Content based Video Querying Technique for Video Retrieval and Video Making from Large Video Compilation

N. Sudha Bhuvaneswari Ph.D.
Associate Professor, School of IT &Science,
Dr. G.R Damodaran College of Science.

M.Madhanika
Research scholar, School of IT &Science
Dr. G.R Damodaran College of Science

ABSTRACT
Content based video querying and video matching systems are popular in the recent technology. The content based video querying takes a sample video clip as an input query and performs the searching operation in the collection of videos which are stored in the video database. This proposal, introduces a novel content-based video matching and copy elimination system that finds the most relevant video segments from video database based on the given query video clip. For effective video copy elimination based on the feature extraction the proposed system applies the scheme names as Dense SIFT_OP (DSIFT_OP). This performs the feature extraction, copy elimination and effective query matching from the video collections.

This thesis overcomes the problem of video frame mining based on effective Meta information's and semantic similarity measures. The semantic similarity contains both textual and visual similarity measures. According to the discovered features and patterns, the query frame can obtain a set of relevant video frames in the refinement process. The proposed approach robustly identifies the duplicate frames and alignsthe extracted frames, which containing the significant spatial and temporal differences.

Based on the feature extraction algorithm and semantic feature identification this applies a motion matching alignment scheme image alignment and video making with extracted clips in the large video database framework. For image analysis and synthesis the image information is transferred from the nearest neighbors to a queryimage according to the distance. This framework is demonstrated through concrete applications, such as motion field prediction

and pattern analysis from a single image, pattern synthesis via object transfer, image registration and object recognition. The proposed sequence of object and distance finding yields better result for video making and video copy elimination

Keywords
SIFT,DSIFT, Dense optical flow,Dueal Threshold

1. INTRODUCTION
The rapid growth of digital information domain, particularly video source, has many challenges in retrieval. The data is being generated has necessitated the development tools for efficient search of these media. Content-based visual queries have been primarily focused on still image retrieval. This proposed system is a novel, interactive system, based on the visual paradigm, with spatiotemporal attributes playing a key role in video retrieval. This proposes an innovative algorithm for automated video object analysis and ranking.

Challenges in video mining:

Each day tens of thousands of video data are generated and published. Among these huge volumes of videos, there exist large numbers of copies or near-duplicate videos. According to the statistics on average, there are 27 percent redundant videos that are duplicate or nearly duplicate to the most popular version of a video in the search results from Google video, YouTube, and Yahoo! video search engines.

- Need to verify each frame for duplicate detection
- Number of frames may vary in every video, so the system may failed to identify that video is duplicate.

The key contributions of this paper are given as follows: this explores the use of Dense SIFT features to model the subjective perception of similarity between two frames that have been extracted from a video database. This presents a Content-based video copy detection Retrieval system which evolves and uses different image similarity measures for different users query video. Specifically, a user-supplied query video allows the system to determine which subset of a set of objective features approximates more efficiently the subjective image similarity of a specific users query.

1. Datasets
The dataset are used as input. The system initially collects and stores some user defined videos in the database. The system stores the video by frames.

The video data set may include any number of video files, about 300 records and data. The query videos are provided which are generated using the method in the following module. Specifically, each query is constructed by taking a segment of variable length from the test video data set. And the segment is embedded into a video which is not in the test data set.

2. Frame analysis and feature extraction:
Feature extraction will transform an image into a matrix of visual objects. Here, this module has used uniform patches to divide the images since they require little computational cost and achieve similar performances compared to the complex features, each image is first divided into 8 _ 8 pixel-sized patches. For each patch, the 8-bin texture histogram is calculated.

a. Histogram analysis

The Align Histogram module is a useful precursor to any image comparison routines such as template matching, cross correlation, stereo depth or simple image subtraction. The module takes as input two images and will align the colors in the "source" image with those specified in the "target" image. Thus if you wish to compare images with respect to colors this module will provide a way to greatly reduce different lighting and color shift effects that can happen between two successive image captures or even in simultaneous capture by different cameras.

b. *Spatio temporal details*

The visual appearance of a semantic concept in video has a strong dependency on the spatio-temporal viewpoint. So the need of spatio temporal detail extraction is high.

