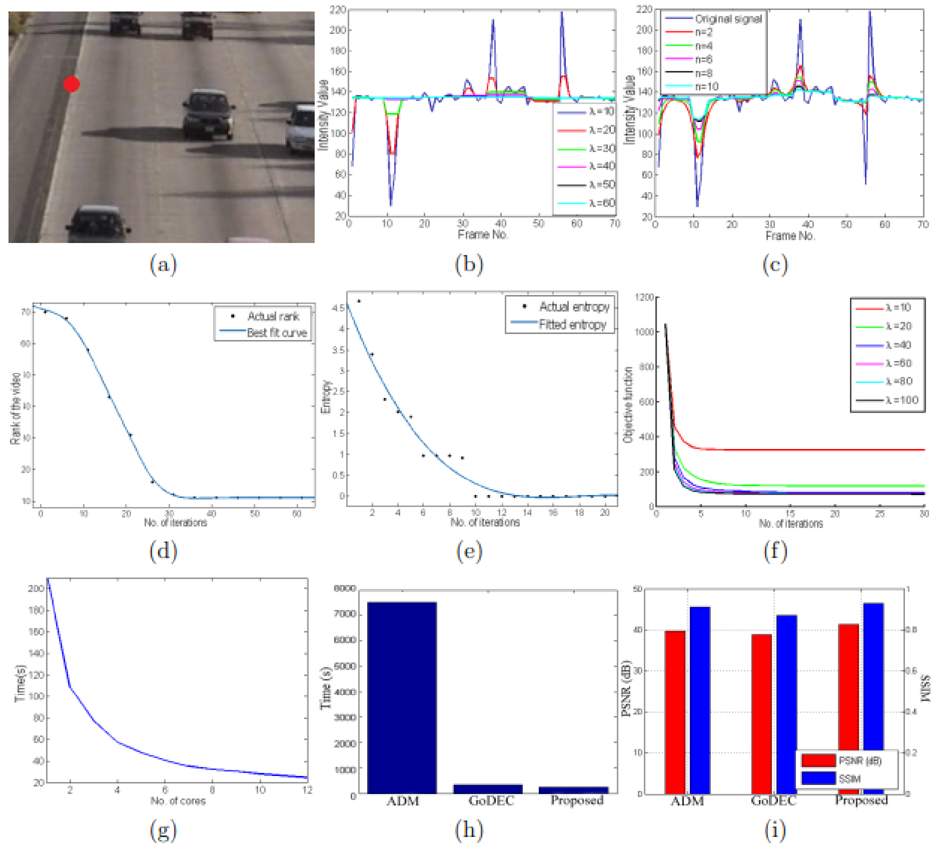
Title:	Novel Video Decomposition Techniques and Their Applications
Author(s):	Bhattacharya, Saumik
Supervisor(s):	Gupta, Sumana Venkatesh, K S
Keyword(s):	Video Decomposition Background Estimation Visual Saliency Restoration
Subject(s):	Video Processing Signal Processing

Abstract: This thesis presents novel algorithms to decompose a video into background and feature videos. Unlike the existing low rank sparse decomposition methods, the proposed decomposition techniques have significantly reduced computational complexity. Moreover, the proposed algorithms are completely parallelizable, which make them better choices for presently available advance multicore processors. The parallelizable property of the proposed decomposition algorithms also ensures that large video shots can be processed, which was not hitherto practically feasible with most of the existing decomposition algorithms. This opens up new possibilities for decomposition based techniques and encourages us to explore the effectiveness of video decomposition in diverse fields of video processing. First, we discuss the novel decomposition schemes and observe the effect of various design parameters on the respective algorithms. This will not only help to understand the proposed decomposition schemes better but also helps to choose necessary parameter values for various applications. The parallelizable nature of the algorithms is discussed, along with necessary timing comparison, with the existing decomposition algorithms. Next, we apply the proposed decomposition algorithm to detect salient regions in a video. At first, we use video decomposition to extract motion salient regions from videos captured with static cameras and show that the extraction of motion saliency using video decomposition is superior to a large class of existing methods. This work is followed by detection of visually salient regions in videos captured with camera motion or dynamic contents. The results demonstrate that the use of video decomposition results in accurate detection of salient objects and the detected regions closely resemble actual human eye tracking data. Finally, we go beyond visual perception and detect motion in scenes where motion is nearly or completely imperceptible due to its low magnitude, low frequency or both. We classify the detected motion to design a simple noninvasive method to monitor physical or biological events. As decomposition algorithms are able to separate spatiotemporal discontinuities in video, we further use video decomposition algorithms to accurately detect several artifacts in archival movies and use the detected mask for localized restoration to ensure maximum quality in the restored videos. We address some common artifacts observed in old movies, like partial color artifacts, blotches, scratches, dust, intensity flicker etc., and also discuss the effectiveness of video decomposition algorithms in various restoration processes. We demonstrate restoration results for some videos naturally degraded by single or multiple artifacts to illustrate the effectiveness of the proposed restoration algorithms in real world applications. Finally, we demonstrate some diverse applications of video decomposition in video processing. We demonstrate how decomposition can be used to summarize a video into an image using motion segmentation, to filter out temporal fluctuation commonly present in cheap depth cameras like Kinect or to generate a video storyboard. We conclude the work by analyzing this diverse set of applications to establish that video decomposition can be an effective tool in many video processing applications and the usage of video decomposition becomes practically feasible in many

