TEC2014-53176-R HAVideo (2015-2017)

High Availability Video Analysis for People Behaviour Understanding

D4.3 v1-2

Results Report

Video Processing and Understanding Lab

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Supported by

AUTHORS LIST

José M. Martínez	josem.martinez@uam.es
Fulgencio Navarro	fulgencio.navarro@uam.es
Diego Ortego	diego.ortego@uam.es
Rafael Martín	rafael.martin@uam.es
Álvaro Martín	alvaro.martin@uam.es
Juan Carlos San Miguel	juancarlos.sanmiguel@uam.es
Marcos Escudero	marcos.escudero@uam.es
Jesús Bescós	j.bescos@uam.es

HISTORY

Version	Date	Editor	Description
0.1	1 December 2015	José M. Martínez	First Draft version for contributions
0.2	2 December 2015	Fulgencio Navarro	Contributions
0.3	2 December 2015	Diego Ortego	Contributions
0.4	3 December 2015	Rafael Martin	Contributions
0.5	5 December 2015	Álvaro García	Contributions
0.6	10 December 2015	Juan Carlos San Miguel	Contributions
0.7	10 December 2015	Marcos Escudero	Contributions
0.8	12 December 2015	Jesús Bescós	Contributions
0.9	14 December 2015	José M. Martínez	Working Draft edition
0.10	15 December 2015	Álvaro García	Working Draft comments
0.11	16 December 2015	Juan Carlos San Miguel	Working Draft comments
0.12	17 December 2015	Marcos Escudero	Working Draft comments
0.13	18 December 2015	Diego Ortego	Working Draft comments
0.14	18 December 2015	Fulgencio Navarro	Working Draft comments
0.15	19 December 2015	Rafael Martín	Working Draft comments
0.16	22 December 2015	José M. Martínez	Final Working Draft
1.0	22 December 2015	José M. Martínez	Editorial checking

CONTENTS:

1. Introduction

This *report* recapitulates the results obtained within the HAVideo project. The results are announced as they are obtained at the Web site (http://www-vpu.eps.uam.es/HAvideo).

The HAVideo project started officially January 2015, nevertheless the official grant notification arrived in September. Nevertheless, Video Processing and Understanding Lab (VPULab) has been working during 2015 in the research lines proposed in the project. Therefore, the project started to produce results before the official announcement of the project start.
Nevertheless, the publication dates of all the deliverables of the project have been rescheduled to December 2015. This decision implied a delay, not affecting the project work planning, of deliverables D1.1, D1.2, D4.1 and D4.3v1 (edited together with v2 –this document-). Also the HAVi Newsletters of year 2015 were published together December 2015 (numbers 1-3).

1.1. Document structure

This document contains the following chapters:

- Chapter 1: Introduction to this document
- Chapter 2: Publications
- Chapter 3: Project Documents
- Chapter 4: Content Sets
- Chapter 5: Workshops and Seminars
- Chapter 6: Main achievements of the project

2. Publications

2.1. Journals

[1] Álvaro García, Juan Carlos San Miguel, **"Context-aware part-based people detection for video monitoring**", Electronic Letters, 51(23):1865:1867, Nov. 2015, IET, ISSN 0013-5194 (DOI 10.1049/el.2015.3099)

[2] Diego Ortego, Juan Carlos San Miguel, José M. Martínez, **"Long-Term Stationary Object Detection Based on Spatio-Temporal Change Detection**", IEEE Signal Processing Letters, 22(12):2368:2372, Dec. 2015, IEEE, ISSN 1070-9908 (DOI 10.1109/LSP.2015.2482598)

2.2. Book Chapters

2.3. Conferences

[3] Matej Kristan et al., "The Visual Object Tracking VOT2015 challenge results", Proc. of 3rd Visual Object Tracking Challenge Workshop at International Conference on Computer Vision, Santiago, Chile, December 2015, pp.564-586.

[4] Michael Felsberg et al., "The Visual Object Tracking VOT-TIR2015 challenge results", Proc. of 3rd Visual Object Tracking Challenge Workshop at International Conference on Computer Vision, Santiago, Chile, December 2015, pp.639-651.

2.4. PhD Thesis

2.5. Master Thesis

[5] **Estimación de la densidad de personas en entornos densamente poblados (People density estimation in crowded environments)**, Rosely Sánchez (advisor: Álvaro García-Martín), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Escuela Politécnica Superior, Univ. Autónoma de Madrid, May 2015.

