HAVi Newsletters

#09 – December 2017

TEC2014-53176-R HAVideo (2015-2017)

High Availability Video Analysis for People Behaviour Understanding

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

Ninth four-months period progress report

The project has been working on plan and five deliverables have been published on schedule:
- D2.1 "People Behaviour understanding in single and multiple camera settings" version 3 (December 2017)
- D2.4 "Exploration and viability studies for people behaviour understanding" version 2 (December 2017)
- D3.1 "Online adaptive people behaviour understanding based on contextual and quality information" version 2 (December 2017)
- D3.2 "Collaborative approaches for people behaviour understanding" version 2 (December 2017)
- D4.2.2 "Applications" (December 2017)
- D4.3 "Results Report" version 6 (December 2017)

All the main objectives of the project have been reached, but due to delays in WP4, we have asked the Ministry for an extension in October 2017 for working further on applications and simulator scenarios.

The main achievements of the different workpackages during these last months are described below. The complete achievements can be found compiled in the latest version of the Results Report, D4.3 "Results Report" version 6 (December 2017).

WP2 "Video analysis tools, models and performance indicators"

The main achievements of WP2 are:
- In the area of People Behaviour understanding, several algorithms have been developed for Automatic classification of videos using features descriptors, Traffic Signs Detection and Recognition, People Detection using Convolutional

Neural Networks, and Object-Background Segmentation; also, a Platform for People Detection Algorithms Evaluation has been developed.
- In the area of Contextual Modelling, approaches for Scene-based automatic initialization of region seeds for skin lesion segregation and characterization, and for Automatic aggregation and refinement of contextual information in scenarios recorded with multiple moving cameras, have been developed.
- With respect to Quality Analysis, a Background Estimation algorithm for video sequences based on the temporal median has been developed.
- Continuation of the visual materials identification work.
- Creation of a prototype of augmented reality.

WP3 "Self-configurable approaches for long-term analysis"

For WP3, the main achievements are:
- Development of Automatic Adaptation of People Detection to the Scene.
- Proposal for Detection Quality estimators for multi-camera setups
- Development of a Probabilistic multi-target multi-camera tracking algorithm
- Improvement of the use of scene context to inhibit and improve people detection
- Adaptation of people detection thresholds during prediction time
- Collaborative multi-camera location estimation.
- Collaborative multi-camera analysis based on quality: a survey
- Guiding video object segmentation with external results
- Improvement of background subtraction algorithms based on quality information

WP4 "Evaluation framework, demonstrators and dissemination"

For this workpackage, the main achievements are the implementation of the following applications/demonstrators:
- A multi-camera pedestrian detector with semantic constraining demonstrator
- A complete abandoned object detection (AOD) system demonstrator
- A long-term tracking demonstrator

