TEC2014-53176-R HAVideo (2015-2017)

High Availability Video Analysis for People Behaviour Understanding

D4.2.2 v1

Applications

Video Processing and Understanding Lab

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Supported by

AUTHORS LIST

Álvaro García Martín (agm)	alvaro.garcia@uam.es

HISTORY

Version	Date	Editor	Description
0.1	10 November 2017	Álvaro García Martín	Initial draft version
0.9	10 December 2017	Álvaro García Martín	Final Working Draft
1.0	12 December 2017	José M. Martínez	Editorial checking

CONTENTS:

1. Introduction

1.1. Motivation

The work package 4 (WP4) aims at evaluating and integrating the algorithms developed within WP1, WP2 and WP3, in order to conform the global analysis chain to provide solutions to long-term video analysis for people behaviour understanding. In particular, this deliverable describes the work related with the task T.4.2: Use Cases and Demonstrators.

The objective of this task is the development of demonstrators of the algorithms developed in the project, providing applications both for developers (for evaluation and testing) as well as for final users (use cases to be defined with the help of the Observing Partners). The use cases will focus on the applications areas of the surveillance and people monitoring challenges targeted by the project (e.g, outdoor surveillance, people monitoring in malls, in-home monitoring).

1.2. Document structure

This document contains the following chapters:

- Chapter 1: Introduction to this document
- Chapter 2: Planning of first set of applications
- Chapter 3: Conclusions

2. Planning of first set of applications

2.1. A complete abandoned object detection (AOD) system demonstrator

A graphical user interface (GUI) has been developed to act as a demonstrator of the complete abandoned object detection (AOD) system. This GUI allows the user to check the system functionality in a visual way and to manually set the system algorithms and parameters [19].

The development environment used for the interface creation has been Qt Creator (version 4.2.1) based on Qt 5.8.0. An example of the working interface is shown in the picture below (see **Figure 1**):

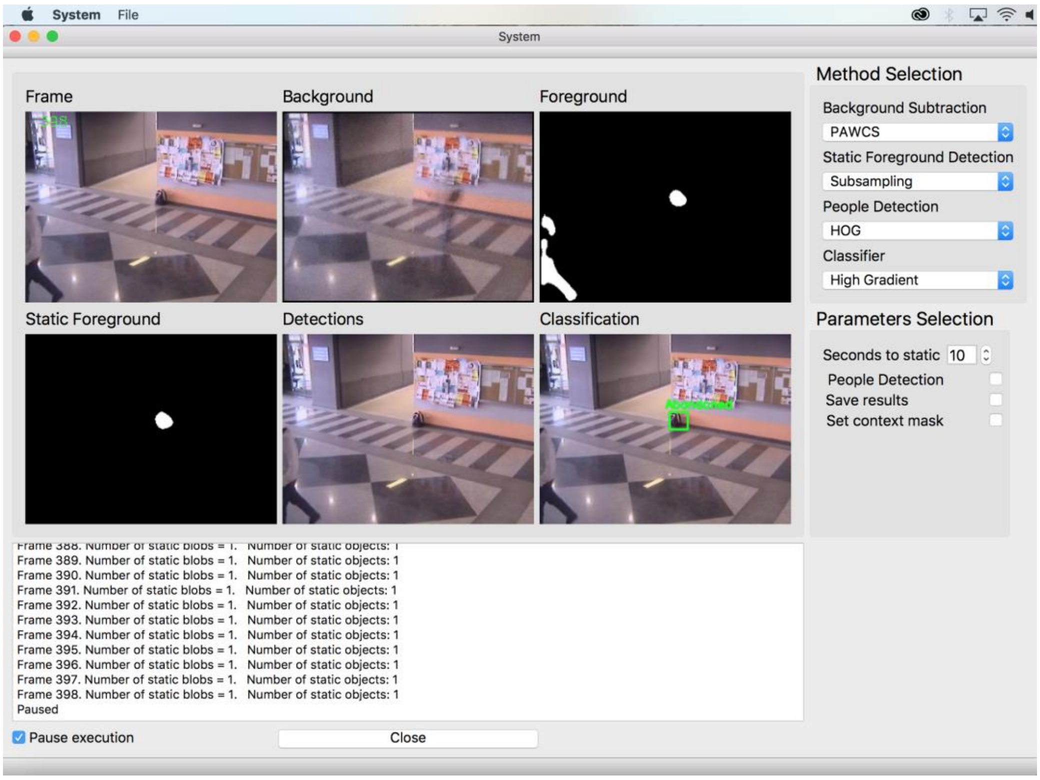

Figure 1. Graphical user interface or demonstrator interface.

