TEC2014-53176-R HAVideo (2015-2017)

High Availability Video Analysis for People Behaviour Understanding

D4.3 v6

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.martinn@uam.es
Álvaro Martín	alvaro.martin@uam.es
Juan Carlos San Miguel	juancarlos.sanmiguel@uam.es
Marcos Escudero Viñolo	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
1.1	13 June 2016	José M. Martínez	First Draft version for contributions (version 3)

1.2	15 June 2016	Juan Carlos San Miguel	Contributions
1.3	16 June 2016	Diego Ortego	Contributions
1.4	18 June 2016	Álvaro García Martín	Contributions
1.5	18 June 2016	Rafael Martín	Contributions
1.6	19 June 2016	Marcos Escudero	Contributions
1.7	20 June 2016	Fulgencio Navarro	Contributions
1.9	21 June 2016	José M: Martínez	Working Draft edition
1.10	22 June 2016	Álvaro García Martín	Working Draft comments
1.11	23 June 2016	Juan Carlos San Miguel	Working Draft comments
1.12	24 June 2016	Rafael Martín	Working Draft comments
1.13	25 June 2016	Marcos Escudero	Working Draft comments
1.14	27 June 2016	Diego Ortego	Working Draft comments
1.15	27 June 2016	Jesús Bescós	Working Draft comments
1.16	29 June 2016	José M. Martínez	Final Working Draft
2.0	30 June 2016	José M. Martínez	Editorial checking
2.1	10 December 2016	José M. Martínez	First Draft version for contributions (version 4)
2.2	11 December 2016	Juan Carlos San Miguel	Contributions
2.3	12 December 2016	Diego Ortego	Contributions
2.4	15 December 2016	Álvaro García Martín	Contributions
2.5	18 December 2016	Rafael Martín	Contributions
2.6	21 December 2016	Marcos Escudero	Contributions
2.7	22 December 2016	Fulgencio Navarro	Contributions
2.9	27 December 2016	José M: Martínez	Working Draft edition
2.10	28 December 2016	Álvaro García Martín	Working Draft comments
3.0	30 December 2016	José M. Martínez	Editorial checking
3.1	1 June 2017	José M. Martínez	First Draft version for contributions (version 5)
3.2	7 June 2017	Juan Carlos San Miguel	Contributions
3.3	9 June 2017	Diego Ortego	Contributions
3.4	12 June 2017	Álvaro García Martín	Contributions
3.5	13 June 2017	Rafael Martín	Contributions
3.6	16 June 2017	Marcos Escudero	Contributions
3.7	20 June 2017	Fulgencio Navarro	Contributions
3.9	24 June 2017	José M: Martínez	Working Draft edition
3.10	26 June 2017	Álvaro García Martín	Working Draft comments

4.0	28 June 2017	José M. Martínez	Editorial checking
4.1	1 December 2017	José M. Martínez	First Draft version for contributions (version 6)
4.2	5 December 2017	Álvaro García Martín	Contributions
4.3	10 December 2017	Diego Ortego	Contributions
4.4	12 December 2017	Juan Carlos San Miguel	Contributions
4.5	13 December 2017	Marcos Escudero	Contributions
4.6	16 December 2017	Rafael Martín	Contributions
4.8	19 December 2017	José M. Martínez	Working Draft edition
4.9	20 December 2017	Diego Ortego	Working Draft comments
4.10	21 December 2017	Álvaro García Martín	Working Draft comments
5.0	28 December 2017	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 both 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).

During the third semester the workplan has been slightly modified:

- January 2016 we decided to delay the start of Tasks 3.1 (till June 2016, to synchronize WP3 tasks) and 4.2 (till January 2017 to focus on applications during the last year).
- June 2016 we have decided to extend Task 2.2 till month 30 (June 2017) in order to further work on context in multi-camera setups.

HAVi Newsletter #4 was published April 2016. With respect to deliverables, D2.2v1 and D2.3v1 were published June 2016, as well as D4.3v3 (this document). D3.1v1 was cancelled as a consequence of the delay in the start of Tasks 3.1; also, accordingly to the delay in the start of Tasks 4.2, all the related deliverables of this task planned for 2016 have been rescheduled.

At the start of the fourth semester, we have decided to update the workplan delaying some milestones and deliverables. With respect to deliverables, the following were rescheduled for March 2017:

- D.2.2 “Contextual modelling and extraction for people behavior understanding” version 2
- D3.1 “Online adaptive people behaviour understanding based on contextual and quality information” version 1
- D.3.2 “Collaborative approaches for people behaviour understanding” version 1

The 2016 Dissemination Workshop was organized for being held November 25th 2016 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid. Due to the reduced number of registered companies, the Workshop was cancelled.
HAVi Newsletter #5 and #6 were published on schedule. With respect to deliverables, D2.1v2 and D4.1v2 were published December 2016, as well as D4.3v4 (this document).

During the fifth semester, once again Task 4.2 has been delayed, expecting to deliver D4.2.2. “Applications” at the end of the project, reporting on the different applications

and demonstrators developed. Deliverables D3.1v1 and D3.2v2 have been also rescheduled to December 2017.
In March 2017, deliverables D2.2v2, D3.1v1, D3.2v1 and D4.2.1v1 where published, followed by the publication of D1.1v2, D1.3v2, D2.3v2 and D4.3v5 (this document) in June 2017. HAVi Newsletter #7 was published on schedule.
The 2017 Developers Workshop was held from May 30th to June 2nd 2017 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid: there were 18 participants in the short course and 26 in the Dissemination Day.

During the sixth semester, the project workplan has been completed, publishing the pending deliverables: D2.1v3, D2.4v2, D3.1v2, D3.2v2, D4.1v2, D4.2.2 and D4.3v6 (this document). HAVi Newsletter #8 and #9 were published on schedule.

