
Video Retrieval using Wavelet and Clustering

S B Singh
B.E (Electrical) M.E (Electronics)
I/C HOD of Electronics Dept
Shah and Anchor Kutchhi Engg. College,
Chembur, MUMBAI-88
hodex@sakp.ac.in

Abstract: A Content-base Video retrieval, which provides convenient ways to retrieve Video from large Video database. While many previous Video retrieval techniques do not look at regions in an Video, region based Video retrieval techniques have been gaining attention recently. Traditional method for retrieval has failed to retrieve multimedia data like Video, audio, video etc. CBIR is very active research area which is used retrieves Videos based on Video visual features. The present method uses a new technique based on wavelet transformations to create feature vector of Video, characterizing texture feature of the Video is constructed. For quick retrieval statically detail of Video like mean of approximation, standard deviation and second moment is utilized for shape in Video. Our method derives feature vector for each Video characterizing the texture feature of sub Video from only three iterations of wavelet transforms. The K Means Clustering Algorithm is used to cluster the group of Videos based on feature vector of Videos by considering the minimum Euclidean distance. Our experiments are performed on texture Videos, and successful matching results are found.Our experiments are performed on texture Videos; Geographical Video, facial Video and medical Videoare finding useful matching results.

Keywords: *Wavelet Transform and Video database K Means Clustering, semantic gap*.

*****

I. INTRODUCTION:

Content-based Video retrieval is a part of Video processing and is also comes under artificial intelligence. We know that interest in digital Videos is growing day by day; Users in many professional fields are exploiting the opportunities offered by the ability to access and manipulate remotely-stored Videos in all kinds of new and exciting ways. The problems in Video retrieval is becoming widely recognized, and the search for solutions an increasingly active area for research and development. So here is the technique Content-based Video retrieval (CBIR), also known as query by Video content (QBIC) and content based visual information retrieval (CBVIR) is the application of computer vision to the Video retrieval problem, that is, the problem of searching for digital Videos in large databases.

The use of digital Videos and pictures in communication is new – what olden days cave-dwelling ancestors did was they use to paint pictures on the walls of their caves, and the use of maps and building plans to convey information almost certainly dates back to pre-Roman times. But the twentieth century has witnessed unparalleled growth in the number, availability and importance of Videos in daily life. They now play an important role in fields as diverse as medicine,
The Content-based Video and video retrieval is a fast growing and increasing relevant research area. The research community recognizes the following main challenges in this field: the bridging of the semantic gap (understanding the meaning behind the query), the content-based retrieval of Video (finding a video similar to another one), and the increasing huge amount of digital data, produced by digital consumer devices (e.g. digital cameras) and computational devices (hard disks, CD-ROMs, etc.), This situation should be improved and we think we should start to develop methods and strategies for the content-based retrieval of information on the web, the largest and most used multimedia database in the world.

II. NEED FOR CBIR

In today's world more and more multimedia Information is stored in databases like audio and video. An Video may be better or more effective than a substantial amount of text. It also aptly characterizes the goals of visualization where large amounts of data must be absorbed quickly and it is a picture is worth of thousand words. If we use text for retrieving/searching an Video what happens is-we can't write whole description of Video this is one thing and if the Video contains geographical data then manual keyword descriptionare not possible ex: Google earth contains geographical data. So then why to go for this is because- It is cultural language dependent and is not possible to describe every Video in database. Keyword matching will not give most relevant Videos.

III. TECHNIQUES USED IN CBIR:

CBIR operates on a totally different principle, retrieving/searching stored Videos from a collection by comparing features automatically extracted from the Videos themselves. To retrieve videos from database efficient video indexing and retrieval mechanisms are required The commonest features used are mathematical measures of color, texture or shape (basic). A system (CBIR) allows users to formulate queries by submitting an example of the type of Video being sought (input), though some offer alternatives such as selection from a palette or sketch input we can also select color textures or any other visual information.

IV. QUERY TECHNIQUES

4.1Semantic retrieval

The ideal CBIR system from a user perspective would involve what is referred to as *semantic* retrieval, where the user makes a request like "find pictures of Abraham Lincoln". This type of open-ended task is very difficult for computers to perform pictures of Chihuahuas and Great Danes look very different, and Lincoln may not always be facing the camera or in the same pose. Current CBIR systems therefore generally make use of lower-level features like texture,colour, and shape although some systems take advantage of verycommon higher-level features like faces (see facial recognition system). Not every CBIR system is generic. Some systems are designed for a specific domain, e.g. shape matching can be used for finding parts inside a CAD-CAM database.

