Crossroad of Control Theory and Vision

Takeshi Hatanaka and Mario Sznaier

Vision sensors provide a large amount of information on situation awareness such as what happens, what a target is, where it is and where it bears especially in dynamic environment. Due to their nature, vision has been and continues to be a powerful tool for completing challenging control problems. The motivating scenarios to fuse control theory and vision also currently spread over the robotic systems into security and surveillance systems, medical imaging procedures, human-in-the-loop systems and even understanding biological perceptual information processing. In addition, driven by the technological innovations of the smart wearable cameras, the vision system is expected as a component of sustainable infrastructures.

These new applications and technological innovations bring us new theoretical and practical challenges beyond the conventional visual servo control. Moreover, in addition to the direction of importing computer vision techniques into control problems, highly matured systems and control theory also has a tremendous amount of potential for contributing to solving difficulties in computer vision.

The goal of this session is thus to draw these new directions at the intersection of the vision and systems and control theory. It is worth noting that we have to deal with at least one of the following characteristics of vision-based systems when we employ data from vision sensors as measurement signals. **(1)** the relative complexity of the 3D world and the 2D image space to the extracted signal due to the projection of a 3D world onto a 2D image, **(2)** the inherent nonlinearity introduced by the projection and the resulting unique noise characteristics associated with the imaging and the signal extraction process, **(3)** extraction of only useful contexts from flood of data and contrastingly acquisition of images with rich information contents **(4)** the registration or matching of the signal across images. Indeed, the papers of the proposed session tackle more than one of these issues and draw cutting edges in this research fields.

The first paper (N. Ozay and M. Sznaier) addresses the problem of robust identification of a class of discrete-time affine hybrid systems, switched affine models, from the data subject to random instantaneous outages. They reduce the challenging problem to a semi-definite optimization problem, called rank minimization problem. The proposed methodology is useful for many applications in activity recognition drawn from computer vision, including detecting contextually abnormal activity for security and assisted living scenarios and human computer interfaces.

The second paper (Y. Zhang, M. Rotea and N. Gans) presents a novel approach to increasing the information content in visual images, where they employ entropy of an image as a measure of the information content. To overcome low entropies caused by poor camera settings, poor lighting conditions, or a scene with little low interest or activity, they present a new control methodology combining the global properties of simplex optimization methods and dynamic properties of extremum seeking control. The novel variation on extremum seeking control has applications even in other fields where the cost function may have many local extrema.

The third paper (A. P. Dani, Z. Kan, N. R. Fischer, and W. E. Dixon) presents a new state observer for estimating structure and motion of a moving object with time-varying velocities seen by a moving camera. To address the problem, the authors employ a nonlinear unknown input observer strategy, where the object's velocity is considered as an unknown input to a perspective dynamical system. This is the first paper to provide the causal, observer-based structure estimation algorithm for a moving camera viewing a moving object with unknown time-varying object velocities.

The fourth paper (T. Hatanaka, K. Hirata and M. Fujita) attempts to combine vision with cooperative control which has been one of the most active research topics for a decade in systems and control society. In the paper, the authors investigate cooperative estimation of 3D target object motion motivated by visual sensor networks consisting of multiple smart cameras. In particular, they present a novel networked observer system to achieve not only tracking of the estimates to the moving objects but also averaging their individual estimate. They also clarify averaging accuracy, convergence speed and tracking performance when the present mechanism is applied to the visual sensor network system.

In the fifth paper (P. Karasev, M. M. Serrano, P. A. Vela and A. Tannenbaum), the authors tackle the issue of non-uniform performance bounds of visual servoing controllers, caused by the nonlinearities introduced by the camera projection. To cancel the distance nonlinearity, they present a novel distance invariant visual servoing controller, where the controller measures the area of a planar visual feature on a given target while tracking the feature.

In the sixth paper (T. Ibuki, T. Hatanaka, M. Fujita and M. W. Spong), the authors present a novel visual feedback pose synchronization law for a network of rigid bodies with vision. The present controller combines a passivity-based pose synchronization with a passivity-based visual observer which are both presented in the authors' previous works. Not only the proof of synchronization but also tracking performance of the group to the moving leader are presented in the paper together with experimental demonstrations by an original multi-robot system. The results point to usability of vision in the research field of cooperative control.

The papers are ordered in such a way that how systems and control theory can impact the problems in computer vision are first presented, then the topics move to inclusion of vision into feedback loops of control systems. The research papers present the viewpoint of how control theory contributes computer vision in a positive way and otherwise how to accurately deal with vision in control problems.