3. Segmentation:

In this module segmentation process is to be handled. According to the image size the segmentation blocks will vary like 8X16, 4X16. Segmentation is a process of extracting and representing information from an image is to group pixels together into regions of similarity.

4. Shot similarity matching:

Dense SIFT: The Dense SIFT features are a set of SIFT features which are extracted uniformly in each spatial location. Normally SIFT features are extracted at the salient locations, but dense SIFT features are proven as quite effective in image similarity matching.In each equally divided grid, an optical flow is extracted even though the location has no salient feature.

5. Video retrieval

In the matching result graph, the vertex variable represents a match between query video and stored data.To determine whether there exists an edge between two vertexes, two measures are evaluated in this module.The graph-based video sequence matching method has good scalability to the application based on sequence matching.

The proposed method can find the best matching sequence in many messy match results, which effectively excludes false “high similarity” noise and compensate the limited description of image low-level visual features. The graph-based method takes fully into account the spatiotemporal characteristic of video sequence, and has high copy location accuracy. The graph-based sequence matching method can automatically detect the discrete paths in the matching result graph. Thus, it can detect more than one copy.

2. RELATED WORK

With the rapid growth of digital devices, internet infrastructures, and web technologies, video data nowadays can be easily captured, stored, uploaded, and shared over the Web. Although general search engines have been well developed for searching videos over the internet is still a tough task. Typically, most Web search engines index only the metadata of videos and search through a text-based approach. However, without the understanding of media content, general search engines have limited capacity of retrieving relevant video information effectively. Thus, there is much scope to improve the retrieval performance of traditional meta-data based search engines through exploiting media content. Content-based video retrieval (CBVR) is becoming a promising direction for implementing reliable video search engines in future. This is retrieved from [38], [13].

Most existing content based video retrieval (CBVR) systems select video shots [38], homogeneous video regions, or semantic video objects. The major problem for using the semantic video objects for video content representation and feature extraction is that automatic semantic video object extraction in general is very hard, if not impossible [42]

The paper [16] addresses the limitations of the novel multimodal and multilevel ranking scheme approach. The first idea in [16] mainly employed text and visual information’s in the ranking tasks. In addition, they have suggested that will include additional information from other modalities.

For example, this can study high-level concept detection techniques [1] and investigate some concept models to improve the ranking performance [41], [8]. It is likely that the proposed scheme could be improved by engaging additional information. Second, in the implementation, the fusion parameters are simply set to default parameters. In future work, this could study more intelligent solutions to determine the optimal parameters for multimodal fusion.

Third, in the current ranking solution, this considers only the query free approach for automatic search tasks. For future work, the query-class dependent weighting methods [41] can also be extended to the solution for further improving the retrieval performance. How to develop an effective query-class dependent algorithm using the graph based ranking framework will be an interesting research issue in future works. Moreover, this will apply the solution to solving other problems of content-based video retrieval, such as interactive video retrieval [7] and image/video annotation [31]. To these problems, this will explore the proposed multimodal and multilevel framework together with active learning techniques [37], [15] to overcome these open challenges. This may also study more effective kernel learning methods, such as the nonparametric kernel learning for improving the retrieval performance [14]. Lastly, this may study more efficient indexing techniques in the current solution. For large-scale video retrieval applications, determining how to index the data is important, a topic which was excluded in the previous discussion. To enable efficient solution of queries, some emerging indexing techniques, such as Locality-Sensitive Hashing [9] and SVM indexing [32], can be investigated when building an efficient ranking and indexing scheme for significant content-based video search engines and retrieval systems.

According to the statistics of [10], on average, there are 27 percent redundant videos that are duplicate or nearly duplicate to the most popular version of a video in the search results from Google video, YouTube, and Yahoo! video search engines. As a consequence, an effective and efficient method for video copy revealing has become more vital. A valid video copy detection method is based on the fact that "video itself is watermark" [10] and makes full use of the video content to detect copies.

As reviewed in [25], many content-based video copy detection methods have been proposed. Furthermore, copy is a subset of near duplicate. Copies have an origin, while near-duplicates may not. Specifically, two news videos on the same event from two broadcasting corporations are not copies, but near duplicates since they deliver the same information to audience, although some variations on the scenes may exist. Also, there are many methods proposed on near-duplicate detection. The methods on copy and near duplicate detection can be grouped into two types.