4.2 Other Query Methods

Other query methods include browsing for example Videos, navigating customized/hierarchical categories, querying by Video region (rather than the entire Video), querying by multiple example Videos, querying by visual sketch, querying by direct specification of Video features, and multimodal queries (e.g. combining touch, voice, etc.)

CBIR systems can also make use of relevance feedback, where the user progressively refines the search results by marking Videos in the results as "relevant", "not relevant", or "neutral" to the search query, then repeating the search with the new information.

V. CONTENT COMPARISION USING VIDEO DISTANCE MEASURES:

The most common method for comparing two Videos in content based Video retrieval (typically an example Video and an Video from the database) is using an Video distance measure. An Video distance measure compares the similarity of two Videos in various dimensions such as color, texture, shape, and others. For example a distance of 0 signifies an exact match with the query, with respect to the dimensions that were considered. As one may intuitively gather, a value greater than 0 indicates various degrees of similarities between the Videos. Search results then can be sorted based on their distance to the queried Video.

VI. FEATURE EXTRACTION AND SIMILARITY CRITERIA

Our *CBIR* algorithm is based on direct wavelet Decomposition of Video in *RGB* color space and utilizes the “query by example” method. With approaches mentioned above, database Videos are decomposed offline into multi-level coefficients from -1 to -J levels, with which, we can generate color feature database and perform similarity match between Videos. After decomposition, each resulting sub Video is in fact a coefficient matrix, where, by special processing, large coefficients with more energy can be distributed in the up left area, therefore, we can well decrease the dimension of Video feature and perform highly efficient Video matching therefore sample result are shown below

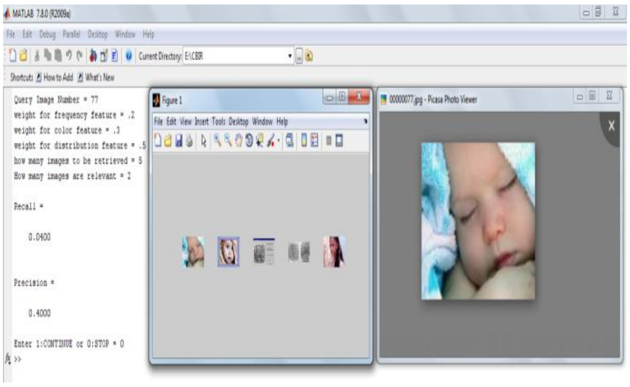

Figure 1.Software Snapshot 1

6.1 Texture Retrieval

What if the Videos are of same color? This will be answered by textures. The ability to retrieve Videos on the basis of texture similarity may not seem very useful-But the ability to match on texture similarity.

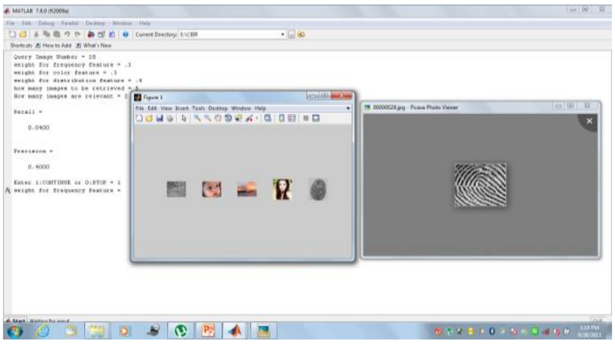

Figure 2. Software Snapshot 2

It can often be useful in distinguishing between areas of Videos with similar color (such as sky and sea, or leaves and grass). A variety of techniques has been used for measuring texture similarity; the best-established rely on comparing values of what are known as *second-order statistics* calculated from query and stored Videos. Essentially, these calculate the relative brightness of selected *pairs* of pixels from each Video. From these it is possible to calculate measures of Video texture such as the degree of *contrast*, *coarseness*, hom*ogeneity* and *regularity* [Tamura et al, 1978], or *periodicity*, *correlation and entropy* [Liu and Picard, 1996]. Alternative methods of texture analysis for retrieval include the use of Gabor filters this is the widely used technique now a days. The system then retrieves Videos with texture measures similar in value. A recent technique is the texture thesaurus developed by Ma and Manjunath [1998], which retrieves textured regions in Videos on the basis of similarity to automatically-derived code word's representing important classes of texture within the collection.