We have asked for an extension in October 2017 for working further on applications and simulator scenarios.

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)

[3] ObaidUllah Khalid, Juan C. SanMiguel, Andrea Cavallaro, **"Multi-tracker Partition Fusion"**, IEEE Transactions on Circuits and Systems for Video Technology, (online March 2016), IEEE, ISSN 1051-8215 (DOI 10.1109/TCSVT.2016.2542699)

[4] Diego Ortego, Juan Carlos San Miguel, José M. Martínez, **"Rejection based Multipath Reconstruction for Background estimation in Video Sequences with Stationary Objects"**, Computer Vision and Image Understanding, 147:23-37, Jun. 2016, Elsevier, ISSN 1077-3142 (DOI 10.1016/j.cviu.2016.03.012).

[5] Víctor Fernández-Carbajales, Miguel Ángel García, José M. Martínez, **"Visual Attention Based on a Joint Perceptual Space of Color and Brightness for Improved Video Tracking"**, Pattern Recognition, 60:571-584, Dec. 2016 (online June 2016), Elsevier, ISSN 0031-3203 (DOI 10.1016/j.patcog.2016.06.007).

[6] Juan C. SanMiguel, Andrea Cavallaro, "Energy Consumption Models for Smart-Camera Networks", IEEE Transactions on Circuits and Systems for Video Technology, (online September 2016), IEEE, ISSN 1051-8215 (DOI 10.1109/TCSVT.2016.2593598)

[7] Álvaro García-Martín, Ricardo Sánchez, José M. Martínez, "Hierarchical detection of persons in groups", Signal, Image and Video Processing, 11(7):1181-1188, Oct. 2017, (first online Feb. 2017), Springer, ISSN 1863-1703 (Print) 1863-1711 (Online) (DOI 10.1007/s11760-017-1073-z)

[8] Juan C. SanMiguel, Andrea Cavallaro, "Networked computer vision: the importance of a holistic simulator", IEEE Computer, 50(7):35-43, Jul. 2017, IEEE, ISSN 0018-9162 (DOI 10.1109/MC.2017.213)

[9] Diego Ortego, Juan C. SanMiguel, José M. Martínez, "Stand-alone quality estimation of background subtraction algorithms", Computer Vision and Image Understanding, 162:87-102, Sept. 2017, Elsevier, ISSN 1077-3142, (DOI 10.1016/j.cviu.2017.08.005)

[10] Marcos Escudero-Viñolo, Jesús Bescós, "Severe-occluded 3D object identification via region-based descriptions", Signal Processing: Image Communication, 58(1):240-257, Oct. 2017, Elsevier, ISSN 0923-5965 (DOI 10.1016/j.image.2017.07.007

2.2. Book Chapters

[11] Juan Carlos San Miguel, Mónica Lozano, José M. Martínez, "Performance Evaluation of Single Object Visual Tracking: Methodology, Dataset and Experiments", cáp. 3, pp. 107-142, Surveillance Systems: Design, Applications and Technology, R. Simmons (ed.), 2017, Nova Publishers. (ISBN 978-1-53610-703-6) (ISBN 978-1-53610-726-5 eBook)

2.3. Conferences

[12] Matej Kristan et al., **"The Visual Object Tracking VOT2015 challenge results"**, Proc. of 2015 IEEE International Conference on Computer Vision, Santiago, Chile, Dec. 2015, pp.564-586 (DOI 10.1109/ICCVW.2015.79).

[13] 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, Dec. 2015, pp.639-651 (DOI 10.1109/ICCVW.2015.86).

[14] Álvaro García-Martin, Rosely Sánchez Ricardo, Jose M. Martinez: "Estimación densidad de personas basada en segmentación persona-fondo (People density estimation based on people-background segmentation)", Actas del XXXI Simposium Nacional de la Unión Científica Internacional de Radio - URSI 2016, Madrid, Spain, Sept. 2016.

[15] Sergio López, Diego Ortego, Juan Carlos Sanmiguel, Jose M. Martinez, "Abandoned Object Detection under Sudden Illumination Changes", Actas del XXXI Simposium Nacional de la Unión Científica Internacional de Radio - URSI 2016, Madrid, Spain, Sept. 2016.

[16] Fulgencio Navarro, Erik Velasco, Jesús Bescós, "Enhancing discriminative tracking via RGBD dual-model" , Actas del XXXI Simposium Nacional de la Unión Científica Internacional de Radio - URSI 2016, Madrid, Spain, Sept. 2016.

[17] Matej Kristan et al. "The Visual Object Tracking VOT2016 Challenge Results", in: Hua G., Jégou H. (eds) Computer Vision – ECCV 2016 Workshops, Lecture Notes in Computer Science, vol. 9914, Springer, Oct. 2016, pp. 777-823 (DOI 10.1007/978-3-319-48881-3_54)

[18] Michael Felsberg et al., "The Visual Object Tracking VOT-TIR2016 challenge results", in: Hua G., Jégou H. (eds) Computer Vision – ECCV 2016 Workshops, Lecture Notes in Computer Science, vol. 9914, Springer, Oct. 2016, pp. 824-849 (DOI 10.1007/978-3-319-48881-3_55)

[19] Diego Ortego, Juan C. Sanmiguel and José M. Martínez, "Rejection based Multipath Reconstruction for Background estimation in SBMnet 2016 dataset", Scene Background Modeling Contest - SBMC2016, Proc. of 2016 International Conference on Pattern Recognition, Cancún, México, Dec. 2016 (DOI 10.1109/ICPR.2016.7899618).

[20] Álvaro García-Martín, Juan Carlos Sanmiguel, "Adaptive people detection based on cross-correlation maximization", Proc. of 2017 IEEE International Conference on Image Processing - ICIP 2017, Beijing, China, Sept. 2017.

