eVent Video

eVi Newsletters

#04 - December 2012

TEC2011-25995 EventVideo (2012-2014)

Strategies for Object Segmentation, Detection and Tracking in Complex Environments for Event Detection in Video Surveillance and Monitoring

http://www-vpu.eps.uam.es/eventvideo/

Fourth trimester progress report

As during the third trimester, the project has been working on plan, with a slight delay in the delivery of Technical Report TR2012.1 "Evaluation results and future research lines". Workpackages WP2 "Video object segmentation", WP3 "People modeling and detection" and WP4 "Video object tracking", have finished testing and evaluating their algorithms over the video sequences content sets (test sequences and associated ground-truths) defined in WP5 and they have contributed to the draft version of TR2012.1 "Evaluation results and future research lines". The Technical Report is still being edited and will be delivered during January 2013. Some enhancements in the algorithms have been done during the evaluation phase.

During this period, the project has submitted several papers for review in international journals and is preparing a additional papers. Additionally, the project has participated in an international evaluation competition, HARL 2012, as described below.

HARL 2012 Human activities recognition and localization competition

The goal of the HARL 2012 competition (http://liris.cnrs.fr/harl2012/) is recognition of complex human activities. In contrast to previous competitions and existing datasets, the proposed tasks focus on complex human behavior involving several people in the video at the same time, on actions involving several interacting people and on human-object interactions. The goal is not only to classify activities, but also to detect and to localize them.

The dataset used for the competition is the LIRIS human activities dataset, which consists of 10 classes. Each class can be a normal (single person) activity, a human-

human interaction, a human-object interaction, or a combination of the latter two types.

Examples of the events proposed in the LIRIS dataset

The dataset consists of two subdatasets: D1 was shot with a moving camera mounted on a mobile robot delivering videos in VGA resolution and depth images from a consumer depth camera (Primesense/MS Kinect); and D2 was shot with a consumer camcorder delivering color videos in DVD resolution. The dataset is composed by 370 (203+167) videos showing 828 (461+367) activities of mid-high complexity.

VPULab participated in the completion making use of the color videos (subdataset D2) obtaining the best results for this dataset. In the global ranking (complete dataset, D1 and D2) obtained the second position with intermediate results, but that in comparison with other participants can be classified as acceptable. The first position was obtained by a collaboration between the University of Illinois (U.S.A) and the National University of Singapore (Singapore) making use of D1, corroborating that depth is a good feature for enhancing activity recognition.

Video Processing and Understanding Lab

UNIVERSIDAD AUTONOMA DE MADRID

Fourth trimester results

Master Thesis

Fabricio A. Córdova Lucero, **"Detección de robo/abandono de objetos en interiores utilizando cámaras de profundidad (***Indoor stolen/abandoned object detection using depth cameras***)"**, Proyecto Fin de Carrera (*Master Thesis*), Ing. Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, December 2012.

Abstract: The detection of abandoned and stolen objects is a very important application of video-surveillance systems. Among the many stages involved, the segmentation of these objects from the background becomes a very fundamental and critical stage to perform the detection. In the same way, another significant step is the objects discrimination stage, which decides the final event (stolen or abandoned).
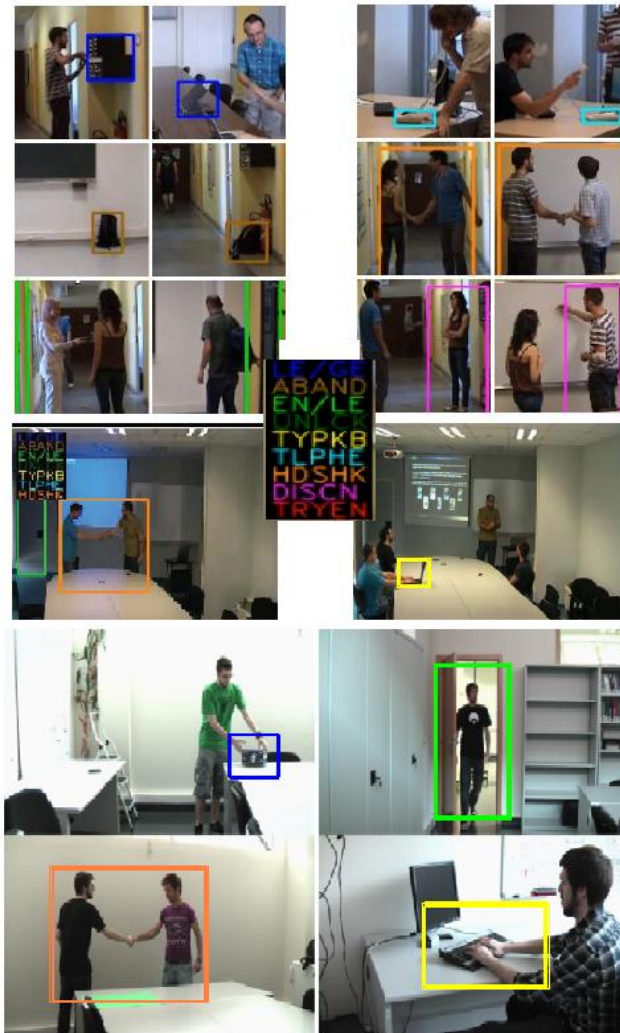
In this work, we provide new procedures for these two stages by making use of a new technology of image capturing: The Kinect Sensor. This device provides us with a new type of information, which indicates the depth of each pixel from a given color image. The combination of depth and color information gives us synchronized color-depth video streams (RGBD) whose data is more accurate and reliable.

