VICPU 2015, Australia

Fusion of Foreground Object, Spatial and Frequency Domain Motion Information for Video Summarization

Authors: Md Musfequs Salehin and Manoranjan Paul

School of Computing and Mathematics

Charles Sturt University, Australia

Abstract

Everyday an enormous amount of video is captured by surveillance system. However, this is almost impossible for human to analyze the vast majority of video data. To solve this problem, a video summarization method is introduced combining foreground object, motion information in spatial and frequency domain. Using these features, a linear weighted fusion method applied to generate a sorted list of frames based on their highest cumulative values. After that a summarized video is produced from this sorted list based on the desired length of video provided by the user.

The Proposed Method

The block diagram of the proposed method is shown in Fig.1. Foreground objects typically provide important information about video contents. To obtain this information, Gaussian mixture-based parametric background modelling (BGM) has been applied. According to the psychological theories of human attention, motion information is more significant than the static attention clues. To obtain motion information in spatial domain, consecutive frame difference is applied. However, it does not work well when global illumination changes occur. To overcome the problem, motion information is also calculated in frequency domain using phase correlation technique. Fig. 2 shows motion estimated using phase correlation technique. In contrast, frequency based motion estimation approach lacks of localization problem. Therefore, motions obtained in both spatial and frequency domains are combined with foreground objects for video summarization.

.

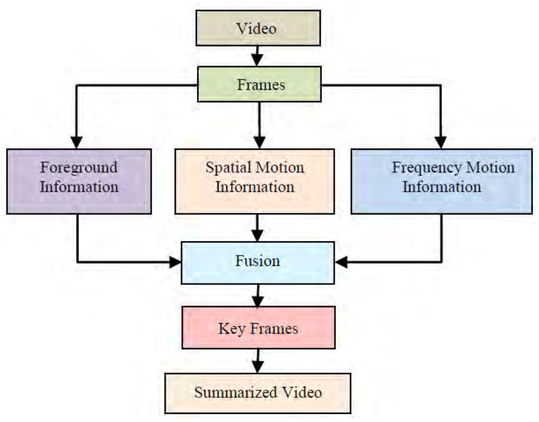

Fig 1: Framework of the proposed method.

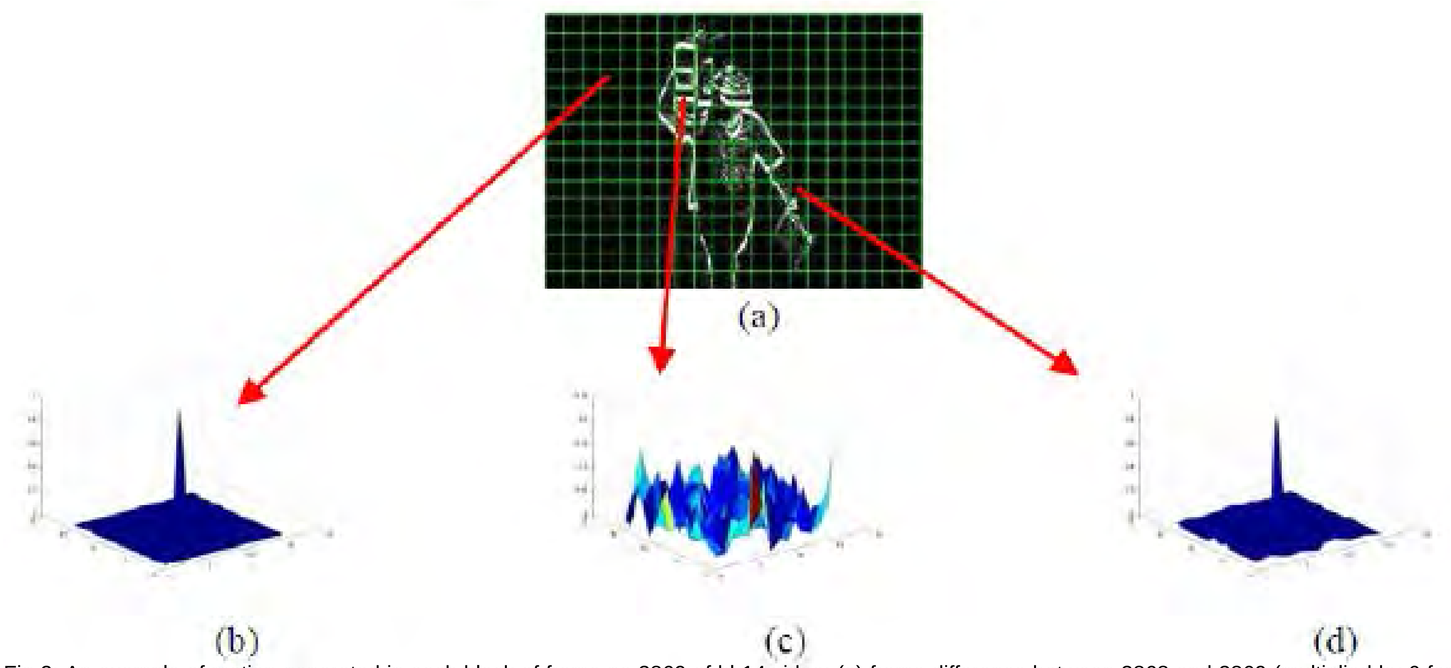

Fig 2: An example of motion generated in each block of frame no 3869 of bl-14 video; (a) frame difference between 3868 and 3869 (multiplied by 6 for better visualization), phase correlation pick with no motion, complex motion and single motion are represented in (b), (c) and (d) respectively.