[21] Juan Carlos Sanmiguel, Andrea Cavallaro, "Efficient estimation of target detection quality", Proc. of 2017 IEEE International Conference on Image Processing - ICIP 2017, Beijing, China, Sept. 2017.

[22] 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.

[23] 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.

[24] 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.

2.4. PhD Thesis

[25] **Contributions to region-based image and video analysis: feature aggregation, background subtraction and description constraining,** Marcos Escudero Viñolo (advisor: Jesús Bescós Cano), Escuela Politécnica Superior, Universidad Autónoma de Madrid, Jan. 2016

2.5. Master Thesis

[26] **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.

[27] **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.

[28] **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.

[29] **Detección de caídas mediante vídeo-monitorización (Fall detection using video)**, David Dean Pulido (advisor: José M. Martínez), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Universidad Autónoma de Madrid, Escuela Politécnica Superior, Mar. 2016.

[30] **Análisis de actividad en vídeos deportivos multicámara (Activity analysis in multicamera sports videos)**, Ángel Mora Sánchez (advisor: Rafael Martín), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Universidad Autónoma de Madrid, Escuela Politécnica Superior, Mar. 2016.

[31] **Estimación de fondo de escena en secuencias de vídeo (Background estimation in video sequences)**, Jaime Gallo de Cal, (advisor: Diego Ortego Hernández), Proyecto fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[32] **Evaluación comparativa de técnicas de detección y descripción de puntos de interés en imágenes (Comparative evaluation of Points of Interest techniques for detection and description in images)**, Miguel Martín Redondo, (advisor: Fulgencio Navarro Fajardo), Proyecto fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[33] **Automatización de funciones en el seguimiento del profesor para la emisión de clases presenciales (Automatization of functions for teacher tracking in lectures broadcasting)**, Alberto Palero Almazán, (advisor: Jesús Bescós Cano), Trabajo Fin de Máster (Master Thesis), Master en Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[34] **Reconocimiento de actividades utilizando información de color y profundidad (Activity recognition using color and depth information)**, Borja Olmo Esteban (advisor: Juan C. SanMiguel), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[35] **Integración y evaluación de sistemas de robo-abandono de objetos en video-seguridad (Integration and evaluation of stolen-abandoned object systems in video-surveillance)**, Jorge Gómez Vicente (advisor: Juan C. SanMiguel), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[36] **Análisis de actividad en vídeos de baloncesto (Activity analysis in basketball videos)**, Rubén García García (advisor: Rafael Martín), Proyecto Fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[37] **Seguimiento de objetos en vídeo a largo plazo (Long-term objects tracking in video sequences)**, Borja Maza Vargas (advisor: Juan Carlos San Miguel), Proyecto fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2016.

[38] **Detección de sombras en secuencias de vídeo-seguridad (Shadows detection in video-surveillance sequences)**, Guillermo Rodríguez Yrezabal (advisor: Juan Carlos San Miguel), Proyecto fin de Carrera (Master Thesis), Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2016.

[39] **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.

[40] **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.

[41] **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.

[42] **Clasificación automática de vídeos utilizando descriptores de características (Automatic classification of videos using features descriptors)**, María Narváez (advisors: Álvaro García-Martín, Tobias Senst), Trabajo Fin de Máster (Master Thesis), Master en Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[43] **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.

2.6. Graduate Thesis

[44] **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, Univ. Autónoma de Madrid, Jun. 2015.

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

[46] **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), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jun. 2015.

[47] **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, Univ. Autónoma de Madrid, Jun. 2015.

[48] **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, Univ. Autónoma de Madrid, Jul. 2015.

[49] **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, Univ. Autónoma de Madrid, Jul. 2015.

[50] **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, Univ. Autónoma de Madrid, Jul. 2015.

[51] **Detección de personas en entornos multicámara utilizando información contextual (People detection in multicamera environments using contextual**

information), Alejandro Miguélez Sierra (advisor: Rafael Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Ene. 2016.

[52] **Detección de personas en entornos residenciales y hospitalarios (People detection in residential and hospital environments)**, Jesús Molina Merchán (advisor: Rafael Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jun. 2016.

[53] **Detección de objetos abandonados para vídeo-vigilancia a largo plazo (Abandoned object detection for long-term video-surveillance)**, Sergio López Álvarez (advisor: Diego Ortego), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jun. 2016.

[54] **Reconstrucción de fondo de escena a partir de secuencias de vídeo (Background reconstruction from video sequences)**, Carolina Fernández-Pedraza (advisor: Diego Ortego), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[55] **Detección de personas utilizando contexto de la escena (People detection using scene context)**, Carlos Chaparro Pozo (advisor: Álvaro García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[56] **Seguimiento de objetos mediante constricción espacial de puntos TILDE (Object tracking via spatial constraining of TILDE points)**, César Augusto Betancur Cruz (advisor: Marcos Escudero), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jun. 2016.

[57] **Diseño y desarrollo de una herramienta para la extracción semi-supervisada de la información de contexto (Design and development of a tolos for semi-automatic extraction of contextual information)**, Raúl García Jiménez (advisor: Marcos Escudero), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[58] **Diseño de herramientas de apoyo para la detección de logotipos en secuencias de video (Design of supporting tolos for logo detection in video sequences)**, Efrén Martín García (advisor: Marcos Escudero), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jul. 2016.

[59] **Seguimiento automático de objetivos con drones mediante algoritmos de tracking (Automatic tracking of targets with drones using tracking algorithms)**, Roi Rico Díaz (advisor: Fulgencio Navarro), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[60] **Control de cámaras PTZ para la reconstrucción de escena basada en puntos de interés (PTZ cameras control for scene reconstruction based on Points of Interest)**, Elisa Martín Pérez (advisor: Fulgencio Navarro), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[61] **Seguimiento de objetos basado en múltiples algoritmos (Object tracking based on multiple algorithms)**, Eduardo Moreno de Pablos, (advisor: Juan C. SanMiguel Avedillo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2016.