One type of copy detection methods uses global descriptor. Literature compared distance measures and video sequence matching methods for video copy detection [10], [11]. They employed convolution for motion direction feature, L1 distance for ordinal intensity signature (OIS), and histogram intersection for color histogram feature. The results show that the method using OIS performs better. Yuan et al. combined OIS with color histogram feature as a tool for describing video sequence [43].

The work in [21] and [4], with [11] as the basis, designed region intensity rank signature along time sequence. Specifically, they divided each video frames along the time sequence into several blocks and proposed average gray values for each block. Then, they linked gray values of these divided blocks separately along the time direction before they use those sequence information to describe the video content. Shen et al. [23], [5] introduced a real-time near-duplicate video detection system, UQLIPS, which globally summarized each video to a single vector.

Huang et al. [18] used global image feature such as color histogram and texture to represent each video frame. Wu et al. [13] adopted the color histogram in HSV color space to detect and remove the majority of duplicates of web videos.

Another type of methods is based on local descriptors. The local descriptors on points, lines, and shape play an important role in image and video copy detection. Among them, descriptors on points are widely used. Specifically, spatiotemporal interest points were employed to classify human actions and to detect periodic movement [24].

Willems et al. [40] presented a robust content-based video copy detection method based on local spatiotemporal features. Ke et al. used local point features for near duplicate image detection and sub image detection [36]. Law-To et al. [27] and Joly et al. [20] adopted Harris corner points [12] as feature points in video frames. And the difference of their methods lies in how to describe the feature points. Specifically, Law-To et al. [27] selected four different locations at the space around interest points (i.e., these four locations are in the same frame) when they describe the feature points, while Joly et al. [20] selected four different locations around interest points in both time and spatial domain. Besides, Law-To et al. [27] also described thetrajectory characteristic of feature points and used labels (such as "background" or "movement") to label some feature points.

This method can effectively improve the robustness and discriminative ability of video signature. Similarly, Satoh et al. [34] detected duplicate scenes by using the trajectory characteristic of the feature points. Zhou et al. [44] proposed a shot-based interest point selection approach for near-duplicate search. Methods based on global descriptor are carried out primarily by using spatiotemporal low-level features of the whole image. The features used include color histogram.

3. METHODOLOGIES

This chapter discus about the contributions and methodologies involved in the proposed work. By using the auto ST segmenting method, continuous video frames can be segmented into temporally continuous and visually similar video segments. Three frames are extracted from each video segment, which are the first frame, the key frame and the last frame of this segment.

- DSIFT_OP **(**DSIFT **+** OP) **-**The similarity score by matching with DSIFT (Dense Scale Invariant Feature Transform) on SIFT features and optical flows.
- Spatio-temporal pattern matching.
- motion matching alignment scheme
- Weight assignment based on the feature.
- Temporal video segmentation method

3.1 Video Extraction Algorithm

Generating semantic template (ST) to support high-level video retrieval.The system performs the above method to perform video extraction. The following steps are involved

Input:Video (V), frame (F)

Output: Key point descriptor (K), matched frames and combined video

Step 1: Segment the video frames and extract features of the key frames.

The first step performs the temporal video segmentation method to segment the video sequences into frames.Frames [F] =Split (V_T)

Step 2: Then extracts SIFT features of the key frames.

Dense SIFT technique has been applied in this step. This collects more features at each location and scale in a frame, this helps at increasing recognition accuracy accordingly.

Step 3: Store the frame in the database.

Step 4: Perform the step 2 for the query video or frame.

Step 5: Match the query video and target video by DSIFT OP method.

This performs the following steps

- Spatio-temporal pattern matching.
- Weight assignment based on the feature.
- Performs visual and textual similarity matching

Step 6: Retrieve the set of frames matched for the query frame

Step 7:Perform the motion matching alignment scheme for video making from the retrieved frames

Step 8: Use DSIFT_OP descriptor and key points for alignment.

To better represent the local content of video frames, the system uses Dense SIFT with OP descriptors to represent the local content of video frames present the video sequences. On the other hand, since the number of Dense SIFT feature points in video sequences is large, it thus exists high computational cost for video copy detection.

By using the spatio temporal indexing methods, the temporal information of the Dense SIFT feature points in different frames will be extracted. The process of this scheme is to match the two Dense SIFT feature sets in two video frames and make use of the temporal information of video frames.