Contrast is the dissimilarity or difference between things (colour, brightness etc.). Homogeneity means "being similar throughout"(like same colour can be said to one part segmentation can also be done through this). Entropy is a measure of the uncertainty associated with a random variable.

6.2 Shape Retrieval

Shape is the most obvious requirement at the primitive level. Unlike texture, shape is a fairly well-defined concept – and there is considerable evidence that natural objects are primarily recognized by their shape .A number of features characteristic of object shape (but independent of size or orientation) are computed for every object identified within each stored Video. Queries are then answered by computing the same set of features for the query Video, and retrieving those stored Videos whose features most closely matchthose of the query. Two main types of shape feature are commonly used – *global* features such as aspect ratio, circularity and moment invariants [Niblack et al, 1993] and *local* features such as sets of consecutive boundary segments [Mehrotra and Gary, 1995]. Alternative methods proposed for shape matching have included elastic deformation of templates (Pentlands et al [1996], del Bimbo et al [1996]), comparison of directional histograms of edges extracted from the Video and *shocks*, skeletal representations of object shape that can be compared using graph matching techniques. Retrieval techniques based on compressed Video data. An individual feature will not be able to describe an Video adequately. For example, it is not possible to distinguish a red car from a red apple based on color alone. Therefore, a combination of features is required for effective Video indexing and retrieval.

6.3 Video Indexing And Retrieval

This section briefly describes seven main approaches, which are based on structured attributes, object-recognition, text, and low-level Video features. The first approach, attribute-based, uses traditional database management systems for Video indexing and retrieval. The second approach color and third based on shape are not mature yet as it relies on automatic Object recognition. Text-based Video retrieval uses traditional IR for Video indexing and retrieval. The low-level content-based approach to Video indexing and retrieval requires the extraction of low level Video features. The common features used are color, object shape, and texture. As most Videos are stored in compressed form, it would be advantageous if we could derive Video features directly from compressed Video data. Table looks into a number of Video indexing and

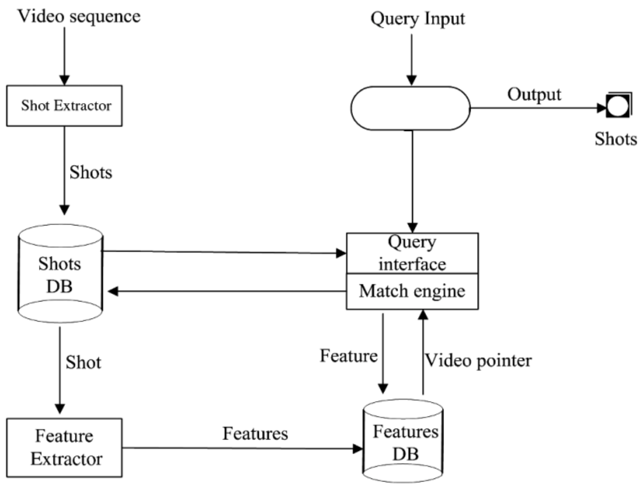

Figure 3. Sequence of Execution

Representation of a function consists of a coarse overallapproximation together with detail coefficients that influence the function at various scaled.

Thewavelet transform has excellent energy compaction and de-correlation properties, which can be used to effectively generate compact representations that exploit the structure of data. By using wavelet subband decomposition, and storing only the most important sub bands (that is, the top coefficients), we can compute fixed-size low-dimensional feature vectors independent of resolution, Video size and dithering effects. In addition, wavelets are robust with respect to colour intensity shifts, and can captureboth texture and shape information efficiently.