[62] **Sistema Multi-cámara Distribuido basado en Unity (Multicamera distributed system based on Unity)**, Mario González Jiménez (advisor: Juan Carlos San Miguel), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Feb. 2017.

[63] **Análisis de consumo energético para aplicaciones de visión artificial (Analysis of power consumption for computer visión applications),** Pablo Sala del Real (advisor: Juan Carlos San Miguel), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jun. 2017.

[64] **Diseño de Redes de Cámaras Inteligentes utilizando Smartphones (Design of Smart Camaras Networks using Smartphones),** Fernando Lahoz Seguido (advisor: Juan Carlos San Miguel Avedillo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jul. 2017.

[65] **Detección y Reconocimiento de Señales Viales (Trafic Signs Detection and Recognition),** José Manuel Esteve de Prada (advisor: Álvaro García Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jul. 2017.

[66] **Aplicación de Realidad Aumentada Multi-Plano mediante Correspondencias entre Puntos de Interés (Multi-plane Augement Reality Application using Points of Interest Correspondences),** Alejandro Núñez Valle (advisor: Marcos Escudero Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jul. 2017.

[67] **Adaptación Automática de la Detección de Personas a la Escena (Automatic Adaptation of People Detección to the Scene)**, Aarón Monedero Grifo, (advisor: Álvaro García-Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[68] **Plataforma de Evaluación de Algoritmos de Detección de Personas (Platform for People Detection Algorithms Evaluation)**, Anthony Bryan Santiago Mendieta, (advisor: Álvaro García-Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[69] **Reconstrucción de Fondo de Escena basada en la Mediana (Median-based Scene Background Reconstruction)**, Emilio Gómez García, (advisor: Diego Ortego Hernández), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[70] **Detección de Personas mediante Redes Convolucionales (People Detection using Convolutional Neural Networks)**, Esther Sánchez Atienza, (advisor: Álvaro García-Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[71] **Segmentación Objeto-Fondo (Object-Background Segmentation)**, Paula Moral de Eusebio, (advisor: Álvaro García-Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[72] **Identificación de Materiales Utilizando el Sensor Kinect (Materials Identification using the Kinect Sensor)**, Miguel Basarte Mena, (advisor: Marcos Escudero Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[73] **Segmentación Espacio-Temporal del Contexto de un Vídeo (Spatio-Temporal Segmentación of Video Context)**, Sergio Serra Sánchez, (advisor: Marcos Escudero Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[74] **Seguimiento de Objetos basado en Múltiples Características (Object Tracking base don Multiple Features)**, Guillermo Luna Aguado, (advisor: Juan Carlos San Miguel Avedillo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

[75] **Estudio e implementación de algoritmos para detección de anomalías en entornos de videovigilancia (Study and implementaion of anomaly detection algorithms in video-surveillance environments)**, José Gil Torrecilla, (advisor: Luis Salgado Álvarez de Sotomayor), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2017.

3. Project Documents

3.1. Deliverables

3.2. First year Deliverables

3.2.1. D1.1 "System Infrastructure" version 1 (December 2015)

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 (December 2015)

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 (December 2015)

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 (December 2015)

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 [26][27] and Graduate Thesis [50]). 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 result two publications [12][13] 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 (December 2015)

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 [44].
The second work focuses on people privacy preservation in video-surveillance. This work is contained in master thesis [28].
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 [45].

3.2.6. D4.1 "Evaluation methodology and datasets" version 1 (June 2015) [cancelled]

This deliverable was cancelled.

3.2.7. D4.1 "Evaluation methodology and datasets" version 1 (December 2015)

In relation with people detection, a new people detection dataset has 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.

3.2.8. D4.3 "Results Report" version 1 (June 2015) [delayed]

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

3.2.9. D4.3 "Results Report" version 1&2 (December 2015)

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

3.3. Third semester Deliverables

3.3.1. D2.2 "Contextual modelling and extraction for people behaviour understanding" version 1 (June 2016)

This deliverable describes the work related with the task T2.2, context modelling and extraction. Eight main works have been developed and are presented below.

The first approach is related to the development of semi-automatic annotation tools [49] based on image and video processing algorithms. The objective is to reduce the effort and time required to annotate video sequences.

The second approach aims to design and develop a set of algorithms able to cope with the problem of modelling the epipolar geometry between two widely separated cameras capturing the same scene [46].

The third approach is focused on the implementation of a system able to segment a video flow generated by a forward-looking camera placed in the frontal of a vehicle, with the aim of detecting vehicles, pavement and lane marks. The system handles contextual information of the scene to generate statistical models of the regions of interest [47].

The fourth approach is related to include context information to people detection and combine information obtained from different cameras [51].

The fifth approach proposes an approach for part-based people detection in images that uses pre-annotated contextual information to weight the contribution of each model part [1]. An upgrade of this approach is also developed in [55].

The sixth approach enhances [19] by designing and developing strategies to propagate the user annotation of contextual objects from the first frame to the rest of the video. The main objective is to provide tight-to-object annotation masks of dynamic contextual objects while further reducing the required amount of user interactions [57].

The seventh approach [60] aims to generate the scene panoramic using a PTZ-recording camera.

Finally, in the eighth approach [25], regions are used to constrain the descriptions of singular points. The advantages of this scheme are illustrated by means of a method for object identification in cluttered scenarios.

3.3.2. D2.3 "Online quality analysis of people behaviour understanding" version 1 (June 2016)

This deliverable describes the work related with the task T2.3, quality analysis. Various activities have been undertaken and are presented below.