3.2 Dense Sift Algorithm

Dense SIFT derives from SIFT algorithm, which is an important key point based approach in the content mining. The dense SIFT features are a set of SIFT features which are extracted uniformly in each spatial location. Normally SIFT features are extracted at the salient locations, but dense SIFT features are proven as quite effective in video similarity matching. The SIFT features are clustered into an arbitrary number of groups by the k-means clustering. The number of clusters is empirically chosen. SIFT finds all the key points in the frame with respect to the gradient feature of each pixel. Every key point contains the information of its location, local scale and orientation. Then, based on each key point, SIFT computes a local image descriptor which shows the gradient feature in the local region around the key point. Combining all the local descriptors, this gets the complete features from the image.

Generally, for generic object category recognition, better results are obtained using dense feature extraction rather than key point-based feature extraction. Dense SIFT collects more features at each location and scale in an image, increasing recognition accuracy accordingly.

The proposed method consists of the following steps.

Step 1: Matrix A^{Nm}= (A1;A2; . . .;Am) represents the feature point set of image A and matrix B^{Nn} = (B1;B2; . ..;Bn) represents the feature points set of image B, respectively. In another word, Ai(i = 1; . . .;m) and Bj(i =1; . .. ; n) represent DesnseSIFT feature points in image A and B respectively. The dimension of Ai and Bj is N(N = 128).

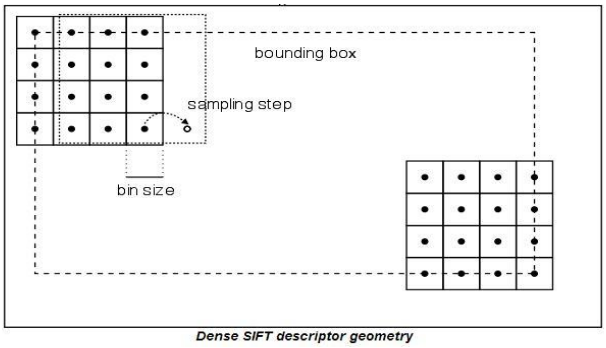

Fig: 3.2.1 Dense SIFT descriptor geometry

The above figure (3.2.1) represents the Dense SIFT descriptor process. In this proposed work the scheme uses the characteristics of the image dense SIFT feature matrix and its optimal temporal threshold value to match the Dense SIFT feature point sets of images. Suppose that A and B represent two images containing m and n Dense SIFT feature points, respectively, the objective of the algorithm is to match two point sets and compute the similarity between two images.

Step 2: Repeat the first step in order to get more description points

DesnsepointDp =D (A, B)

Step 3: Generate the global descriptor and match the data with the resource result graph according to the matching results. In the matching result graph, the vertex Mij represents a match between CQ i and CT j. To determine whether there exists an edge between two vertexes, two measures are evaluated.

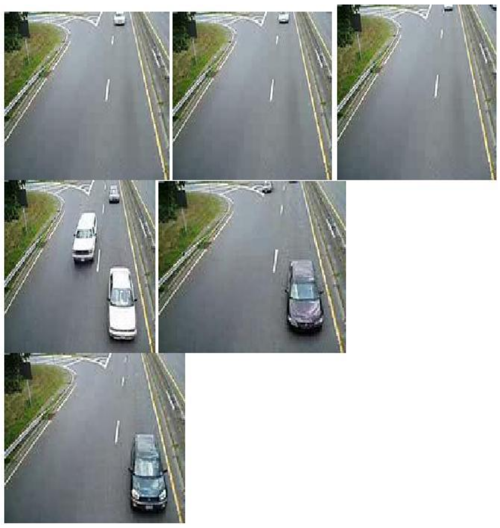

Fig: 3.2.2 Temporal video segmentation method

3.3 Dense Optical Flows

In each equally divided grid, an optical flow is extracted even though the location has no salient feature. For the temporal dimension, the optical flows are calculated in the original video clips. Frame extraction and video making with effective duplicate elimination is the highlight of the proposed system.

When the query given by the user the system extracts the DSIFT descriptor and matches with the video repository. In the application of video retrieval, the problem of searching for frames from huge or large databases is still a challenging one. The proposed implements the visual and textual similarity matching and aligning techniques in large set of frames using graph based approach. This means that the search will analyze the actual content of the video by using colors, shapes and textures.