1. **GUI elements.**

- Menu bar. This bar is placed on the top of the screen and the “file” option allows the user to select the desired vile file to process from the file explorer. Several formats (*.mpg, *.avi, *.mov, *.mp4, ...) are supported.

- Inside the main window of the interface the user can find:
    - Display area. This area allows the user to visualize the results of the system after each stage thereof. On top row from left to right, the current frame, the computed background and the foreground mask are displayed. On bottom row, the static foreground mask is displayed first, the people and object detections and finally the stolen/abandoned object classification.
    - Algorithm selection. Allows the user to choose a different algorithm for each module of the system by selecting it in a drop-down menu. Implemented algorithms at each stage of the system are listed below.
- Parameters selection:
    - Seconds to static. This option allows the user to modify the number of seconds until an object is considered as stationary. It is set to 10 seconds by default.
    - People detection. If this checkbox is activated, the people detection is running throughout the full sequence, otherwise it only will be run when something static is detected.
    - Save results. If it is checked a .xml file with the detections found along the sequence will be created and saved.
    - Set context mask. This option allows the user to select an area of the frame where the detections will be ignored (non-interest area).
- A dialog box where results are displayed in real time.
- Pause and close execution options.

2. Current functionalities.

- Reading stored video sequences in several video formats (*.mpg, *.avi, *.mov, …).
- Algorithms selection for each system stage.
- Parameters selection.
- Displaying results after each system stage.

- Saving results in .xml files.

3. **Implemented algorithms.**

   a. Background subtraction module

[1] Mixture Of Gaussians (MOG) background subtraction models the background by using several Gaussian distributions at each pixel location.

[2] K-Nearest Neighbors background subtraction is an improved version of MOG.

[3] Kernel Density Estimation (KDE) is a non-parametric modeling method that estimates the density function directly from the data.

[4] Independent Multimodal Background Subtraction (IMBS) algorithm is based on the discretization of the color distribution of each pixel by applying a grouping algorithm to generate the background model.

[5] LOcal Binary Similarity segmenTER (LOBSTER) based on Local Binary Similarity Pattern (LBSP) features and color information.

[6] Pixel- based Adaptive Word Consensus Segmenter (PAWCS) is an improved version of LOBSTER and based also on background word consensus.

[7] Self-Balanced SENsitivity Segmenter (SUBSENSE) is on the same basis than PAWCS, but in addition it includes automatic adjustments of local sensitivity.

   b. Static foreground detection module

[8] Temporal accumulation of the pixel's persistence.

[9] Subsampling approach also analyzes persistence via foreground subsampling.

[10] History Images is a multi-feature detector combining foreground, motion and structural information.

[11] Dual Background Model computes static foreground by comparing two background models (short and long-term)

[12] Triple Background Model computes static foreground by comparing three background models (short, medium and long-term)

   c. People detection module

[13] Histogram of Oriented Gradients (HOG) applies exhaustive search based on holistic appearance descriptors along the whole image.

[14] Haar-like features classifier is based on a trained holistic person model.

[15] Deformable Part-based Model (DPM) is a part-based person model.

[16] Aggregated Channel Features (ACF) is a detector also based on exhaustive search and a holistic model.

d. Abandoned object classification module

[17] High Gradient (HG) and Color Histogram (CH) approaches. The former analyzes high-gradient value points along the object shape while the second only considers color information.

[18] Pixel Color Contrast (PCC) method combines edge and color information.

2.2. A multi-camera pedestrian detector with semantic constraining demonstrator

The main functionality of the developed application is to facilitate the use and configuration to the final user of the different algorithms and strategies implemented in the final application. This application aims to implement a multi-camera pedestrian detector with semantic constraining. Two main interfaces have been created depending on the final user [20].