The first activity considers a video tracking application where multiple algorithms are employed to track a single target [3]. Tracker's quality is estimated to cluster trackers over time in order to identify the ones following the right target. Intra-tracker and inter-tracker spatio-temporal measures are proposed to estimate the quality of each algorithm.

The second activity is related to the study of good object properties to estimate the quality of Background Subtraction algorithms without using ground-truth. A set of 21 measures is analysed in a heterogeneous dataset using four algorithms, demonstrating that fitness between objects and segmented image regions has a great potential to replicate performance.
The third activity consists in the development of a Background Estimation algorithm for video sequences [4] robust to stationary objects. The main novelty consists in a multipath reconstruction scheme that allows improving performance compared to related literature.
The fourth activity involves the implementation of two algorithms for Background estimation in video sequences [54] from the literature. The experiments carried out demonstrates superior capabilities of block-wise approaches against pixel-wise ones in challenging scenarios.

3.3.3. D4.3 "Results Report" version 3 (June 2016)

This *report* recapitulates the results obtained within the HAVideo project during the first semester of 2016. The results are announced as they are obtained both at the Web site (http://www-vpu.eps.uam.es/HAvideo).

3.4. Fourth semester Deliverables

3.4.1. D2.1 "People Behaviour understanding in single and multiple camera settings" version 2 (December 2016)

This deliverable describes the work related with the task T.2.1 Analysis Tools for human behaviour understanding. The people behaviour understanding in this project has been already designed as a sequential combination of object segmentation, people detection, object tracking and behaviour recognition. In particular, during the second year of the project there have been a focus on developing different approaches for people detection, tracking and behaviour recognition.
In relation with local detection and description features, an evaluation framework and a comparative analysis of state of the art local features detection and description algorithms has been developed [32].
In relation with people detection, [52] propose a sitting person model with the aim of completing a detector for a nursing home scenario and explore the possibility of creating synthetic images datasets reducing the amount of resources needed and save the cost of having to record sequences for a detector in this specific nursing home scenario.
In relation with tracking, a video object tracker based on the point-of-interest TILDE has been implemented and evaluated [56]. Also, a new visual attention model based on a joint perceptual space of both color and brightness for improved video tracking is proposed [5]. In addition, an enhanced version of a teacher tracking system in a classroom has been implemented and tested [33]. And finally, a fusion algorithm [37] has been implemented which examines the behaviour of the combination of different tracker from the SoA under a long-term framework.
In relation with behaviour recognition, a configurable abandoned-stolen object detection system in security-video that integrates the most relevant techniques in each one of its

stages is proposed [35]. Also, a fall detection video-based system to promote independent living for the elderly is proposed [29]. In addition, a unified prototype Graphical User Interface (GIU) for sport video-content analysis has been implemented (soccer [30] and basketball [36]).

3.4.2. D4.1 "Evaluation methodology and datasets" version 2 (December 2016)

This deliverable updates the work related with the task T.4.1: Evaluation framework.
In relation with segmentation, a new dataset containing 4 categories or challenges with 10 video sequences for the task of Background Estimation has been generated [54]. In addition, we have participated in an International background estimation challenge (SBMnet 2016 challenge) [4]. This deliverable includes a description of the proposed methodology and dataset presented in the SBMnet challenge.
In relation with tracking, we have presented a tracker to an international tracking challenge (VOT challenge 2015 [12][13] and 2016 [17][18]). 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 updates the description of the proposed methodology and dataset presented in the VOT 2016 challenge.
In relation with behaviour recognition, a configurable abandoned-stolen object detection system in security-video that integrates the most relevant techniques in each one of its stages is proposed [35] and described in task T.2.1. This deliverable includes a description of the proposed methodology and dataset presented in [35].

3.4.3. D4.3 “Results Report” version 4 (December 2016)

This *report* recapitulates the results obtained within the HAVideo project during the second semester of 2016. The results are announced as they are obtained both at the Web site (http://www-vpu.eps.uam.es/HAvideo).

3.5. Fifth semester Deliverables

3.5.1. D1.1 “System Infrastructure” version 2 (June 2017)

This deliverable updates the work related with the task T1.1. To mimic the operation modules of smart-cameras, new portable equipment has been acquired and documented, which is composed of four mid-range smart-phones. Moreover, an API has been developed to enable the use of this portable infrastructure following the client-server architecture. Finally, some contributions of the second year have been applied to the new equipment, such as the development of the consumption models.

3.5.2. D1.3 “Simulator documentation: description, programmer manual and user guide” version 2 (June 2017)

This deliverable updates the work related with the task T1.3. During the second year, a multi-camera simulator has been developed for virtual data based on the previous experience of the first year. A recently released-for-free game engine has been included

in the simulator to provide highly realistic graphics and the software framework to manage and control cameras has been improved to support up to 8 cameras running at 25 fps. To do so, we have taken advantage of the hardware infrastructure acquired during the second year. The framework has been fully documented and tutorials for its use have been created.

3.5.1. D2.2 "Contextual modelling and extraction for people behaviour understanding" version 2 (March 2017)

This deliverable updates the work carried out for context modelling and extraction. The contributions in this line of research are divided regarding the scenario into mono and multi-camera approaches. Regarding mono-camera related approaches we further explore the use of context to describe points-of-interest, incorporating motion statistics to isolate a point from its non-related neighbours. Besides, a new descriptor to fight camouflage in background subtraction scenarios is proposed. The descriptor is based on the use of a dynamically truncated DCT for characterisation. Moreover, initial studies on contextual transferring have been carried out. Regarding multi-camera related approaches, we refine existing strategies to relate people detections across cameras; using offline extracted contextual information on the cameras' position. We have also designed a method to incorporate, refine and associate semantic segmentations extracted for each camera in a multi-camera scenario. The segmentation spatially define areas-of-interest in the camera views (e.g. floor, doors, and walls) on which search for particular related people behaviours.