One process involved in the frame extraction is the Dense SIFT (DSIFT) descriptor, which is extracted at a single scale for all the pixels in the image. Establishing correspondences between two images is then performed either locally or by using global optimization schemes such as the Dense SIFT-Flow algorithm

In each equally divided grid, an optical flow is extracted even though the location has no salient feature. For the temporal dimension, the optical flows are calculated in the original video clips. Frame extraction and video making with effective duplicate elimination is the highlight of the proposed system.

When the query given by the user the system extracts the DSIFT descriptor and matches with the video repository. In the application of video retrieval, the problem of searching for frames from huge or large databases is still a challenging one.

The proposed implements the visual and textual similarity matching and aligning techniques in large set of frames using graph based approach. This means that the search will analyze the actual content of the video by using colors, shapes and textures.

One process involved in the frame extraction is the Dense SIFT (DSIFT) descriptor, which is extracted at a single scale for all the pixels in the image. Establishing correspondences between two images is then performed either locally or by using global optimization schemes such as the Dense SIFT-Flow algorithm

By using the temporal video segmenting method that will perform the continuous video frames temporally that can be segmented into temporally continuous and visually similar video segments. The frame extraction technique extracts all frames from each video segment, and performs the descriptor for each key frame, the key frame and the last frame of this segment.

The keyframe is determined by the frame that is the most similar to the average frame (i.e., the average feature value of all the frames within the segment). The keyframe is used for video sequence matching, while the first and the last frames for accurately determining the segment location for copy detection and assisting matching. Each segment is assigned a continuous ID number along the time direction.

Fig: 3.3.1 Query Video

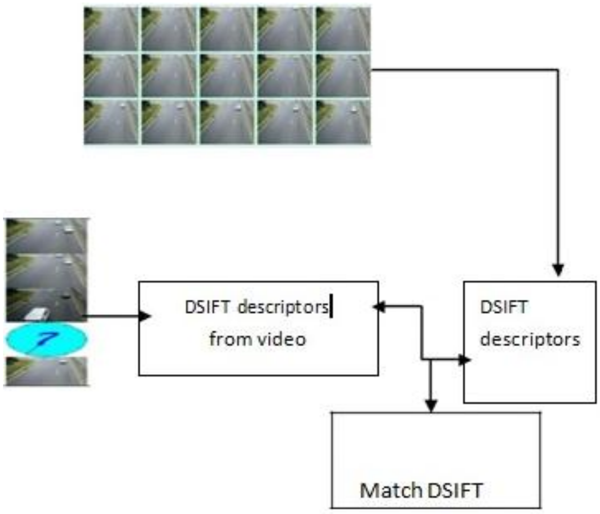

Fig: 3.3.2 Extracted Frame From Query Video

For continuous frame matching, Dense Sift Invariant Feature Transform (DSIFT) and optical flow have been implemented. SIFT has been widely used to match two or more frames. DSIFT features extracted from key frames may not good enough to capture the movement of objects in the frames. Thus, this also uses the optical flow to capture the continuous frames and objects in the sequence of frames in a video clip.

3.4 Evaluation Of The Dense Scene Alignment

The system needs further evaluation for Dense SIFT which performs compared to human perception of scene alignment. DSIFT optical flow is such a well-defined problem that it is straightforward for humans to annotate motion for evaluation In the case of Dense SIFT flow, however, there may not be obvious or unique pixel-to-pixel matching as the two images may contain different objects, or the same object categories with very different instances. To evaluate the matching obtained by Dense SIFT flow.

3.5 Predicting Motion Fields from a Single Image

The frames are aligned by predicting the motion fields from a single image from video. When predicting the motion fiends from a single image, namely to know the frames and its sequence. This adds potential temporal motion information onto a single image for further applications, such as animating a still image and event analysis. A scene retrieval infrastructure is established to query still images over a database of videos containing common moving objects. The database consists of sequences depicting common events, such as cars driving through a street and kids playing in a park. Each individual frame was stored as a vector of dense SIFT features, as described. In addition, this stores the temporal motion field estimated using between every two consecutive frames of each video. This compares two approaches for predicting the motion field for a query still image. In the first approach, using the DenseSIFT based optimal content matching this retrieves nearest neighbors (similar video frames) that are roughlyspatially aligned with the query, and directly transfer the motion from the nearest neighbors to the query.