of these cases because of the parallelizable nature of the proposed algorithms.



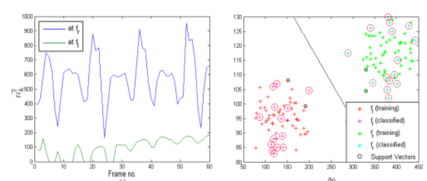


Figure 3.16: (a) n_k^T for low frequency vibration f_l and high frequency vibration f_r for *platform* dataset; (b) Output of the SVM trained for *platform* dataset. The figure shows the training data, classified data and the support vectors.

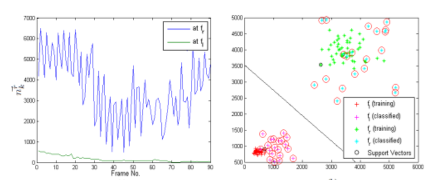
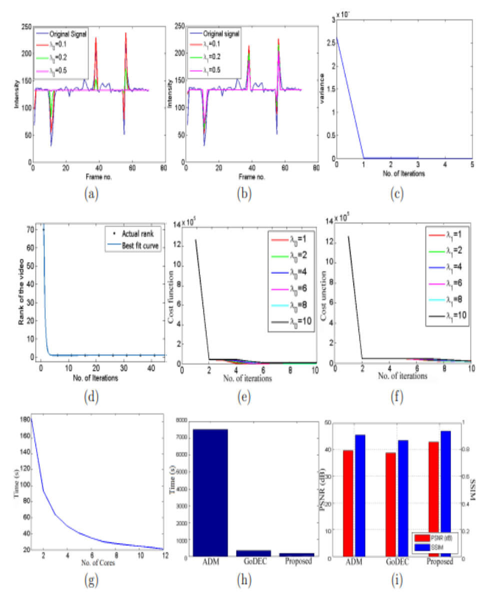


Figure 3.17: (a) n_k^T for low frequency vibration f_l and high frequency vibration f_r for *box* dataset; (b) Output of the SVM trained for *platform* dataset. The figure shows the training data, classified data and the support vectors.



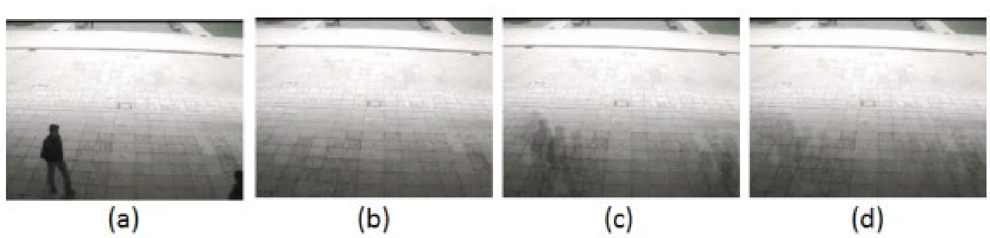


Figure 2.4: (a) Frame of an input video; (b) Ground truth of the background; respective frames from background video **L** for algorithms (c) TVD and (d) IRND.