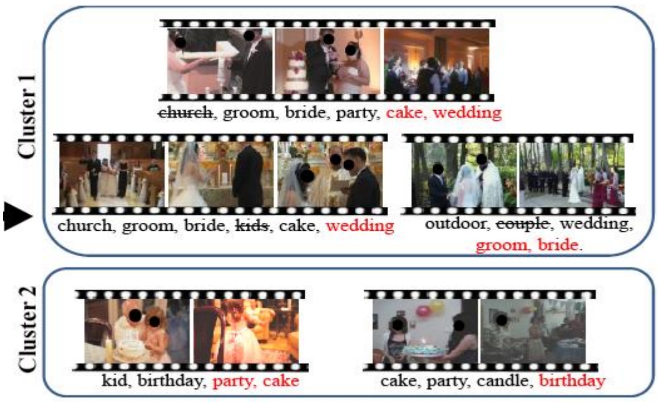

Figure 4. Video clustering

In each iteration Ii2...4 Videos are saved and Ii1 sub Video is again subjected to wavelet Transformation instead of entire Video for three iterations, by which 10 sub Videos of input Video are obtained. Sub Video I11 is further divided into sub Videos I21 ... I24 in the second iteration. The sub Video I21 is further divided into I31 I32 I33 I34 in the third iteration. All sub Videos are normalized to maintain the uniform size.

VII. REAL-WORLD REQUIREMENTS

Building real-world systems involve regular user feedback during the development process, as required in any other software development life cycle. Not many Video retrieval systems are deployed for public usage, save for Google Videos or Yahoo! Videos (which are based primarily on surrounding meta-data rather than content). There are, however, a number of propositions for real-world implementation. For brevity of space we are unable to discuss them in details, but it is interesting to note that CBIR has been applied to fields as diverse as Botany, Astronomy, Mineralogy, and remote sensing with so much interest in the field at the moment, there is a good chance that CBIR based real-world systems will diversify and expand further. We have implemented an IRM-based publicly available similarity search tool on an on-line database of over 800, 000 airline-related Videos. Another on-going project is the integration of similarity search functionality to a large collection of art and cultural Videos. Screen-shots can be see. Based on our experience with implementing CBIR systems on real world data for public usage, we list here some of the issues that we found to be critical for real-world deployment. **Performance:** The most critical issue is the quality of retrieval and how relevant it is to the domain-specific user community. Most of the current effort is concentrated on improving performance in terms of their precision and recall.

Semantic learning: To tackle the problem of semantic gap faced by CBIR, learning Video semantics from training data and developing retrieval mechanisms to efficiently leverage semantic estimation are important directions.

Volume of Data: Public Video databases tend to grow into unwieldy proportions. The software system must be able to efficiently handle indexing and retrieval at such scale.

Heterogeneity: If the Videos originate from diverse sources, parameters such as quality, resolution and color depth are likely to vary. This in turn causes variations in color and texture features extracted. The systems can be made more robust by suitably tackling these variations.

Concurrent Usage: In on-line Video retrieval systems, it is likely to have multiple concurrent users. While most systems have high resource requirements for feature extraction, indexing etc., they must be efficiently designed so as not to exhaust the host server resources. Alternatively, a large amount of resources must be allocated.

Multi-modal features: The presence of reliable metadata such as audio or text captions associated with the Videos can help understand the Video content better, and hence leverage the retrieval performance. On the other hand, ambiguous captions such as "wood" may actually add to the confusion, in which case the multi-modal features together may be able to resolve the ambiguity.

User-interface: As discussed before, a greater effort is needed to design intuitive interfaces for Video retrieval such that people are actually able to use the tool to their benefit.

Operating Speed: Time is critical in on-line systems as the response time needs to be low for good interactivity. Implementation should ideally be done using efficient algorithms, especially for large databases. For computationally complex tasks, off-line processing and caching the results in parts is one possible way out.

System Evaluation: Like any other software system, Video retrieval systems are also required to be evaluated to test the feasibility of investing in a new version or a different product. The design of a CBIR benchmark requires careful design in order to capture the inherent subjectivity in Video retrieval. One such proposal can be found in.

VIII. EXPERIMENTAL RESULTS

The general flow of the experiments starts with the decomposition of data base Video using Haar wavelet and with the calculation of moment, standard deviation, and second moment for creating featurevector in offline. We focus on comparison of all three points. The text Video content 200 Videos of 5 Bit

