page Wi from Ws
C = Read content of the web page Wi
C = Apply html parser to remove html tags from C
Ts = Split C with pattern single space
For each term Ti in the term set Ts
If Ti present in stop word list Sw then
Remove Ti from Ts
End
If (Ti contains "ing")
Ti = Remove "ing" from Ti
End
If (Ti contains "ed")
Ti = Remove "ed" from Ti
End
End
End
Step4: For each Ti from Ts
Identify presence of bigram Bi
If Bi Presents
Update Ts
Else
Continue
End
Step5: Return set of textual term set
Step6: Stop
Combination-Based Matching:

This approach "leverages semantic concepts by learning the combination strategies from a training collection, It is useful for combination-based queries that are adaptable to multimodal searches. The merits of the combination-based matching approach are that concept weights can be automatically determined and hidden semantic concepts can be handled to some extent. The limitation of this approach is that it is difficult to learn query combination models.

1.Feature Matching:

The most direct measure of similarity between two videos is the average distance between the features of the corresponding frames. Query by example usually uses low-level feature matching to find relevant videos. However, video similarity can be considered in different levels of resolution or granularity. According to different user' demands, static features of key frames, object features, and motion features all can be used to measure video similarity.

The merit of feature matching is that the video similarity can be conveniently measured in the feature space. Its limitation is that semantic similarity cannot be represented because of the gap between sets of feature vectors and the semantic categories familiar to people.

2) Text Matching:

Matching the name of each concept with query terms is the simplest way of finding the videos that satisfy the query. Normalize both the descriptions of concepts and the query text and then ompute the similarity between the query text and the text descriptions of concepts by using a vector space model. Finally, the concepts with the highest similarity are selected.

The merits of the text-matching approach are its intuitiveness and simplicity of implementation. The limitation of this approach is that all related concepts must be explicitly included in the query text in order to obtain satisfactory search results.

3) Ontology-Based Matching:

This approach achieves similarity matching using the ontology between semantic concepts or semantic relations between keywords. Query descriptions are enriched from knowledge sources, such as ontology of concepts or keywords. The ontology is used to determine which concepts are mostly related to the original query text.

Based on the fact that the semantic word similarity is a good approximation for visual co-occurrence. Utilize semantic word similarity measures to measure the similarity between text annotated videos and users' queries. Videos are retrieved based on their relevance to a user-defined text query. The merit of the ontology-based matching approach is that extra concepts from knowledge sources are used to improve retrieval results.

Active Buckets:

There is a semantic concept available as metadata which is directly related to the search query we have a good starting point. Depending on whether the concept relates to data that shares visually similar characteristics or not defines whether from this initial starting point the search is easy (one cluster in feature space) or difficult (various clusters in feature space). When there is no directly related metadata we can still find results indirectly if the data is visually similar, but it is very difficult otherwise. This brings us to the four types of retrieval conditions:

A metadata available, visually similar
B no direct metadata, visually similar
C metadata available, visually diverse
D no direct metadata, visually diverse

The basis for our categorization approach is formed by MediaTable. MediaTable provides different views on a collection, with a spreadsheet-like table interface, showing the entire collection. This table shows per row a shot of video material, together with a preview image, and all associated metadata. On top of this, several visualization techniques enhance the effectiveness of MediaTable for video categorization and video retrieval. Most relevant video is retrieved by the user.

Explicit Relevance Feedback:

This feedback asks the user to actively select relevant videos from the previously retrieved videos. The user can label sample videos as "highly relevant," "relevant," "no-opinion," "nonrelevant," or "highly nonrelevant,"a relevance feedback response technique that can adjust the weights of different features and different spatial locations in key frames according to the user's feedback.

The user can directly select the features important for the user and the image sections that the user wants to search for. The merit of explicit feedback is that it can obtain better results than implicit feedback or the pseudo feedback discussed later as it uses the user feedback directly. Its limitation is that it needs more user interaction, which requires more user patience and cooperation. If the user is not satisfied with the retrieved video, use the explicit relevance feedback user may actively select the video from the buckets.

