



DEPARTMENT OF COMPUTER SCIENCE

PhD Degree Oral Presentation

PhD Candidate:	Mr Qinmu PENG
Supervisor:	Professor Yiu Ming Cheung
External Examiner:	Dr Yan LIU Dr Koon Kau CHOI (Proxy for Dr. Hujun YIN)
Time:	3 Aug 2015 (Monday) 10:30 am – 12:30 pm (35 mins presentation and 15 mins Q & A)
Venue:	SCT714, Cha Chi Ming Science Tower, HSH Campus

“Visual Attention: Saliency Detection and Gaze Estimation”

Abstract

Visual attention is an important characteristic in the human vision system, which is capable of allocating the cognitive resources to the selected information. Many researchers are attracted to the study of this mechanism in the human vision system and have achieved a wide range of successful applications. Generally, there are two tasks encountered in the visual attention research including visual saliency detection and gaze estimation. The former is normally described as distinctiveness or prominence as a result of a visual stimulus. Given images or videos as input, saliency detection methods try to simulate the mechanism of human vision system, predicting and locating the salient parts in them. While the later involves physical device to track the eye movement and estimate the gaze points.

As for saliency detection, it is an effective technique for studying and mimicking the mechanism of the human vision system. Most of saliency models can predict the visual saliency with the boundary or the rough location of the true salient object, but miss the appearance or shape information. Besides, they paid little attention to the image quality problem such as low-resolution or noises. To handle these problems, in this thesis, we propose to model the visual saliency from the local and global perspectives for better detection of the visual saliency. The combination of the local and global saliency scheme employing different visual cues can make fully use of their respective advantages to compute the saliency. Compared with existing models, the proposed method can provide better saliency with clear appearance and shape information, and can work well even in the low-resolution or noisy images. The experimental results demonstrate the superiority of the proposed algorithm.

Next, video saliency detection is another issue for the visual saliency computation. Numerous works have been proposed to extract the video saliency for the tasks of object detection. However, one might not be able to obtain desirable saliency for inferring the region of the foreground objects when the video presents low contrast or complicated background. Thus, this thesis develops a salient object detection approach with less demanding assumption, which gives higher detection performance. The method computes the visual saliency in each frame using a weighted multiple manifold ranking algorithm. It then computes motion cues to estimate the motion saliency and localization prior. By adopting a new energy function, the data term depends on the visual saliency and localization prior; and the smoothness term depends on the constraint in time and space. Compared to existing methods, our approach automatically segments the persistent foreground object while preserving the potential shape. We apply our method to challenging benchmark videos, and show competitive or better results than the existing counterparts.

Additionally, to address the problem of gaze estimation, we present a low cost and efficient approach to obtain the gaze point. As opposed to eye gaze estimation techniques requiring specific hardware, e.g. infrared high-resolution camera and infrared light sources, as well as a cumbersome calibration process. We concentrate on visible-imaging and present an approach for gaze estimation using a web camera in a desktop environment. We combine intensity energy and edge strength to locate the iris center and utilize the piecewise eye corner detector to detect the eye corner. To compensate for head movement causing gaze error, we adopt a sinusoidal head model (SHM) to simulate the 3D head shape, and propose an adaptive weighted facial features embedded in the pose from the orthography and scaling with iterations algorithm (AWPOSIT), whereby the head pose can be estimated. Consequently, the gaze estimation is obtained by the integration of the eye vector and head movement information. The proposed method is not sensitive to the light conditions, and the experimental results show the efficacy of the proposed approach.

***** ALL INTERESTED ARE WELCOME *****