[6] **Detección de personas en presencia de grupos (People detection in presence of groups)**, Sergio Merino Martínez, (advisor: Álvaro García Martín), Trabajo Fin de Máster (Master Thesis), Master en Investigación e Innovación en TIC, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Sep. 2015.

[7] **Preservación de privacidad de personas en vídeo-seguridad (People privacy preservation in video-surveillance)**, Jaime Mateo Herrero (advisor: José M. Martínez), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Universidad Autónoma de Madrid, Escuela Politécnica Superior, Oct. 2015.

2.6. Graduate Thesis

[8] **Detección de ritmo cardíaco mediante vídeo (Heart rate detection using video)**, Erik Velasco Salido (advisor: José M. Martínez), Trabajo Fin de Grado

(Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jun. 2015.

[9] **Identificación automática de materiales usando el sensor Kinect (Material identification through the Kinect technology)**, Alejandro López Cifuentes (advisor: Marcos Escudero Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jun. 2015.

[10] **Detección de la posición relativa de una cámara en situaciones de grandes desplazamientos (Detection of camera relative position in wide-baseline scenarios)**, María Narváez Encinal (advisor: Marcos Escudero Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jun. 2015.

[11] **Detección de vehículos mediante el análisis de imágenes (Image-based vehicle detection)**, Juan Ignacio Bravo (advisor: Luis Salgado), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jun. 2015.

[12] **Simulador virtual para sistemas multi-cámara distribuidos (Distributed multicamera systems virtual simualtor),** Luis Pérez Llorente (advisor: Juan C. San Miguel), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jul. 2015.

[13] **Desarrollo de herramienta para la anotación manual de secuencias de vídeo (Developemt of a video sequences manual annotation tool),** Yoel Witmaar (advisor: Juan C. San Miguel)), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jul. 2015.

[14] **Detección de Personas en Grupos (People detection in groups),** Marta Villanueva Torres (advisor: Álvaro García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jul. 2015.

3. Project Documents

3.1. Deliverables

3.2. First year Deliverables

3.2.1. D1.1 "System Infrastructure" version 1

This deliverable describes the current infrastructure and purchase plans for the second year, as due to budget constraints and grant notification delay, the originally planned acquisitions have been postponed to the beginning of the second year. This delayed purchase allows to define priorities over the acquisitions in the proposal by discussing the needs with the package leaders. As a result, additional equipment is going to be purchased: a new server with GPU support and laptops for the research team. The acquisition of battery-powered cameras will be considered during the third year subject to budget availability. The second year report will contain the planned milestones and deliverables corresponding to the acquisition of portable equipment (to be performed in the third year).

3.2.2. D1.2 "DiVA documentation" version 1

This deliverable describes the activities performed for the task T.1.2: Update and maintenance of the Distributed Video Analysis (DiVA) software architecture. During the first year, an existing annotation tool has been improved to provide semiautomatic information which will be integrated in DiVA. Due to the low priority of this task (no dependencies exist during the first year), project resources have been reassigned to more critical tasks and therefore, the completion of the developments for the contextual information management have been postponed to the second and third year, where integration of new sensors, reliability testing and improvements of the annotation tools will be explored. For the second year, the deliverable D1.2 will be updated with the planned improvements of the employed annotation tool in the project.

3.2.3. D1.3 "Simulator documentation: description, programmer manual and user guide" version 1

This deliverable describes the activities performed for the task T.1.3: Development and maintenance of a camera network simulator. During the first year, two multi-camera simulators have been developed for virtual data and for network conditions. The former allows to repeatedly visualize the same activities with a network of cameras whereas the latter provides features to simulate real networks (communication protocols, wireless/wired channels, transmission delays/errors,…). Proper documentation has been generated for both simulators. For the second year, the objective is twofold: improve the virtual simulator and integrate with the network simulator.

3.2.4. D2.1 “People Behaviour understanding in single and multiple camera settings” version 1

This deliverable describes the work related with the task T.2.1 Analysis Tools for human behavior understanding. The people behavior understanding in this project has been already designed as a sequential combination of object segmentation, people detection, object tracking and behavior recognition. In particular, during the first part of the project there have been a focus on developing different approaches for segmentation, people detection and tracking in single camera.