Ninth four-months period results

Journals

Abstract: Foreground segmentation is a key stage in multiple computer vision applications, where existing algorithms are commonly evaluated making use of ground-truth data. Reference-free or stand-alone evaluations that estimate segmented foreground quality are an alternative methodology to overcome the limitations inherent to ground-truth based evaluations. In this work, we survey and explore existing stand- alone measures proposed in related research areas to determine good object properties for estimating the segmentation quality in background subtraction algorithms. We propose a new taxonomy for stand-alone evaluation measures and analyze 21 proposals. We demonstrate the utility of the selected measures to evaluate the segmentation masks of eight background subtraction algorithms. The experiments are per- formed over a large heterogeneous dataset with varied challenges (CDNET2014) and identify which properties of the measures are the most effective to estimate quality. The experiments also demonstrate that qualitative performance levels can be distinguished, and background subtraction algorithms can be ranked without the need of ground-truth.
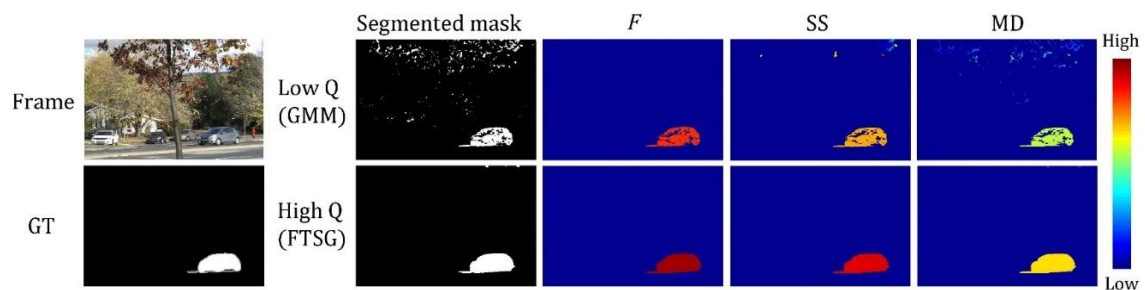
Abstract: This paper describes a region-based strategy for part-based object identification with independence of the external factors that affect its captured image: light variations, capture point-of-view or occlusions. Starting from color images and depth estimations, i.e. not requiring 3-dimensional models, we focus on the identification of learned objects in severe-occlusion scenarios. To face this problem, we assume that objects have been preliminarily segregated from the scene. Strong changes of appearance—due to one or several of the aforementioned factors or to the object nature, e.g. deformable objects—substantially increase the problem complexity. The proposed algorithm operates by splitting segregated objects in successively coarser region-partitions, with each region representing a part of the object from which it was extracted. For the characterization of these parts, two region-driven descriptors are proposed: R-DAISY and R-SHOT. Their novelty relies on the use of a size-and-shape variable description support which is automatically defined by the object part itself. Descriptions obtained in this way are self-organized in a single neural structure by an unsupervised learning process. Experimental results are promising in the identification of severe-occluded objects using a small set of training instances—1-to-8 short-varied views per object.
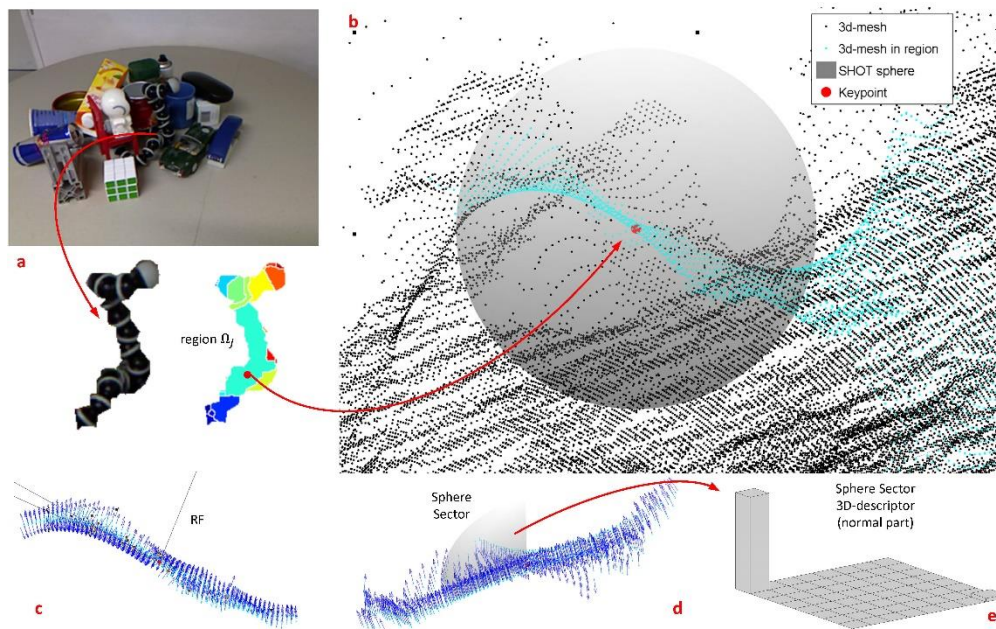
Conferences

Rafael Martín Nieto, Jesús Molina Merchán, Álvaro García-Martín, José M. Martínez, **"Generation and evaluation of synthetic models for training people detectors"**, Proc. of 2007 International Carnahan Conference on Security Technology – ICCST 2017, Madrid, Oct. 2017.

Abstract: There is a large demand in the area of video-surveillance, especially in people detection, which has caused a large increase in the number of researches and resources in this field. As training images and annotations are not always available, it is important to consider the cost involved in creating the detector models. For example, for elderly people detection, the detector must have into account different positions such as standing, sitting, in a wheelchair, etc. Therefore, this work has the main objective of reducing the amount of resources needed to generate the detection model, saving the cost of having to record new sequences and generate the associated annotations for a detector training. To achieve this, three synthetic image datasets have been created in order to train three different models, evaluating which model is optimal and finally analyzing its feasibility by comparing it with a people detector for wheelchair users trained with real images. Other people detection scenarios in which this technique could be applied are, for example, people riding horses or motorbikes, or people carrying supermarket carts. The synthetic datasets have been generated by combining images of standing people with wheelchair images, combining image patches, and segmenting sections of people (trunk, legs, etc.) to add them to the wheelchair image. As expected, the obtained results have a reduction of efficiency (between 21 and 25%) in exchange for the enormous saving in human annotation and resources to record real images.

Juan Carlos Sanmiguel, José M. Martínez, Luis Caro-Campos, **"Object-size invariant anomaly detection in video-surveillance"**, Proc. of 2007 International Carnahan Conference on Security Technology – ICCST 2017, Madrid, Oct. 2017.