More specifically, we propose to take advantage of the features of this new information (depth map) such as immunity to changes in illumination or shadows discrimination to increase the robustness of the current systems. The improvements are applied to the already existing prototype in the VPU-Lab group. The improved system is then subjected to several tests, whose results are then compared to those obtained with the previous system.
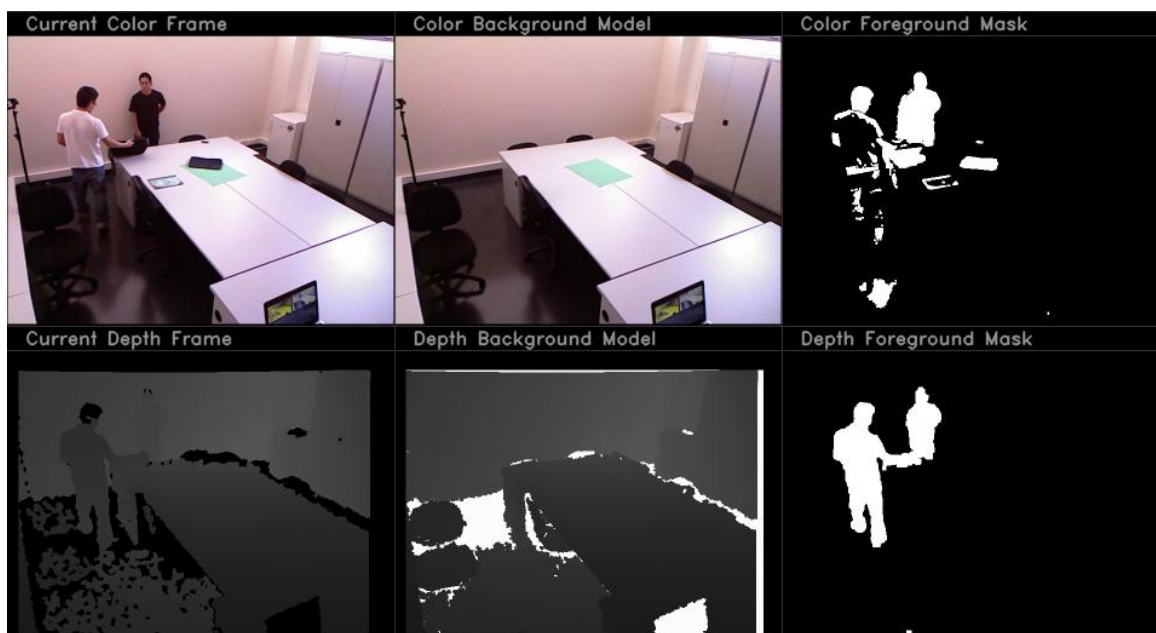
Sergio Suja Garrido, **"Análisis de interacciones y actividades en entornos controlados (*Analysis of interactions and activities in controlled environments*)"**, Proyecto Fin de Carrera (*Master Thesis*), Ing. Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, December 2012.

Abstract: In this work, we propose to contribute to the state of art in the field of automatic event recognition focusing on the analysis of interactions and activities of persons in highly controlled environments such as meeting rooms or indoor video-surveillance. The goal is to improve the prototype developed by the Video Processing and Understanding Lab of the Universidad Autónoma de Madrid (VPU-UAM) considering the framework of the international competition ICPR - HARL 2012 for human-related events. First, the related state of art and the available prototype are studied. Then, feature extraction is explored for the task of skin detection where different methods are proposed and implemented. Later, a deterministic method for event detection based on finite-state-machines (FSMs) is developed for the competition events (person-object interactions and activities for two persons) by analyzing, among others, skin regions of persons. Finally, such improvements are included in the initial prototype, which is evaluated on available datasets and the results are presented to the international competition ICPR - HARL 2012.