3.5.2. D2.3 "Online quality analysis of people behaviour understanding" version 2 (June 2017)

This deliverable updates the work carried out for online quality analysis in various research areas. We keep on exploring the standalone evaluation of foreground segmentation maps in video by providing an in-depth study of the related state-of-the-art and proposing solutions to tackle the existing challenges. Moreover, we also continue with the combination of video tracking algorithms based on quality measures and propose a combination based on colour and edges based on background information.

3.5.3. D3.1 "Online adaptive people behaviour understanding based on contextual and quality information" version 1 (March 2017)

This deliverable summarizes the work during the first and half year(s) of the project for the task T3.1 "Adaptive approaches" (WP3 "Self-configurable approaches for long-term analysis").whose goal is to analyse alternatives to include contextual and quality information in the developed algorithms to adapt their operation to the changing environment/conditions. Adaptation is targeted at three different levels: model, algorithm configuration and processing strategy. This task T3.1 depends upon developments within WP2 (T2.1 Analysis tools for human behavior understanding, T2.2 Contextual modelling and extraction and T2.3 Quality analysis). The results of this task T3.2 will provide self-configurable approaches for long-term analysis and WP4 Evaluation framework, demonstrators and dissemination.

3.5.4. D3.2 "Collaborative approaches for people behaviour understanding" version 1 (March 2017)

This document summarizes the work during the first and half year(s) of the project for the task T3.2 "Collaborative approaches" (WP3 "Self-configurable approaches for long-term analysis"). whose goal is to exploit interactions among multiple entities to optimize the overall performance (accuracy or resource-usage). First, we consider the processing stages where interactions are based on quality and contextual information. Second, we investigate approaches in camera networks where the quality and contextual information of each camera have to be distributed and used by other cameras in order to coordinate them. This task T3.2 depends upon developments within WP2 (T2.1 Analysis tools for human behavior understanding, T2.2 Contextual modelling and extraction and T2.3 Quality analysis). The results of this task T3.2 will provide self-configurable approaches for long-term analysis and WP4 Evaluation framework, demonstrators and dissemination.

3.5.5. D4.1 "Evaluation methodology and datasets" version 3 (June 2017)

This deliverable updates the work related with the task T.4.1: Evaluation framework. In relation to detection, a new dataset for cars and parking lot detections has been recorded. The dataset consists of two main image sets, a training set used with 6616 frames, and a test set with its associated Ground-truth. The test set includes a multicamera synchronized scenario.

3.5.6. D4.2.1 "Framework and Guidelines for the development of applications" version 1 (March 2016) [rescheduled March 2017]

This document presents the existing framework and guidelines for the development of applications and demonstrators in relation with the project. First, we describe the image processing infrastructure and software available in the Video Processing and Understanding Lab. And then, we describe the guidelines for the development of applications, including recommendations for the visualization of the results or outcomes of the applications or demonstrators.

3.5.7. D4.3 "Results Report" version 5 (June 2017)

This *report* recapitulates the results obtained within the HAVideo project during the first semester of 2017. The results are announced as they are obtained both at the Web site (http://www-vpu.eps.uam.es/HAvideo).

3.6. Sixth semester Deliverables

3.6.1. D1.1 "System Infrastructure" version 3 (December 2017, if required)

This deliverable has not been generated as no acquisitions of new equipment have been performed during the sixth semester.

3.6.2. D1.2 “DiVA documentation” version 3 (December 2017, if required)

This deliverable has not been generated as no new functionalities have been included during the sixth semester (and therefore, there is no need of improving the existing documentation).

3.6.3. D1.3 “Simulator documentation: description, programmer manual and user guide” version 3 (December 2017, if required)

This deliverable has not been generated as no new functionalities have been included during the sixth semester (and therefore, there is no need of improving the existing documentation).

3.6.4. D2.1 “People Behaviour understanding in single and multiple camera settings” version 3 (December 2017, if required)

This deliverable updates the work related with the task T.2.1 Analysis Tools for human behaviour understanding.

In relation with people/object detection, [70] presents a comparison between people detection approaches based in hand craft features versus those based on convolutional neural networks. [71] presents a generalization of a previous people-background segmentation approach to any object model: a generic object-background segmentation. [68] presents an automatic benchmarking platform for people detection algorithms evaluation. Finally, [65] propose a basic traffic signs detection and recognition approach based in signs shapes and colour.

In relation with behaviour recognition, an automatic classification of videos using features descriptors (SIFT, SP-SIFT and Lagrarian) has been implemented [42].

3.6.5. D2.4 "Exploration and viability studies for people behaviour understanding” version 3 (December 2017, if required)

This deliverable includes an update with the advances of the materials identification work, presented in the previous version, and a new research line that aims to create a naïve prototype of augmented reality.

3.6.6. D3.1 "Online adaptive people behaviour understanding based on contextual and quality information” version 2 (December 2017)

This deliverable includes an update on three recent developments regarding the quality-based improvement of foreground masks from background subtraction approaches, multi-target tracking employing probabilistic data association filters and the estimation of detection quality in multi-camera setups. Regarding the use of contextual information, it also includes updated results on the use of context to constrain and refine people detection in multi-camera setups.

3.6.7. D3.2 "Collaborative approaches for people behaviour understanding" version 2 (December 2017)

This deliverable includes an update on recent works for multiple cameras dealing with automatic discovery of camera location, collaborative object detection for pedestrians and a survey for task-based collaboration employing quality indicators.

3.6.8. D4.1 "Evaluation methodology and datasets" version 4 (December 2017, if required)

This deliverable has not been generated as no new evaluation methodology or dataset have been used during the sixth semester.