In relation with segmentation, a long-term stationary object detection based on spatio-temporal change detection has been implemented and evaluated, as a result a publication in an international journal has been presented ([2]).

In relation with people detection several approaches have been implemented and tested in relation with people behavior understanding: people density estimation in crowded scenarios and people detection in groups (Master theses [5][6] and Graduate Thesis [14]). There is also an ongoing work related with integrating wheelchair users in people detection.

In relation with tracking, we have presented a tracker to an international tracking challenge (VOT challenge), as a result two publications[3][4] in an international conference have been presented. In addition, there is an ongoing work on high-precision shape tracking.

3.2.5. D2.4 "Exploration and viability studies for people behaviour understanding” version 1

This deliverable describes the work related with the task T2.4, exploration and viability studies. Three main works has been developed and are presented below.

The first work is related to heart rate detection using video, and the developed system is explained in graduate thesis [8].

The second work focuses on people privacy preservation in video-surveillance. This work is contained in master thesis [7].

The third work combines colour, texture and depth information to identify materials in a Kinect-like controlled scenario. This work is contained in master thesis [9].

3.2.6. D4.1 "Evaluation methodology and datasets” version 1

In relation with people detection, a new people detection dataset have been recorded and evaluated, in particular, multiple camera video sequences has been recorded in a real in-door senior residence environment containing wheelchairs users and standing people and it has been released together with the associated ground-truth.

In relation with tracking, we have presented a tracker to an international tracking challenge (VOT challenge). The VOT challenges provide the visual tracking community with a precisely defined and repeatable way of comparing trackers as well as a common platform for discussing the evaluation and advancements made in the field of visual tracking. This deliverable includes a description of the proposed methodology and dataset presented in the VOT challenge.

3.2.7. D4.3 “Results Report” version 1 (June 2015)

This deliverable was delayed and merged with December’s version (this document).

3.2.8. D4.3 "Results Report" version 2 (December 2015)

This *report* recapitulates the results obtained within the HAVideo project during 2015. The results are announced as they are obtained at the Web site (http://www-vpu.eps.uam.es/HAvideo).

3.3. Technical Reports

4. Content Sets

4.1.1. Wheelchair Users dataset - WUds

http://www-vpu.eps.uam.es/DS/WUds/

5. Workshops and Seminars

5.1. Workshops

5.1.1. 1st Workshop

The first workshop was held May 2015 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid.
It was bot announced as HAVideo Workshops as the project proposal approval notification was not received at that time.
The Workshop consisted of two parts:

- Dissemination Day (May 25th 2015): Video Analysis Technologies and applications for Computer Vision.
- Short course: (May 26th-28th 2015): Introduction to Computer Vision applications programming with OpenCV

5.2. Seminars

6. Main Achievements of the Project

6.1. First year main achievements

For the workpackage WP.1: Video Analysis Framework, the main achievements are:

- Development of a virtual camera network simulator
- Improvement of an existing camera network simulator to include embedded hardware modelling, sensor models and real communication protocols
- Improvement of an existing manual annotation tool, to provide semiautomatic tools that speed up the annotation process.

For the workpackage WP.2: Video analysis tools, models and performance indicators, the main achievements are:

- Development of a long-term stationary object detection based on spatio-temporal change detection [2].
- Development of a people density estimation in crowded scenarios and a people detection in groups [5].
- Development of a context-aware people detector [1] which uses static contextual annotation to integrate and weight partial evidences from a part-based people detector algorithm.
- Development of a camera calibration algorithm [10] to detect the relative position of a camera respect to another in wide-baseline scenarios.
- Development of a vehicle detection algorithm [11] that uses on-line generated context to separate between object classes.
- Development of a tool and of its associated user interface to annotate static contextual objects in video sequences [13].
- Development of a heart rate detector using video [8].
- Implementation of a privacy preserving module for existing video surveillance systems [7].
- Development of a system for material identification through the Kinect technology [9].

For the workpackage WP.4: Evaluation framework, demonstrators and dissemination, the main achievements are:

- International tracking challenge (VOT challenge) evaluation [3][4].
- Development of a new dataset for wheelchairs users and standing people detection in a real in-door senior residence environment.
- The web page has been created

The first workshop was held May 2015 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid (see 5.1.1)