International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Survey on Nearest Duplicate Video Retrieval

Shital Bhirud[1], P. S. Desai[2]

[1]PG Student, Pune University, STES's, Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

[2]Professor, Pune University, STES's, Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: ***Creating video recordings of events such as lectures or meetings is increasingly less expensive and easy. Thus the Video data is increasing in a great deal on World Wide Web (www) and so thus the need of more efficient and correctly functioning method of video indexing, grouping and video retrieval in WWW or Large video archives is necessary. Similarly a method is needed for the Near-Duplicate Video Retrieval (NVDR). There are many applications for this technique, in copyright protection, online video monitoring, video tagging etc. Though, lot of techniques has been developed for NDVR over the past few years. Many of these techniques use the single feature for video representation for NVDR. Clearly using only one feature is not a best choice for any technique. Similarly, previous techniques were focused on the accuracy of the results for the NDVR. But no technique was successfully managed large scale video datasets. Thus, these techniques were having less usage in real world applications.***

Keywords: Near-duplicate videos (NDVs), video retrieval, optimization, video indexing, temporal correlation, video copy detection.

1. Introduction

Near-duplicate video retrieval is getting to be more vital with the exponential development of the Web. With the exponential development of video imparting websites, e.g. YouTube, the quantity of videos being searchable on the Web has colossally expanded. Examinations on heading social video offering stages uncover a high measure of redundancy with covering or duplicate substance [1]. Wu et al. [2] demonstrated that there are as a rule 27% duplicates among the query items of 24 prominent questions from YouTube, Google Video and Yahoo Video. To abstain from being overpowered with duplicates in the video output and use such redundancy for different assignments, for example, mining the interior structure of video database, automatic video tagging and so forth, it is crucial to creating a near-duplicate Web video retrieval framework. Despite the fact that different methodologies have been proposed to address this issue, they are fundamentally concentrating on the retrieval exactness while infeasible to inquiry on Web scale video database progressively. To rundown a couple of issues that happens because of near-duplicate videos are, an ordinary situation could be that a Web client needs to get some novel videos yet winds up with bunches of near-duplicate videos in the top-positioned indexed lists returned by the web search tool. An alternate circumstance could be that a video maker hopes to evade their copyright secured videos being imparted on the Web. Both events oblige NDVR procedures to attain their objectives. To detect nearly duplicated videos from the large collection of videos, general procedure is-

1) Firstly, a video is divided into sequence of frames.
2) On the basis of the visual features in the video keyframes are extracted..
3) This signature of the video is then compared with other videos .
4) The most similar videos are returned as NDVs.

In traditional ways for finding NDVs, single features were extracted for similarity. But as the complexity of video is increasing, single feature cannot represent the video completely, and the results will be incomplete or even incorrect. So many techniques are found for efficient NDVR. Some of them find more than one feature, some use more efficient techniques for feature extraction. Such techniques are discussed in section 2.The remaining paper can be summed as: section 2 studies some previously developed methods for NDVR. Section 3 describes the proposed methodology on the basis of the analysis of section 2. Section 4 concludes our paper, with some future works.

2. Literature Survey

2.1 NDVR

NDVR have been effectively considered in numerous types of genuine applications [3], [4]. Different methodologies utilizing diverse features and matching algorithms have been proposed for NDVR in recent years. The current methods on NDVR can be generally isolated into two categories, first is global feature based methodologies and another is local feature based methodologies. A local feature based methodologies have likewise pulled in much research consideration. Features, for example, color, surface and shape removed at the key-frame level are further portioned into numerous area units, and are especially suitable for recovering NDVs with complex varieties. The well-known local feature based methodologies incorporate key-frames based local feature recognition approach (e.g., Filter [6]). As an option to global feature based methodologies most of the current global feature based NDVR methodologies accentuate the quick distinguishing proof of NDVs [2], [5]. In these works, videos are spoken to by conservative global features. These methodologies perform well in consideration of just about identical videos. Case in point, in [2], the creators receive HSV to speak to their key-frames and further produce a feature signature by cumulating all the key-frames in the feature. This representation attains quick recovery speed and high precision in their dataset.

2.2 Multiple Feature Fusion:

The key issue lying in multiple feature fusion is the identification of the similarity or correlation between two perceptions spoke to by multiple features [7]. Late fusion methods and early fusion procedures are the conventional techniques for multiple-source fusion. The late fusion algorithms [8], [9], [10] first create separate results from distinctive features, and afterward consolidate these results together by diverse techniques. The strategies don't consider the correlation among features. Furthermore, late fusion is computationally more costly for preparing. The early fusion systems attempt to join multiple features at the data stage [11]. For instance, the algorithms proposed in [12] venture diverse features into a brought together space in which their applications can be performed. The [13] system used the local descriptors for compact representations. Though, this method is very efficient for the duplication detection, the compact generation process may cause some information loss.

Table 1: Comparison between different NDVR techniques

Name of the Paper	*Methodology*	*Merits*	*Demerits*
Effective Multiple Feature Hashing for Large-Scale Near-Duplicate Video Retrieval [7]	Multiple Feature Fusion	Use of multiple visual features	Correlations between key-frames are not used.
Understanding Near-Duplicate Videos: A User-Centric Approach [3]	Human perception in Near-duplicate video clips	User centric approach	Does not clearly determine the semantic similarity between two videos.
Compact video description for copy detection with precise temporal alignment [13]	Compact video description with temporal alignment	Use of precise temporal alignment	Might cause information loss.
Exposing digital forgeries in video by detecting duplication [14]	Use of overlapping sub-sequences	Uses both, temporal and spatial correlations	Use of over-lapping subsequence, instead of key-frames.
An image-based approach to video copy detection with spatio-temporal post-filtering [16]	Image based approach	Use of spatial and temporal filtering	Selects fixed number of frames.

