Albert Cruz

IGERT Fellow
Electrical Engineering
University of California, Riverside

Date: Friday, October 14, 2011
Location: EBU2 Room 205/206
Time: 12:10am

A Psychologically-Inspired Match-Score Fusion Model for Video-Based Facial Expression Recognition

Abstract. Communication between humans is rich in complexity and is not limited to verbal signals; emotions are conveyed with gesture, pose and facial expression. Facial Emotion Recognition and Analysis (FERA), the set of techniques by which non-verbal communication is quantified, is an exemplar case where humansconsistently outperform computer methods. While the field of FERA has seen many advances, no system has been proposed which scales well to very large data sets. The challenge for computer vision is how to automatically and non-heuristically downsample the data while maintaining a minimum representational power that does not sacrifice accuracy. In this paper, we propose a method inspired by human vision and attention theory [1]. Video is segmented into temporal partitions with a dynamic sampling rate based on the frequency of visual information. Regions are homogenized by an experimentally selected match-score fusion technique. The approach is shown to increase classification rates by over baseline with the AVEC 2011 video-subchallenge [2].

References.

1. Findlay, J., & Gilchrist, I.: Active Vision: The Psychology of Looking and Seeing. Oxford University Press, Oxford (2003)
2. Schuller, B., Valstar, M., Eyben, F., McKeown, G., Cowie, R., & Pantic, M.: AVEC 2011 – The First International Audio/Visual Emotion Challenge. First International Audio/Visual Emotion Challenge and Workshop. Springer LNCS. (2011)