Chapter 1

Content-Based Keyframe Clustering Using Near Duplicate Keyframe Identification

Ehsan Younessian
Nanyang Technological University, Singapore

Deepu Rajan
Nanyang Technological University, Singapore

ABSTRACT

In this paper, the authors propose an effective content-based clustering method for keyframes of news video stories using the Near Duplicate Keyframe (NDK) identification concept. Initially, the authors investigate the near-duplicate relationship, as a content-based visual similarity across keyframes, through the Near-Duplicate Keyframe (NDK) identification algorithm presented. The authors assign a near-duplicate score to each pair of keyframes within the story. Using an efficient keypoint matching technique followed by matching pattern analysis, this NDK identification algorithm can handle extreme zooming and significant object motion. In the second step, the weighted adjacency matrix is determined for each story based on assigned near duplicate score. The authors then use the spectral clustering scheme to remove outlier keyframes and partition remainders. Two sets of experiments are carried out to evaluate the NDK identification method and assess the proposed keyframe clustering method performance.

INTRODUCTION

Effective clustering of video shots is an important step in applications involving content-based video analysis and retrieval. Clustering of video shots also has been applied to understand associated semantics in video organization which can lead to detecting of scenes in the video (Rasheed & Shah, 2005). Furthermore a wide range of other video-related applications from content-based annotation of video to video summarization can benefit from...
effective clustering of similar shots (Gao & Dai, 2008). Generally speaking, the different shot clustering approaches utilize either all frames in a video (Chen, Wang, & Wang, 2009; Zhang, Sun, Yang, & Zhong, 2005) or only a particular frame representing the shot, called the keyframe (Odobez, Gatica-Perez, & Guillemot, 2003) as the initial unit of video. In this study, we consider keyframe as the representative of a video shot and tackle the keyframe clustering problem.

One of the most critical issues in keyframe clustering is the similarity measure of visual information. In the video retrieval literature, scholars refer to this problem as Near Duplicate Keyframe identification. Near-Duplicate Keyframe (NDK) refers to the pair of keyframes in a video dataset that are closely similar to each other despite the minor variations due to capturing conditions, lighting, motion and editing operations. The task of NDK detection involves finding NDK pairs while that of NDK retrieval involves ranking all keyframes in the dataset with respect to their probability of being NDK to an input query keyframe. The former is useful for multimedia search (Zhao & Ngo, 2007) and linking news stories and grouping them into threads (Zhang & Chang, 2004), while the latter finds applications in query by example applications and copyright infringement detection (Ke, Suthankar, & Huston, 2004). The real challenge in identifying NDK is the moderate to significant degree of the variations caused by zooming and object motion (Zhu, Hoi, Lyu, & Yan, 2008). Figure 1 shows four pairs of NDKs varied in term of view point, camera lens and object locations.

In near duplicate analysis, local features invariant to the kind of variations mentioned above are becoming more important compared to global features. Local features are composed of detected keypoints and their associated descriptors extracted from local patches in the image. One of the most popular among them is the scale invariant feature transform (SIFT) descriptor (Lowe, 2004) and their variants such as PCA-SIFT. The former is shown to be scale invariant as well as robust to a certain degree of affine transformation. Similarly, the latter is known to be tolerant to color and photometric changes (Chang et al., 2005).

We use an efficient keypoint matching technique followed by matching pattern analysis in the NDK identification algorithm (Younessian, Rajan, & Chng, 2009) to handle extreme zooming and significant object motion. For each pair of keyframes within the story we calculate a near-duplicate score based on which a weighted adjacency matrix for the graph representation of keyframes is determined. Then we use spectral clustering to remove outlier keyframes and partition the rest of keyframes. The number of clusters is determined automatically so as to obtain well-separated and also balanced clusters.

Figure 1. Examples of NDKs with zooming, object motion and the color change
Related Content

Query Adaptation Techniques in Temporal-DHT for P2P Media Streaming Applications
[www.igi-global.com/article/query-adaptation-techniques-temporal-dht/72892?camid=4v1a](www.igi-global.com/article/query-adaptation-techniques-temporal-dht/72892?camid=4v1a)

Peer-to-Peer Filesharing Systems for Digital Media
[www.igi-global.com/chapter/peer-peer-filesharing-systems-digital/17332?camid=4v1a](www.igi-global.com/chapter/peer-peer-filesharing-systems-digital/17332?camid=4v1a)

Matching Word-Order Variations and Sorting Results for the iEPG Data Search
Denis Kiselev, Rafal Rzepka and Kenji Araki (2014). *International Journal of Multimedia Data Engineering and Management* (pp. 52-64).
[www.igi-global.com/article/matching-word-order-variations-and-sorting-results-for-the-iepg-data-search/109078?camid=4v1a](www.igi-global.com/article/matching-word-order-variations-and-sorting-results-for-the-iepg-data-search/109078?camid=4v1a)

Virtual Learning Environment Blends
[www.igi-global.com/chapter/virtual-learning-environment-blends/19851?camid=4v1a](www.igi-global.com/chapter/virtual-learning-environment-blends/19851?camid=4v1a)