2.2.1 Categories of Near-Duplicate Videos:

a) Formatting differences:

1. Encoding format: flv, .wmv, .avi, .mpg, .mp4, .ram ...
2. Frame rate: 15fps, 25fps, 29.97fps …
3. Bit rate: 529kbps, 819kbps …
4. Frame resolution: 174x144, 320x240, 240x320 …

b) Content differences:

1) Photometric variations: Color change, lighting change, etc.
2) Editing: Logo insertion, adding borders around frames, superposition of overlay text.
3) Content modification: Adding unrelated frames with different content at the beginning, end, or in the middle.
4) Versions: Same content in different lengths for different releases.

2.3 Temporal Correlation

The temporal correlations are used to detect the similarity between the two frames. A spatial correlations and temporal correlations techniques are used widely in image and video processing applications like [14], [15]. Both these features are used in the techniques like [16], to measure the similarity between the two videos

3. Proposed Methodology

We have proposed a new algorithm for the near-duplicate video retrieval (NDVR) for more accurate and more efficient retrieval of the videos. This technique is based on the multiple visual features, unlike most of the previous techniques. We have used the temporal correlation method for the detection of the duplication. In this technique, duplicate frames from the same video are detected using the temporal correlations. Only single instance is taken for such duplicate frames. Once the frames are obtained, these frames are compared with the key frames of the other videos. This method optimizes NDVR. The proposed method overcomes the problem of near-duplicate video retrieval problem. This method will be more successful than older ones, because it can be easily implemented on the large scale video databases. And can also be used in the real world applications. As, the system makes use of multiple features of the videos, the result seems to be more clear and correct.

4. Conclusion

In this paper, we have studied the different aspects of the near-duplicate video retrieval. And have also studied the importance of the NVDR problem and the different techniques that has been developed and used in past years. We have also seen, the generation of the duplicated videos. And we have finally proposed a method to overcome the drawbacks of the previous methods in NDVRs. We have used the multiple feature fusion and the temporal correlations to determine the duplicated video contents. We are not claiming to have developed a complete method to absolutely detect the duplicated videos, but we have indeed taken a step ahead.

References

[1] S. Poullot, M. Crucianu, and O. Buisson, “Scalable Mining of Large Video Databases Using Copy Detection”, In ACM MM, 61–70, 2008.
[2] X. Wu, C.-W. Ngo, and A. G. Hauptmann, “Practical Elimination of Near-Duplicates from Web Video Search”, In ACM MM, 218–227, 2007.
[3] M. Cherubini, R. de Oliveira, and N. Oliver, “Understanding Near-Duplicate Videos: A User-Centric Approach,” in Proc. ACM Multimedia, 2009, pp. 35–44.

[4] J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford, "Video Copy Detection: A Comparative Study," in Proc. CIVR, 2007, pp. 371–378.
[5] H. T. Shen, X. Zhou, Z. Huang, J. Shao, and X. Zhou, "UQLIPS: A Real Time Near-Duplicate Video Clip Detection System," in Proc. VLDB, 2007, pp. 1374–1377.
[6] K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 10, pp. 1615–1630, 2005.
[7] J. Song, Y. Yang, Z. Huang, H. T. Shen, and J. Luo, "Effective Multiple Feature Hashing for Large-Scale Near-Duplicate Video Retrieval", IEEE Transactions on Multimedia, vol. 15, no. 8, December 2013.
[8] S. Tollari and H. Glotin, "Web Image Retrieval on Imageval: Evidences on Visualness and Textualness Concept Dependency in Fusion Model," in Proc. CIVR, 2007, pp. 65–72.
[9] D. R. Turnbull, L. Barrington, G. Lanckriet, and M. Yazdani, "Combining Audio Content and Social Context for Semantic Music Discovery," in Proc. SIGIR, 2009, pp. 387–394.
[10] M.Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, and Y. Song, "Unified Video Annotation Via Multi-graph Learning," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 5, pp. 733–746, 2009.
[11] C. Snoek, M. Worring, and A. W. M. Smeulders, "Early Versus Late Fusion in Semantic Video Analysis," in Proc. ACM Multimedia, 2005, pp. 399– 402.
[12] B. Cui, A. K. H. Tung, C. Zhang, and Z. Zhao, "Multiple Feature Fusion for Social Media Applications," in Proc. SIGMOD, 2010, pp. 435– 446.
[13] M. Douze, H. Jégou, C. Schmid, and P. Pérez, "Compact video description for copy detection with precise temporal alignment," in Proc. ECCV, 2010, pp. 522–535.
[14] Wang, W., Farid, H., "Exposing Digital Forgeries in Interlaced and De-interlaced video", IEEE Transactions on Information Forensics and Security 2(3), 438–449 (2007).
[15] Wang, W., Farid, H., "Exposing digital forgeries in video by detecting duplication", In Proc. of Workshop on Multimedia & security in International Multimedia Conference, pp. 35–42 (2007).
[16] M. Douze, H. Jegou, and C. Schmid, "An image-based approach to video copy detection with spatio- temporal post-filtering," IEEE Trans. Multimedia, vol. 12, no. 4, pp. 257–266, 2010.