Results and Discussion

The proposed method is applied on the publicly available BL-7F dataset. The foreground object, spatial and frequency domain motion information are extracted by the proposed method are shown in Fig. 3.

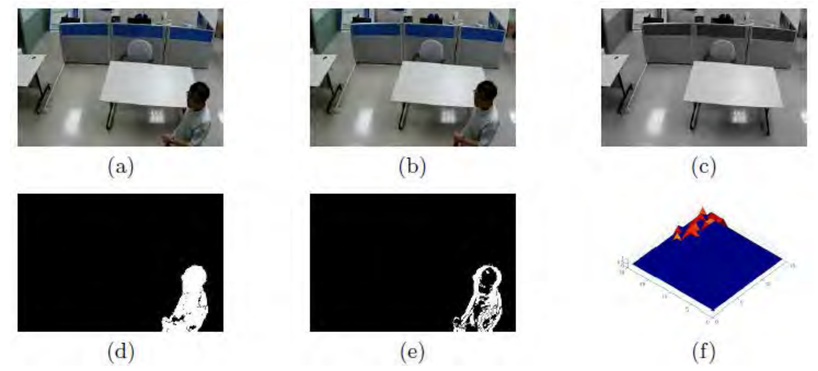

Fig 3: An example of foreground and motion information extracted by the proposed method; (a) and (b) are frame no 740 and 741 of bl-0 video, (c) is the background image of (b), (d) is the foreground image of (b) , (e) is the object motion between two frames and (f) is the motion obtained by phase correlation technique on frame no 741.

In Fig. 4, the results obtained by GMM based method as well as the proposed method are shown. The graphical representations of F1-measures are shown in Fig. 5. Some anomalies in ground truth of bl-12 video in BL-7F dataset are shown in Fig. 6.

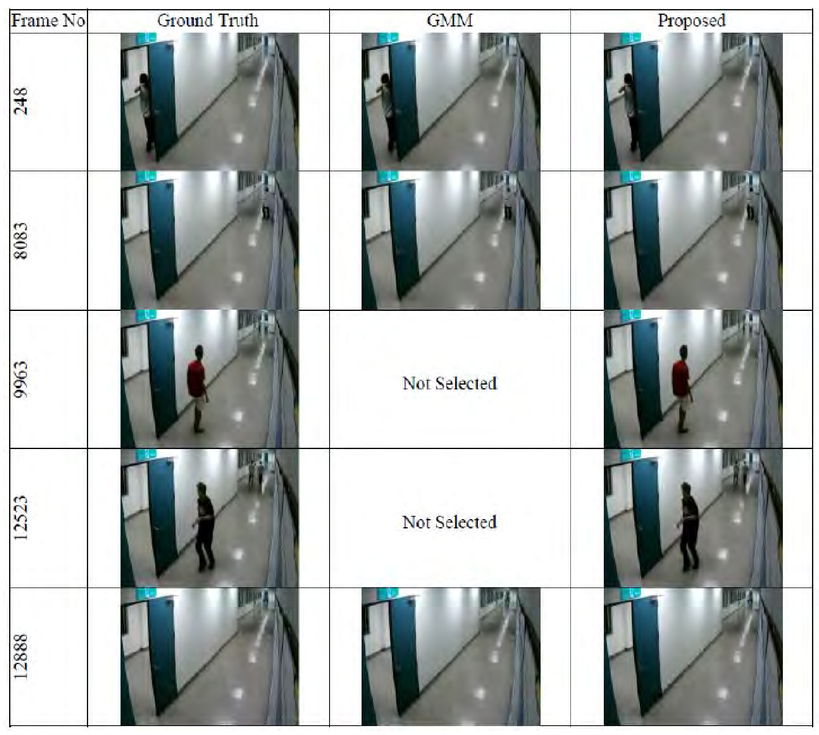

Fig 4: Evaluation of key frames extraction of bl-11 video of BL-7F dataset; first, second, third, and forth columns indicate frame no, ground truth, results obtained by GMM based method and the proposed method respectively.

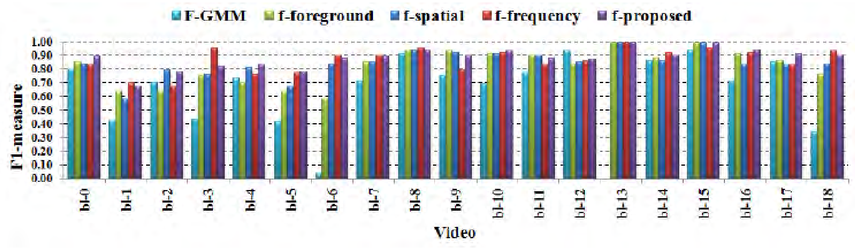

Fig 5: F1-measure of the proposed and GMM based approach.

Fig 6: Sample frames of bl-12 are not selected as ground truth (first row) and considered as key frame (second row).

Contact details:

Md. Musfequs Salehin

Phone: +61 2 6338 6172

Email: msalehin@csu.edu.au