The developer interface version allows to:

First, visualize all the video sequence from the video cameras. These sequences may be loaded through the dedicated menu for opening video files, or cameras IP can be loaded so network streams are used in the case of IP cameras (see **Figure 2**).

In addition to the video sequence, the interface allows to visualize both intermediate results and outputs from the developed algorithms (pedestrian detection, semantic segmentation, usage rate extraction…) over the camera frames. This information is intended for the developer user.

Also, the interface represents the common reference plane with all the projected information that is used in the multi-camera system. In this plane one can observe projected pedestrian in addition to the desired semantic information. In the image it has been decided to project on to the reference plane those path-labelled areas by the semantic segmentation algorithm.

Finally, the developer version has a configuration area that allows the user to:

Choose the desired configuration to perform pedestrian detection task. Configuration includes both the algorithm and the online selection of the detection threshold. In the options one can choose between several algorithms such as Histogram of Oriented Gradients (HOG), Deformable Parts Models (DPM), Aggregate Channel Features (ACF) or Fast-RCNN. The interface also allows to enable or disable the semantic constraining or the multi-camera fusion algorithm.

From the interface one can modify the projection representation for pedestrian detections onto the reference plane. Detections may be represented as lines obtained from the bounding boxes or as Gaussian functions.

Furthermore, and if files have been previously loaded, the interface allows the user to visualize simultaneously ground-truth information.

Finally, there exists an information area from where the user gets all the necessary information to proceed if the interface is in idle state. When the application is working this area will display several information about the video files, the algorithms and the output paths.

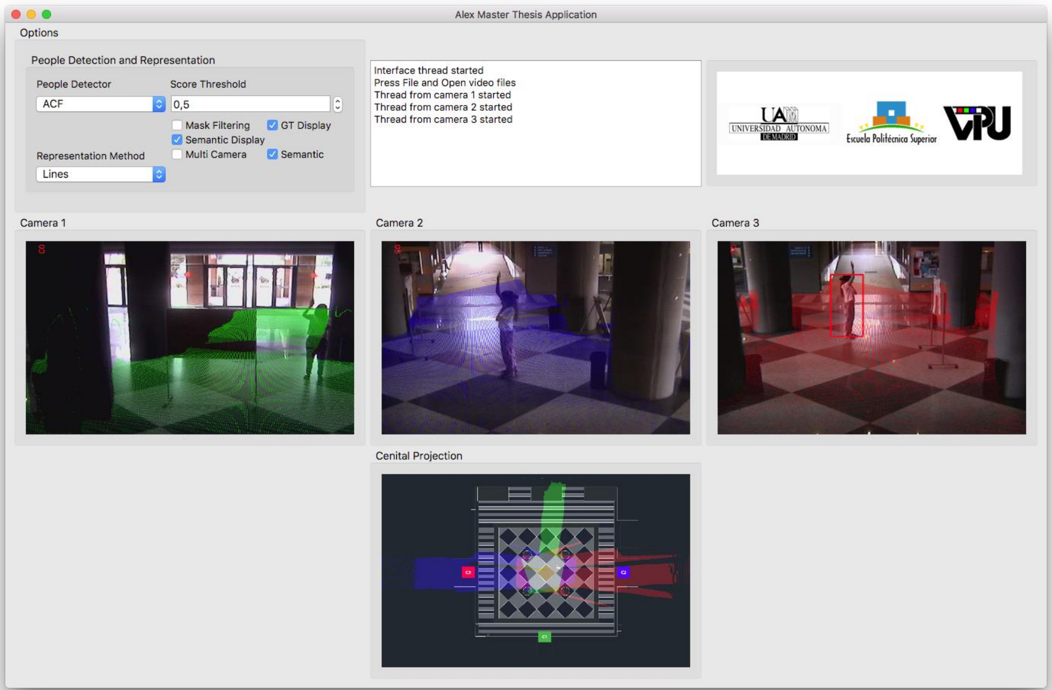

Figure 2. Developer demonstrator interface.

Client version:

This version of the interface is intended for the final user. It allows the user to run the program without any knowledge about the algorithms behind the application. Due to this reason, it only contains the display and the information area. The configuration area has disappeared, and all the tuneable (see **Figure 3**).

