論 文 の 内 容 の 要 旨

Toward Constraint-Free Gaze Estimation with Unrestricted Head Motion
（自由な頭部運動を伴う拘束の少ない視線推定技術の開発）

氏　名　陸 峰

We investigate the appearance-based gaze estimation problem. Eighty percent of a human's sensory information is received by the eyes. As a result, the ability to track human gaze direction/visual attention is essential for use in an intelligent system. Among recent computer vision-based approaches, the appearance-based methods are promising because they only require a single camera and work under common conditions. However, problems exist and result in serious constraints on real systems and users. To deal with these problems, we present a framework aiming at developing a practical appearance-based gaze estimation system with minimum constraints.

First, we introduce our main technique to perform appearance-based gaze estimation, with respect to its essential difficulty in reducing the number of required training samples while still achieving high accuracy. We cast the problem as mapping high-dimensional eye image features to low-dimensional gaze positions, and propose an adaptive linear regression (ALR) method as the key to our solution. The ALR method adaptively selects an optimal set of sparsest training samples for the gaze estimation via l1-optimization. In this sense, the number of required training samples is significantly reduced. We carefully evaluate the proposed method by conducting experiments with multiple users and variant conditions.

Then, we introduce a pre-processing technique to our system to allow free head motion. Most existing appearance-based methods assume a fixed head pose because head motion significantly changes the appearance of the eye, and thus training images captured under the original head pose do not work well for test images captured under different head poses. To overcome this difficulty, our method

pre-synthesizes training images for unseen head poses before estimation. The key is to use only one set of original training images and just four additional eye images taken under four reference head poses to synthesize new training images for any unseen head poses in estimation. To this end, we propose a 1D pixel flow model that efficiently handles eye image variations due to head motion. Evaluation of the method is conducted through experiments to demonstrate its effectiveness.

Third, during gaze estimation, we further consider slight head motion, image resolution variation and eye blinking. We enhance our basic ALR technique by integrating sub-pixel alignment and blink detection techniques into its optimization framework and handle the above mentioned issues. We also evaluate the proposed method by conducting experiments with multiple users and variant conditions.

Finally, we introduce a post-processing technique that also helps to handle free head motion. This technique can be used as an alternative to the pre-processing technique, depending on the scenario. The key idea is to decompose the head pose-free gaze estimation problem into subproblems, including an initial fixed head pose problem, which can be solved by using our ALR method, and the subsequent compensations for correcting the initial estimation biases. These compensations are then performed by using either learning-based regression or geometric-based calculation. The additional training requirement is only to capture a 5-second video clip. Experiments are conducted to verify the effectiveness of the proposed method.