Introduction to the First International Workshop on Video Processing for Security (VP4S-06)

Dmitry O. Gorodnichy and Lijun Yin

The research in video processing and recognition has received outstanding attention over the last several years, in particular due its relevance to the security-relates applications. This workshop is organized to address the demand to this area, with the goal of both providing Computer Vision scientists a market-driven inspiration and prospective to their work, and providing the Security and Biometrics industry with the references and the update of the state of the art research in the field.

Organizationally, this workshop is an extension of the past IEEE-published workshops on Face Processing in Video: FPiV'04 and FPiV'05, with its interest extended from face detection, tracking, recognition, coding to people, objects, scene and action detection, tracking and recognition. Its focus remained dealing with video data, in particular such as coming from TV, surveillance cameras or web/PDA cameras, and its theme has shifted towards security-related applications. Submissions addressing any video processing task with respect to a security application were welcomed.

As a result, fifteen papers have been submitted to the workshop, of which seven have been accepted for the publication in this proceedings. Below follows the overview of these papers.

All of the accepted papers consider a specific video processing application for security - that of observing people in video, which shows how important and challenging this application is. The first three papers address the problem of detecting and tracking people using a video camera. The next two papers extend the people surveillance problem to the multiple-sensor setups. And the last two papers approach the final task of video surveillance - that of person recognition.

The object size is one of the main attributes that allow to detect an object and, in particular, a person. To make the size of an observed object invariant to the distance between the camera and the object is the goal of the paper by Mookyung Park, Namsu Moon, Sang Rim Ryu, Jeong Pyo Kong, Y.J. Lee, and W.J. Mun, where a novel pixel-weighting method is developed for the purpose based on geometrical properties of the observes scene.

One of the advantages of surveillance video is in its stationary setup which allows one to compute the information about the observed background so that it can be used to facilitate the detection of moving objects and persons in front of it. A technique to do it using the K-means clustering of pixel values is proposed in the paper of S. Indupalli, M.A. Ali, and B. Boufama.

Once an object such as a person is detected, it can be tracked, and the object colour and texture information is known to be of the most use for this purpose. The paper by Rafik Bourezak and Guillaume-Alexandre Bilodeau introduces an approach to perform such tracking for multi-part objects using adaptive image subdivision and spacial-temporal relationship among the object parts.

The paper by Trevor Ahmedali and James J. Clark proposed a way to improve person detection by using several cameras with the overlapping fields of view and an unsupervised technique for inter-camera calibration.

In their paper, Valery A. Petrushin, Omer Shakil, Damian Roqueiro, Gang Wei, and Anatole V. Gershman combine data obtained from pan-tilt-zoom camera with the data obtained from infrared and fingerprint readers. Colour information and log-likelihood are used to track a person and the Bayesian framework is used to combine data over time and over several sensors.

It is generally known that video does not provide data sufficient for person identification [1,2]. One way to improve video-based face recognition is to use 3D face models. Reconstruction of such 3D face models from stereo video is the task addressed in the paper by Unsang Park and Anil K. Jain, where facial features are manually extracted from video and a generic 3D face model is used to make face recognition more robust to pose and illumination variations.

Video-based face recognition becomes also viable when the observed faces are forced to be close and

facing the camera. This is the case in the paper of Paolo Abeni, Madalina Baltatu, Rosalia D'Alessandro, where high recognition rates are achieved by using Fourier features and combining global classifier with four local ones, corresponding to eyes, nose, and mouth.

It can also be mentioned that in addition to the oral session where the papers accepted for IEEE publication are presented, this workshop also includes a poster session. The papers presented in the poster session are not included in the proceedings, but are intended to facilitate the discussion of new emerging techniques. The workshop website (http://computer-vision.org/4security) provides links to the pages of both poster and oral presentation authors.

To conclude, we would like to acknowledge the work of the workshop program committee and additional reviewers. Thanks to their timely and thorough reviews. Each submitted paper received three or more reviews! We would also like to thank the organizers of CRV 2006 and the IEEE for helping us organize this workshop and prepare the proceedings.

[1] D.O. Gorodnichy, “Seeing faces in video by computers”, *Image and Video Computing*, Vol. 24, No. 5 (Special Issue on Face Processing in Video Sequences. Editor: D.O. Gorodnichy), pp. 1-6, May 2006.

[2] The Face Recognition Grand Challenge at the Face Recognition Vendor Tests site: http://www.frvt.org/FRGC.