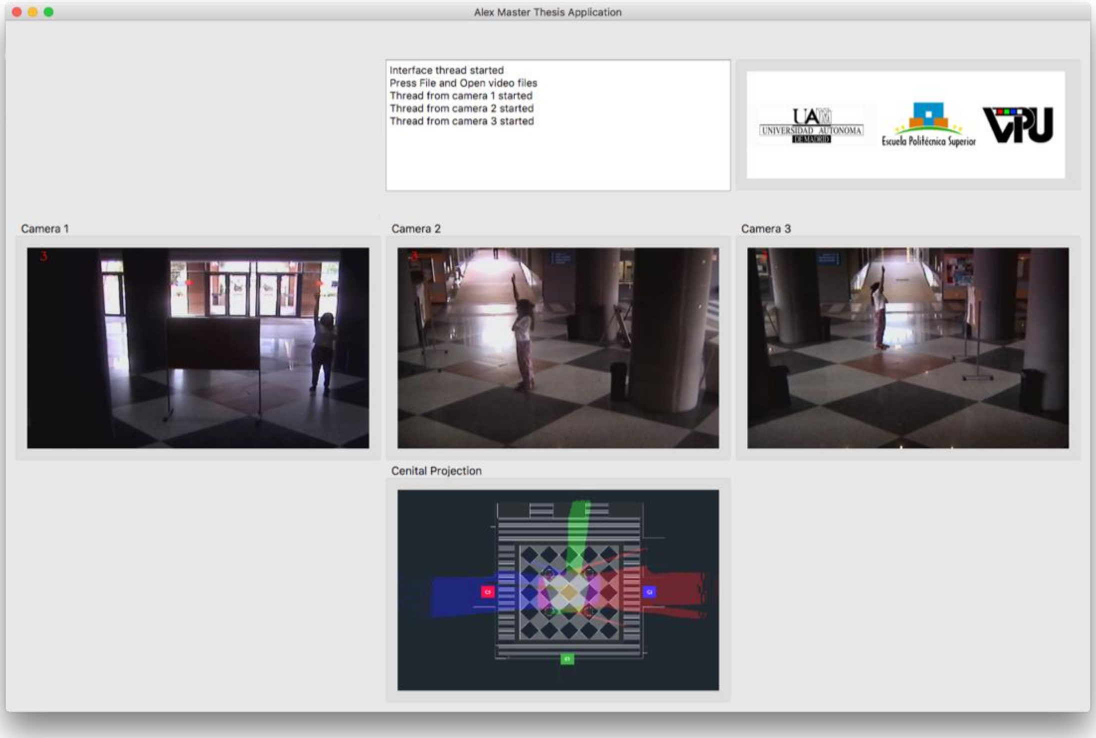

Figure 3. Client demonstrator interface.

2.3. A long-term tracking with target re-identification demonstrator

The objective of this work is to improve the performance of an existing tracker, called PKLTF (Point-based Kanade Lucas Tomasi colour-Filter). A newly improved tracker is designed considering the problems that affect the base tracker. Several improvements are tested, some of which are integrated into the proposed version SAPKLTF (Scale Adaptive Point-based Kanade Lucas Tomasi colour-Filter) 14[21].

A demonstrator application is developed in order to show the operation of the tracker in real situations, to facilitate the understanding of the algorithm and the influence of the parameters on the algorithm performance (see **Figure 4**).

In this case, the PKLTF demonstrator developed in a previous work [22] is used as starting point. This application has been updated with the SAPKLTF tracking algorithm.

The application gives a simple way to interact with the algorithm. Firstly, the user must choose the camera from which the application receives the video streaming; it can be a local camera or an IP camera. Once the camera, is selected the user can start running the demonstrator. To initialize the algorithm, the user must select the bounding box that defines the target with the mouse.

Through the configuration button, the user can change the parameters of the tracker, as well as different display options.

Figure 4. SAPKLTF demonstrator interface.

3. Conclusions

This deliverable updates the work related with the task T.4.2: Use Cases and Demonstrators. Following the guidelines for the development of applications and demonstrators in relation with the project. This document includes the description of three different demonstrator implemented during the last period of this project. [20] presents a multi-camera pedestrian detector with semantic constraining demonstrator. [19] proposes a complete abandoned object detection (AOD) system demonstrator. And **¡Error! No se encuentra el origen de la referencia.** describes a long-term tracking with target re-identification demonstrator.