III. CONCLUSION AND FUTURE WORK

The videos are arranged in the buckets based upon the level of their similarity. The most relevant video is retrieved from the bucket. If the user is not satisfied with the retrieved video, user can actively select the relevant video from the buckets. Thus the user can retrieve the video with high accuracy.

By additionally adding low level features, we can retrieve the audio files related to the search query. Then it will be the efficient method of audio and video retrieval technique.

IV. REFERENCES

[1] Arthur Pope, Rakesh Kumar, Harpeet Sawhney and Charles Wan "Video Abstraction: Summarizing Video Content for Retrieval and Visualization" 1998.

[2] Milind Ramesh Naphade and Thomas S.Huang "A Probabiliistic Framework for Semantic Video Indexing, Filtering and Retrieval" Vol.3, NO 1, March 2001.

[3] Manolis Wallace, Thonas Athanasiasdis and Yannis Avrithis "Knowledge Assisted Analysis and Categorization for Semantic Video Retrieval" 2004.

[4] Dar-Shyang Lee, Joanthan J.Hull,Berna Erol " Meeting Video Retrieval using Dynamic HMM Model Similarity" 2005.

[5] Sergio Benini, Pierangelo Miglorati and Ricardo Leonardi "Retrieval of Video Story Units by Markov Entropy Rate" 2008.

[6] Steven C.H.Hoi, Michael E.Lyu "A Multimodal and Multilevel Ranking Scheme for Large –Scale Video Retrieval"2008.

[7] Alan F.Smeaton, Peter Wilkins, Marcel Worring, Ork de Rooji, Tat-Seng Chua, Huanbo Luan"Content Based Video Retrieval: Three Example Systems from TRECVid" 2008.

[8] Lili N.A. "Hidden Markov Model for Content-Based Video Retrieval" 2009.

[9] Lifeng Shang, Linjun Yang, Fei Wang, Kwok-Ping Chan, Xian-Sheng Hua "Realtime Large Scale Near-Duplicate Web Video Retrieval" 2010.

[10] Weiming Hu, Nianhua Xie , Li Li, Xianglin Zeng and Stephen Maybank " A Survey on Visual Content Based Video Indexing and Retrieval" Vol.41,No.6, November 2011.

[11] Barbara Andre, Tom Vercauteren, Anna M. Buchner, Michael B. Wallace, Nicholas Ayache "Learning Semantic and Visual Similarity for Endomicroscopy Video Retrieval" August 2011.

[12] S. Padmakala and G.G. AnandhaMala "A Technique to Content-Based Video Retrieval Utilizing Diverse Set of Features"2012.

[13] Ms.Kainjan S.Kotecha , Mr.Mahesh R.sanghavi Dr Mrs.A.M.Rajurkar " Content Based Video Retrieval using Ranking Correlation, Motion and Color" 2012.

[14] R.Kanagavalli and K.Duraisamy "Shot Detection using Genetic Edge Histogram and Object Based Video Retrieval using Multiple Features" 2012.

[15] Xu Chen, Hero. A.O, Savarese. S "Multimodal Video Indexing and Retrieval using Directed Information" Volume :14, Feb 2012.

Author's Details:

M.Geetha, received her B.Sc degree from university of Bharathidasan and M.Sc(IT) degree from Bharathidasan university. She has completed her M.Phil at Annamalai university. She is having 8 years of experience, 6 years in Muthayammal College of arts & Science, Rasipuram affiliated by Periyar University. Already had 3 in years experienced in PGP College of arts & science, Namakkal affiliated by periyar university. Her main research interests include algorithm and data mining & ware housing.

R.Priyanga, received her B.Sc., degree in Thiruvalluvar government Arts college from periyar university,salem(2008-2011)Tamilnadu(India).Then finished M.Sc., degree in vivekanandha college of Arts and sciences for women from periyar university ,salem(2011-2013)Tamilnadu(India). She is the M.Phil Research Scholar of Muthayammal College of Arts and Science ,Rasipuram. Periyar university, salem. Her Area of interest is Data mining.