3.6 Image Alignment

This has demonstrated that Dense SIFT flow can be effectively used for retrieval and synthesis purposes. In this section this shows that Dense SIFT flow can be applied to traditional image alignment scenarios to handle challenging registration problems in the video retrieval. Same-scene image registration Image registration of the same 3D scene can be challenging when there is little overlap between two images or drastic appearance changes due to phenomena such as changes of seasons and variations of imaging conditions geographical deformations, and human activities. Although sparse feature detection and matching has been a standard approach to image registration of the same scene.

4. CONCLUSION AND FUTURE WORK

In this paper, the study addressed the problem of extracting and storing video clips for content-based video query. The frame descriptors are found using a strong descriptor such as Dense SIFT learnt from a learning algorithm. Then, the similarity of video clips is calculated by the spatio-temporal pattern matching which includes spatio temporal dimension into the matching schema using the local and global descriptors. the experimental evaluation using traffic related videos and other synthetic dataset shows that the spatio temporal dimension is an effective feature to match video clips improves the accuracy and performance. This also performed a comprehensive comparison with existing methods and showed that the Dense SIFT OP with spatio-temporal matching systems outperformed.

The main objective of the study is to remove the duplicate data and extracting the meaningful information to the human expected needs based on the query image or video. The images are extracted with DSIFT descriptor techniques and the visual score calculation is highly focused. Here, images are segments based on RGB Components. The threshold method based on the key frame is used to compare the images. This application can be used in future to extract and make video in animation and video image retrieval engines.

In this study, this introduces a new scheme using data mining framework that supports spatial temporal data based video retrieval and image matching. Feature descriptors are extracted from image tiles and summarized into visual thesaurus. Visual thesaurus allows us to record spatial relationships among labeled descriptors using DSIFT approach.

In future the system may implement strong data mining technique to deal the storage problems. The current study does not deal with the storage based issues. The system may improve with the above constraint in future. The proposed scheme is implemented as an application, in future this will be implemented in real time. Secondly if the video input is very large the frames will be more so it takes very long time to split the video into frames so this should be reduce in future. The system can also extend with some other descriptors other than dense descriptor. Effective pattern matching and object matching in the frames will helps us to identify the duplicate frames and videos

5. ACKNOWLEDGMENTS

At outset I owe my gratitude to the Almighty for giving me the courage and insight to complete this work. Especially, I would like to give my special thanks to my parents for their patient love and encouragement enabled me to complete this work. I would also thank the editor and the members who accepted my paper for publishing in“International journal for computer Applications (IJCA)”Thank you.

6. REFERENCES

[1] Amir, G. Iyengar, J. Argillander, M. Campbell, A. Haubold, S. Ebadollahi, F. Kang, M. R. Naphade, A. Natsev, J. R. Smith, J. Tesic, and T. Volkmer, “IBM research trecvid-2005 video retrieval system,” in Proc. TRECVID Workshop, Washington, DC, 2005.

[2] Anjewierden, A., Koll¨offel, B., and Hulshof C., “Towards educational data mining: Using data mining methods for automated chat analysis to understand and support inquiry learning processes”. International Workshop on Applying Data Mining in e-Learning, ADML'07, Vol-305, Page No 23-32Sissi,LassithiCrete Greece, 18 September, 2007.

[3] Chapman, P., Clinton, J., Kerber, R., Khabaza, T.,Reinartz, T., Shearer, C. and Wirth, R... “CRISP-DM 1.0 : Step-by-step data mining guide, NCR Systems Engineering Copenhagen (USA and Denmark), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringenen BankGroup B.V (The Netherlands), 2000”.

[4] L. Chen and F.W.M. Stentiford, “Video Sequence Matching Based on Temporal Ordinal Measurement,” Pattern Recognition Letters, vol. 29, no. 13, pp. 1824-1831, Oct. 2008.

[5] R. Cheng, Z. Huang, H.T. Shen, and X. Zhou, “Interactive Near- Duplicate Video Retrieval and Detection,” Proc. ACM Int’l Conf. Multimedia, pp. 1001-1002, 2009.