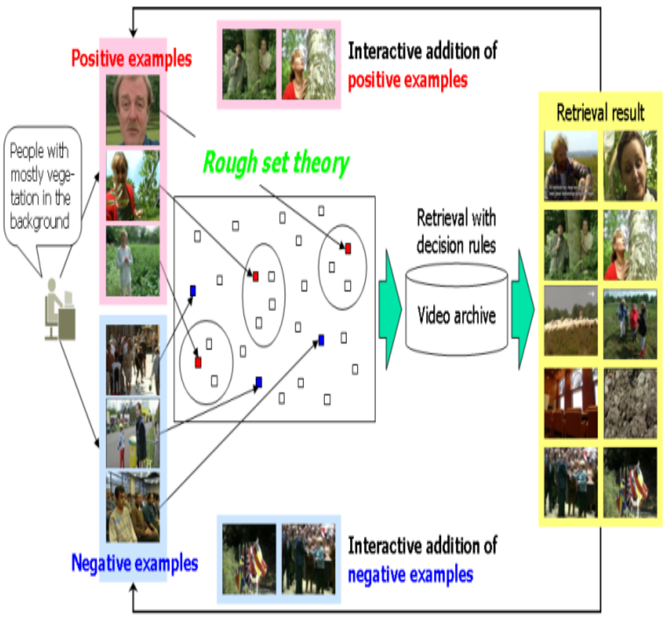

Figure 5. Haar Wavelet Transform

The above result is find using Haar wavelet Transform and existing wavelet histogram for each category of color Video. This is also the most apparent advantage of the wavelets in the real time application of retrieval cue on its own.

Figure 6. Sample data base video

Simple retrieved Videos for the given query Video using Haar and wavelets are shown in fig. the number at the foot of each Video indicates its similarity to the example Video.

Figure 7. Retrieved Sample

IX. CONCLUSION

We have presented a brief survey on work related to the young and exciting fields of content-based Video retrieval (CBIR) and automated Video annotation, spanning more than 100 publications in the current decade. We believe that the field

Will experience a paradigm shift in the foreseeable future, with the focus being more on application-oriented, domain-specific work, generating considerable impact in day-to-day life. These aspectsshould also be considered equally important. Meanwhile, the quest for robust and reliable Video understanding technology needs to continue as well. The future of this field depends on the collective focus and overall progress in each aspect of Video retrieval, and how much the ordinary individual stands to benefit from it.

CBIR at present is still very much a research topic. The technology is exciting but immature, and few operationalVideo archives have yet shown any serious interest in adoption .CBIR technology replacing more traditional methods of indexing and searching. Use of text and Video features might well yield better performance than either type.

X. ACKNOWLEDGMENTS

The future of this field depends on the collective focus and overall progress in each aspect of Video retrieval, and how much the ordinary individual stands to benefit from it. The trends indicate that while systems, feature extraction, and relevance feedback have received a lot of attention, application-oriented aspects such as interface, visualization, scalability, and evaluation have traditionally received lesser consideration. We feel that for all practical purposes, these aspects should also be considered equally important.

REFERENCES

[1] Mrs.Y. M. Latha and Dr.B.C.Jinaga ” Content Based Color Video R-etrieval via Wavelet Transforms ”
[2] Pasnur M.A and P. S. Malge,” Video Retrieval Using Modified Haar Wavelet transform and K Means Clustering”
[3] Prof. Bhavesh V.Patel & Prof. B. B. Meshram,”*Techniques for Text, Video, Audio and Video Indexing and Retrieval”*
[4] Punpiti Piamsa-NGA, Nikitas A. Alexandridis “MULTI-FEATURE ONTENT-BASED VIDEO RETRIEVAL”
[5] V.S. Subrahmanian, Principles of Multimedia Database Systems, 2001.
[6] Waleed Farag and Sanwar Ali. “A Survey of Current Challenges And Future Directions For Multimedia Information Retrieval Systems.”
[7] Menzo Windhouwer and Roelof van Zwol. “Combining Concept- With Content-Based Multimedia Retrieval.”
[8] Linda Null, Julia Lobur, The Essentials of Computer Organization and Architecture.
[9] Willam Stallings, Computer Architecture and Organisation: Designing and Performance.
[10] Lina Sun; Yihua Zhou, "A key frame extraction method based on mutual information and image entropy," Multimedia Technology (ICMT), 2011 International Conference on , vol., no., pp.35,38, 26-28 July 2011
[11] Farouk, H.; Elsalamony, H. A., "Digital library creation based on wavelet coefficients for video stream indexing and retrieving," Signal Processing Systems (ICSPS), 2010 2nd International Conference on , vol.1, no., pp.V1-158,V1-162, 5-7 July 20