4. References

[1] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in 17th International Conference on Pattern Recognition (ICPR 2004), vol. 2, pp. 28–31, 2004.

[2] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern recognition letters, vol. 27, no. 7, pp. 773–780, 2006

[3] Elhabian, S. Y., El-Sayed, K. M., & Ahmed, S. H. (2008). Moving object detection in spatial domain using background removal techniques-state-of-art. Recent patents on computer science, 1(1), 32-54.

[4] D. D. Bloisi, A. Grillo, A. Pennisi, L. Iocchi, and C. Passaretti, "Multi-modal background model initialization," in International Conference on Image Analysis and Processing, pp. 485–492, Springer, 2015

[5] P.-L. St-Charles and G.-A. Bilodeau, "Improving background subtraction using local binary similarity patterns," in IEEE Winter Conference on Applications of Computer Vision (WACV 2014), pp. 509–515, 2014.

[6] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "A self-adjusting approach to change detection based on background word consensus," in IEEE Winter Conference on Applications of Computer Vision (WACV 2015), pp. 990–997, 2015

[7] St-Charles, P. L., Bilodeau, G. A., & Bergevin, R. (2014). Flexible background subtraction with self-balanced local sensitivity. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 408-413).

[8] Guler, S., Silverstein, J. A., & Pushee, I. H. (2007). Stationary objects in multiple object tracking. In Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on (pp. 248-253).

[9] H.-H. Liao, J.-Y. Chang, and L.-G. Chen, "A localized approach to abandoned luggage detection with foreground-mask sampling," in IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance, AVSS 2008., pp. 132–139, 2008

[10] D. Ortego and J. C. SanMiguel, "Stationary foreground detection for video- surveillance based on foreground and motion history images," in IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2013, pp. 75–80, 2013

[11] Porikli, F., Ivanov, Y., & Haga, T. (2008). Robust abandoned object detection using dual foregrounds. EURASIP Journal on Advances in Signal Processing, 2008

[12] Cuevas, C., Martinez, R., Berjón, D., & Garcia, N. (2017). Detection of stationary foreground objects using multiple nonparametric background-foreground models on a finite state machine. IEEE Transactions on Image Processing, 26(3), 1127-1142

[13] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893)

[14] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-I)

[15] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. IEEE transactions on pattern analysis and machine intelligence, 32(9), 1627-1645

[16] Dollár, P., Appel, R., Belongie, S., & Perona, P. (2014). Fast feature pyramids for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(8), 1532-1545

[17] San Miguel, J. C., & Martínez, J. M. (2008, September). Robust unattended and stolen object detection by fusing simple algorithms. In Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on (pp. 18-25).

[18] Sanmiguel, J. C., Caro, L., & Martínez, J. M. (2012). Pixel-based colour contrast for abandoned and stolen object discrimination in video surveillance. Electronics letters, 48(2), 86-87.

[19] Abandoned Object Detectionin Long-Term Video-Surveillance, Elena Luna García, (advisor: Juan Carlos San Miguel Avedillo), Trabajo Fin de Máster (Master Thesis), Master en Investigación e Innovación en TIC – Programa Internacional de Múltiple Titulación IPCV (Image Processing and Computer Vision), Univ. Autónoma de Madrid, Jul. 2017.

[20] Online Contextual Updating in Multi-Camera Scenarios, Alejandro López Cifuentes, (advisor: Marcos Escudero Viñolo), Trabajo Fin de Máster (Master Thesis), Master en Investigación e Innovación en TIC – Programa Internacional de Múltiple Titulación IPCV (Image Processing and Computer Vision), Univ. Autónoma de Madrid, Jul. 2017.

[21] Long-Term Tracking with Target Re-Identification, Erik Velasco Salido, (advisor: José M. Martínez), Trabajo Fin de Máster (Master Thesis), Master en Investigación e Innovación en TIC – Programa Internacional de Múltiple Titulación IPCV (Image Processing and Computer Vision), Univ. Autónoma de Madrid, Jul. 2017.

[22] J. Sanjuan, “Seguimiento de objetos en tiempo real,” Bachelor Thesis, Univ. Autónoma de Madrid, 2014.