[6] Chenxia Wu, Jianke Zhu, Jiemi Zhang College of Computer Science, Zhejiang University, China. A Content-based Video Copy Detection Method with Randomly Projected Binary Features

[7] M. Christel and R. Yan, “Merging storyboard strategies and automatic retrieval for improving interactive video search,” in Proc. Int. Conf. Image and Video Retrieval (CIVR), Amsterdam, The Netherlands, 2007.

[9] S. Dagtas, W. Al-Khatib, A. Ghafoor, and R. Kashyap, “Models for motion-based video indexing and retrieval,” IEEE Trans. Image Process., vol. 9, no. 1, pp. 88–101, Jan. 2000.

[10]M. Datar, N. Immorlica, P. Indyk, and V. S. Mirrokni, “Locality-sensitive hashing scheme based on p-stable distributions,” in Proc. 20thAnnu.Symp. Computational Geometry, New York, 2004, pp. 253–262.

[11] A. Hampapur and R. Bolle, “Comparison of Distance Measures for Video Copy Detection,” Proc. IEEE Int’l Conf. Multimedia and Expo (ICME), pp. 188-192, 2001.

[12] A. Hampapur, K. Hyun, and R. Bolle, “Comparison of Sequence Matching Techniques for Video Copy Detection,”Proc. SPIE, Storage and Retrieval for Media Databases, vol. 4676, pp. 194-201, Jan.2002

[13] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” Proc. Fourth Alvey Vision Conf., pp. 147-151, 1988.

[14] X. Wu, C.-W.Ngo, A. Hauptmann, and H.-K. Tan, “Real-Time Near-Duplicate Elimination for Web Video Search with Content and Context,” IEEE Trans. Multimedia, vol. 11, no. 2, pp. 196-207, Feb. 2009.

[15] S. C. Hoi, R. Jin, and M. R. Lyu, “Learning non-parametric kernel matrices from pairwise constraints,” in Proc. 24th Int. Conf. Machine Learning (ICML’07), OR, June 20–24, 2007.

[16] S. C.H. Hoi, R. Jin, and M. R. Lyu, “Large-scale text categorization by batch mode active learning,” in Proc. 15th Int. World Wide Web conference (WWW’06), CITY?, U.K., May 23–26, 2006.

[17] Hoi, Steven CH, and Michael R. Lyu. "A Multimodal and Multilevel Ranking Scheme for Large-Scale Video Retrieval." IEEE TRANSACTIONS ON MULTIMEDIA 10.4 (2008): 607.

[18] Hong Liu, Hong Lu, Member, IEEE, and XiangyangXue, Member, IEEE

[18] Z. Huang, H.T. Shen, J. Shao, B. Cui, and X. Zhou, “Practical Online Near-Duplicate Subsequence Detection for Continuous Video Streams,” IEEE Trans. Multimedia, vol. 12, no. 5, pp. 386-397, Aug. 2010.

[19] Ji Zhang Wynne Hsu Mong Li Lee Image Mining: Trends and developments.

[20] A. Joly, O. Buisson, and C. Frelicot, “Content-Based Copy Retrieval Using Distortion-Based Probabilistic Similarity Search,” IEEE Trans. Multimedia, vol. 9, no. 2, pp. 293-306, Feb. 2007.

[21] C. Kim and B. Vasudev, "Spatiotemporal Sequence Matching for Efficient Video Copy Detection," IEEE Trans. Circuits and Systems for Video Technology, vol. 15, no. 1, pp. 127-132, Jan. 2005.

[22] KuratThearling .Foundation of data mining www.thearling .com

[23] Kwang-Ting Cheng Dept. of Electrical and Computer Engineering University of California, Santa Barbara , CA 93106, USA.

[24] I. Laptev and T. Lindeberg, "Space-Time Interest Points," Proc. Int'l Conf. Computer Vision, pp. 432-439, 2003.

[25] Larose, D. T., "Discovering Knowledge in Data: An Introduction to Data Mining", ISBN 0-471-66657-2, ohn Wiley & Sons, Inc, 2005.

[26] J. Law-To, C. Li, and A. Joly, "Video Copy Detection: A Comparative Study," Proc. ACM Int'l Conf. Image and Video Retrieval, pp. 371-378, July 2007.