Abstract: Nowadays, there is a growing demand for automated video-based surveillance systems due to increase security concerns. Anomaly detection is a popular application in this area where anomalous events of interest are defined as observed behavior that stands out from its context in space and time. In this paper, we present an approach for the detection of anomalous motion based on the extraction of object-size features that is independent of object size and video resolution. The proposed approach relies on a variable spatial window based on

object size that has shown robustness in scenarios that present motion of objects of different sizes. We propose a system composed of four building blocks: background subtraction, feature extraction, event modeling and outlier detection. The proposed approach is evaluated on publicly available datasets

Sample frame of test case
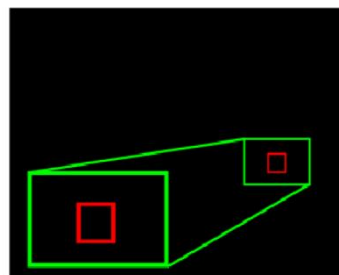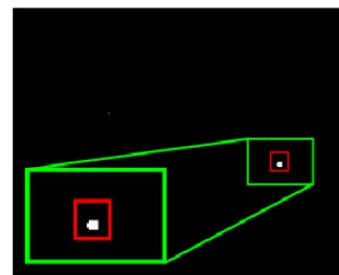
Related work   Proposed approach

which contain instances of abandoned objects of different sizes (considered as anomalies). The experiments carried out demonstrate that our approach outperforms the related state-of-the-art in the selected datasets. The proposal can identify anomalies associated to objects with different sizes and motion without increasing the number of false positives.

Adrián Tomé, Luis Salgado, "Detection of anomalies in surveillance scenarios using mixture models", Proc. of 2007 International Carnahan Conference on Security Technology – ICCST 2017, Madrid, Oct. 2017.

Abstract: In this paper we present a robust and simple method for the detection of anomalies in surveillance scenarios. We use a "bottom-up" approach that avoids any object tracking, making the system suitable for anomaly detection in crowds. A robust optical flow method is used for the extraction of accurate spatio-temporal motion information, which allows to get simple but discriminative descriptors that are employed to train a Gaussian mixture model.

We evaluate our system in a publicly available dataset, concluding that our method outperforms similar anomaly detection approaches but with a simpler model and low-sized descriptors.
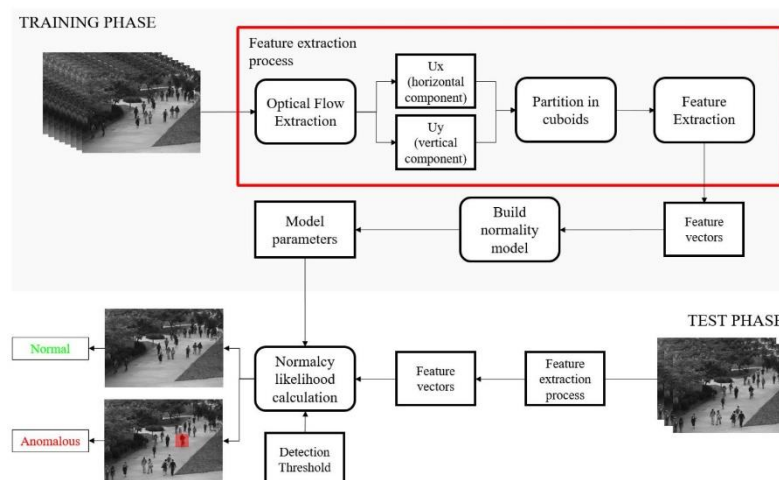
Master thesis

Desarrollo de un Toolbox para Posicionamiento 2D de cámaras (**Development of a Toolbox for cameras 2D positioning**), Sandra Gaytán Grande, (advisor: Juan Carlos San Miguel Avedillo), Trabajo Fin de Máster (Master Thesis), Master en Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2017.

Abstract: In recent years, video-surveillance systems, and more specifically camera positioning systems, are now being deployed widely thanks to their ability. In this context, optimal camera configuration will reduce the total number of cameras used achieving the same or greater level of utility, as well as better results and lower cost associated with future modifications. The objective of this project is the development of a Matlab toolbox which allows the calculation of the camera's position according to the constraints associated to the problem's formulation: continuous or discrete domain. The toolbox makes use of an interface that allows the user to select the floor map that is to be analyzed, the coverage's restrictions and the type of optimization to be used. Additionally, a graphical interface functionality has been included which allows the user to insert objects on the floor map by directly clicking on the interface, as well as fixing their height relative to the ceiling's. The final result will allow the user to visually examine the location of the cameras and the field of view.