3.6.9. D4.2.2 "Applications" (December 2017)

This deliverable describes 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. [39] presents a multi-camera pedestrian detector with semantic constraining demonstrator. [40] proposes a complete abandoned object detection (AOD) system demonstrator. And [41] describes a long-term tracking demonstrator.

3.6.10. D4.3 "Results Report" version 6 (December 2017)

This *report* recapitulates the results obtained within the HAVideo project during the last semester of 2017. The results are announced as they are obtained both at the Web site (http://www-vpu.eps.uam.es/HAvideo).

3.7. Technical Reports

4. Content Sets

4.1. Wheelchair Users dataset - WUds

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

This dataset was recorded for the work extending a people detector to wheelchair users, in order to provide a more general solution to detect people in environments such as independent and assisted living, hospitals, healthcare centres and senior residences.

4.2. Background Estimation dataset - BEds

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

The Background Estimation dataset (BEds) is a corpus of video sequences generated from publicly available video-surveillance datasets to cover several Background Estimation challenges. The dataset is focused on 4 challenges or categories conformed by 10 video sequences each and the associated ground-truth background image.

4.3. Parking Lot dataset - PLds

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

The Parking Lot dataset (PLds) was recorded due to the lack of public parking car datasets. The sequences were recorded in a real environment, the Pittsburgh International Airport parking, in order to work with an environment as realistic as possible. The dataset consists of two main image sets, a training set used with 6616 frames, and a test set with its associated Ground-truth. The test set consists of a long (named 'All') and a short (named 'Multicamera') version of the images, with 1000 and 100 frames, respectively. The short version (multicamera sets) is contained in the long version and has the frames synchronized between the two cameras, allowing to evaluate experiments combining the information of both cameras.

In addition to generating the images, the vehicles of all images have been manually annotated. The training images have been annotated for its use in the generation of the parked vehicle model, and the test images for the evaluation of the system. In the case of the synchronized multicamera set, the vehicle occupancy matrix has been manually generated.

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 not 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

5.1.2. 2016 Developers Workshop

The 2016 Developers Workshop was held May 2016 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid.
The Workshop consisted of two parts:

- Dissemination Day (May 23rd 2016): Video Analysis Technologies and applications for Computer Vision.
- Short course: (May 24th-26th 2016): Introduction to Computer Vision applications programming with OpenCV

There were 16 participants.

5.1.3. 2016 Dissemination Workshop

The 2016 Dissemination Workshop was organized for being held November 25th 2016 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid. Due to the reduced number of registered companies, the Workshop was cancelled.

5.1.4. 2017 Developers Workshop

The 2017 Developers Workshop was held from May 30th to June 2nd 2017 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid.
The Workshop consisted of two parts:

- Dissemination Day (June 1st 2017): Video Analysis Technologies and applications for Computer Vision.
- Short course: (May 30th-31st and June 2nd 2017): Introduction to Computer Vision applications programming with OpenCV

There were 18 participants in the short course and 26 in the Dissemination Day.

5.2. Seminars

6. Main Achievements of the Project

6.1. First year main achievements

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

- Task 1.3: Update and maintenance of a camera network simulator
  - Development of a virtual camera network simulator [48]
  - Improvement of an existing camera network simulator to include embedded hardware modelling, sensor models and real communication protocols

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

- Task 2.1: Analysis Tools for human behaviour understanding
  - Development of a long-term stationary object detection based on spatio-temporal change detection [2].
  - Development of a people density estimation in crowded scenarios [26] and a people detection in groups [27][50].
- Task 2.2: Contextual modeling and extraction
  - Development of a context-aware people detector [1][55] which uses static contextual annotation to integrate and weight partial evidences from a part-based people detector algorithm.
  - Improvement of an existing manual annotation tool, to provide semiautomatic tools that speed up the annotation process.
  - Development of a camera calibration algorithm [46] to detect the relative position of a camera respect to another in wide-baseline scenarios.
  - Development of a vehicle detection algorithm [47] 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 [49].
- Task 2.4: Exploration and viability studies
  - Implementation of a privacy preserving module for existing video surveillance systems [28].
  - Development of a heart rate detector using video [44].
  - Development of a system for material identification through the Kinect technology [45].
  -

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

- Task 4.1: Evaluation framework
  - International tracking challenge (VOT challenge) evaluation [12][13].
  - Development of a new dataset for wheelchairs users and standing people detection in a real in-door senior residence environment.
- Task 4.3: Dissemination
  - 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)

6.2. Third semester main achievements

For workpackage WP.1: Video Analysis Framework, the work on the third semester has been focused on improving the virtual camera network simulator using the Unity tool. This work is expected to be completed in the fourth semester (Task 1.3: Update and maintenance of a camera network simulator)

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

- Task 2.1: Analysis Tools for human behaviour understanding
  - Development of a new Visual Attention Model enhancing discrimination among objects and its application to tracking [5]
  - Development of a fail detection algorithm [29]
  - Development of two applications for sport analysis of multicamera sports videos [30].
  - Enhancements to a long-term teacher-tracking system for the real-time distribution of classroom activities [33].
  - Compiled and generated training image sets (real and synthetical) to train people detectors adapted to hospitalary environments and nursing homes [52].
  - Development of a robust to appearance changes tracker based on the use of a state-of-the-art point-of-interest detector [56].
  - Development of an evaluation framework and a comparative analysis of state of the art local features detection and description algorithms in a Master Thesis [32]
- Task 2.2: Contextual modeling and extraction
  - Study to include context information to people detection and combine information obtained from different cameras [51].
  - Improvement of an existing tool and of its associated user interface to propagate the annotation of static and dynamic contextual objects in video sequences [57].
  - Development of an online panoramic background initialization algorithm for PTZ cameras in a Graduate Thesis [60].
- Task 2.3: Quality Analysis
  - Study of good object properties for stand-alone evaluation of Background Subtraction algorithms.
  - Study of online people detection algorithms quality analysis using correlation metrics.
  - Development of a Background Estimation algorithm for video sequences robust to stationary objects [4].
  - Development of a Background Estimation algorithm for video sequences in a Master Thesis [31].
  - Development of a long-term abandoned object detector robust against sudden illumination changes and stationary pedestrians [53].

