Chapter 32

Efficient Key Frame Selection Approach for Object Detection in Wide Area Surveillance Applications

Almabrok Essa
University of Dayton, USA

Paheding Sidike
University of Dayton, USA

Vijayan K. Asari
University of Dayton, USA

ABSTRACT

This paper presents an efficient preprocessing algorithm for object detection in wide area surveillance video analysis. The proposed key-frame selection method utilizes the pixel intensity differences among subsequent frames to automatically select only the frames that contain the desired contextual information and discard the rest of the insignificant frames. For improving effectiveness and efficiency, a batch updating based on a modular representation strategy is also incorporated. Experimental results show that the proposed key frame selection technique has a significant positive performance impact on wide area surveillance applications such as automatic object detection and recognition in aerial imagery.

INTRODUCTION

With advances in sensor development and the need for high-performance data analysis, automated video summarization has become a more and more attractive method to tackle the ever-increasing volume of multimedia. The concept of video summarization relies on the extraction of the most representative frames (key frames) which deliver the salient content and ignore irrelevant information. Key frame is the frame that can represent prominent content and information of the video shot which must summarize

and describe the features of the video. Due to the fact that the number of key frames is considerably less than the total number of frames in the dataset, key frame extraction techniques can reduce data storage space, accelerate data processing speed, and decrease false positive rates in object detection application (Essa et al., 2015) and a great deal of other applications (Liu et al., 2013; Ma et al., 2015).

Over the past decade, many key frame selection algorithms have been proposed in the literature. Earlier methods of key frame selection only retained the first or last frame for each shot (shot is an un-interrupted video sequences captured by a sensor). For example, Nagasaka and Tanaka (1991) divided the video into shots and then used the first frame of each shot as a key frame. Even though this approach is comparatively fast, its disadvantages are that the number of key frames for each shot is limited to one and the method does not capture the foremost visual content of the shot.

To overcome first or last frame issue, several cluster-based methods (Doulamis et al., 1998; Hanjalic & Zhang, 1999; Girgensohn & Boreczky, 2000; Zhang et al., 2003; Zeng et al., 2008) have been proposed. In the cluster-based method, all the frames of a shot are arranged into groups with similar visual content before performing the key frame selection process. For instance, Zhuang et al. (1998) utilized the color histogram of frames as visual content to measure similarity of two frames and extract key frames from relationship between frames within a cluster and at the cluster centroid. Sun et al. (1998) analyzed frames globally in multi-feature spaces for selecting the representative frames in a video. Chang et al. (1999) introduced a tree-structured key frame hierarchy which is not constrained to the video shot level. Girgensohn & Boreczky (2000) applied temporal constraints for extracting key frames from the frame clusters. Zhang et al. (2003) proposed a motion-based clustering algorithm to classify key frames utilizing motion compensation error. The weaknesses of these approaches are that they neglect temporal information and use iterative techniques to perform minimization that are computationally expensive.

Alternately, sequential processing based methods (Vermaak et al., 2002; Lee & Kim, 2003; Gong et al., 2014) consider both temporal and visual information of frames. Vermaak et al. (2002) employed the entropy of the normalized histogram with a time based frame distance function to obtain maximally dissimilar frames as key frames, which does not require iteration and is thus more computationally efficient. Lee and Kim (2003) presented an iteration-based key frame selection method by minimizing the temporal visual content redundancy. Zhang et al. (1997) employed color feature and motion based criteria to define key frames from video shots. Constrains of these type of algorithms is that they focus on the properties of local frame sequences compare to global attributes (Asghar et al., 2014).

In the previous work (Essa et al., 2015), a modular key frame (MKF) selection approach was developed for wide area aerial surveillance video processing. In this technique, the entire video frame dataset is divided into subset batches and then divided each frame into a few equal sized sub-regions, then applied key frame selection criteria for each sub-region (i.e., modular). Although this method is capable of capturing dominant structural information while discarding redundant ones, the breakpoints (i.e., number of frames in a batch) in a video subset needs to be predefined and tuned to achieve optimal key frame selection. This drawback is mitigated by introducing an improvement to MKF selection by adding an additional batch updating functional stage named the Batch Updated Modular Key Frame selection (BUMKF), as illustrated in Figure 1. BUMKF is designed such a way that it considers both local and global frame properties of video sequences.

In the remainder of the paper, the details of the BUMKF method are presented. The performance evaluations and comparisons using real-world datasets is also provided.

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/efficient-key-frame-selection-approach-for-object-detection-in-wide-area-surveillance-applications/213823

Related Content

Who Watches?
(2022). *Modern Day Surveillance Ecosystem and Impacts on Privacy (pp. 41-63).*
www.irma-international.org/chapter/who-watches/287143

Opposing Viewpoints on Youth Social Media Banning in the U.S. for the Combatance of Extremist Recruiting: Constitutionality and Societal Implications
Lindsay A. West, Richard V. Martin, Courtney Perkins, Jennifer M. Quateland Gavin Macgregor-Skinner (2019). *Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications (pp. 1986-1997).*
www.irma-international.org/chapter/opposing-viewpoints-on-youth-social-media-banning-in-the-us-for-the-combatance-of-extremist-recruiting/213894

The Right to Privacy Is Dying: Technology Is Killing It and We Are Letting It Happen
Sam B. Edwards III (2019). *Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications (pp. 111-134).*
www.irma-international.org/chapter/the-right-to-privacy-is-dying/213797

Political Communication and Twitter in Greece: Jumps on the Bandwagon or an Enhancement of the Political Dialogue?
Stamatis Poulakidakosand Anastasia Veneti (2019). *Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications (pp. 1125-1152).*
www.irma-international.org/chapter/political-communication-and-twitter-in-greece/213847

Citizen Perspectives on the Customization/Privacy Paradox Related to Smart Meter Implementation
Jenifer Sunrise Winter (2019). *Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications (pp. 938-954).*
www.irma-international.org/chapter/citizen-perspectives-on-the-customizationprivacy-paradox-related-to-smart-meter-implementation/213839