  - Implementation of two algorithms for Background Estimation from the literature [54].
  - Considering the potential of measures based on fitness to regions, new measures for evaluation of Background Subtraction without ground-truth are being developed.
  - Development of a measure to estimate tracking quality based on reverse trajectory analysis employed for combining multiple algorithms [3].
  - Development of an adaptation of state of the art tracking algorithms to perform long-term on UAVs sequences in a Graduate Thesis [59].
- Task 2.4: Exploration and viability studies
  - Development of pre-processing algorithms to constrain logo detection [58].

Workpackage WP.3: Self-configurable approaches for long-term analysis, has been delayed till the second semester of 2016.

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

- Task 4.1: Evaluation Framework
  - Generation of a dataset containing 4 categories or challenges with 10 video sequences for the task of Background Estimation [54].
  - International tracking challenge (VOT challenge) evaluation (results will be available in the next months).
- Task 4.3: Dissemination
  - The 2016 HAVideo Developers workshop was held May 2016 at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid (see 5.1.2)

6.3. Fourth semester main achievements

For workpackage WP.1: Video Analysis Framework, the work on the fourth semester has been focused on improving the virtual camera network simulator using the Unity tool. This work has been completed in the fourth semester (Task 1.3: Update and maintenance of a camera network simulator). Moreover, a network camera simulator has been developed to model the communication and computational resources required by collaborative cameras **¡Error! No se encuentra el origen de la referencia.** .

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

- Task 2.1: Analysis Tools for human behaviour understanding
  - Development and tested a people density estimation in crowded scenarios [14].
  - Development and tested of an interactive detection and tracking application for basketball players [36].
  - Analysis of tracking objects in long-term sequences [37].
- Task 2.2: Contextual modeling and extraction

  - Use of semantic segmentation to constraint people detection proposals (publication pending).
- Task 2.3: Quality Analysis
  - Design and implementation of a shadow detection algorithms that perform quality analysis for self-adaptation [38].
  - Implementation of an abandoned object detection system under sudden illumination changes [15].
  - Development of a discriminative tracking using the combination of RGB and depth information [16].
- Task 2.4: Exploration and viability studies
  - The goal of this task is to explore research alternatives not directly considered within the project. There are not updates for this task during fourth semester.

Workpackage WP.3: Self-configurable approaches for long-term analysis

- Task 3.1
  - Adaptation of people detection thresholds during prediction time (publication pending).
- Task 3.2
  - Design and implementation of models to estimate the resources required for collaboration, enabling the adaptation of self-configurable approaches in multi-camera settings [6].

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

- Task 4.1: Evaluation Framework
  - International tracking challenge (VOT 2016 challenge) evaluation [17][18].
  - International background estimation challenge (SBMnet 2016 challenge) evaluation [19].
  - Design and evaluation of a configurable abandoned-stolen object detection system in security-video that integrates the most relevant techniques in each one of its stages [35].
- Task 4.3: Dissemination
  - The 2016 HAVideo Dissemination workshop, organized to be help Novermber 25th at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid, was cancelled (see 5.1.3).

6.4. Fifth semester main achievements

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

- Task 1.1: Acquisition of portable equipment
- Task 1.3: Update and maintenance of a camera network simulator
  - Improvement of the virtual camera network simulator with a new graphics engine and client-server architecture [48][62]

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

- Task 2.2: Contextual modelling and extraction
  - A new descriptor to fight camouflage (publication pending).
  - Use of region-constrained descriptors to identify objects in 2.5D scenarios (publication pending).
  - Use of region-and-motion-constrained descriptors to detect events of interest in video sequences (publication pending).
  - Combination of semantic segmentations in multi-camera scenarios (publication pending).
- Task 2.3: Quality Analysis
  - Study in depth of quality estimators for foreground segmentation maps (publication pending).
  - New techniques to improve precision of foreground segmentation maps have been developed. This completed work is part of the ongoing work in performance improvement.
  - Estimation of detection quality in multi-camera networks [21]
  - Estimation of detection quality for multiple people detectors [21]
- Task 2.4: Exploration and viability studies
  - The goal of this task is to explore research alternatives not directly considered within the project. There are not updates for this task during fifth semester.

Workpackage WP.3: Self-configurable approaches for long-term analysis

- Task 3.1
  - Context-based adaptation of people detectors [1]
  - Abandoned object detection robust to illumination changes [15]
  - Video tracking based on dual RGB-D models [16]
  - People detection based on adaptive scale selection [27]
- Task 3.2
  - Collaborative multi-camera tracking via modelling resource usage [6].
  - Collaborative multi-algorithm tracking based on quality estimation [61]
  - Study for performance of video trackers for long-term operation [37]
  - Detection threshold adaptation during runtime for shadow detection [38]

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

- Task 4.1: Evaluation Framework
  - A new Parking Lot dataset (PLds) was recorded due to the lack of public parking car datasets. The sequences were recorded in a real environment, the Pittsburgh International Airport parking, in order to work with an environment as realistic as possible. The dataset includes a multicamera environment.
- T.4.2: Use Cases and Demonstrators
  - Definition of the framework and guidelines for the development of applications.
  - Planning of the first set of applications.

6.5. Sixth semester main achievements

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

- Task 2.1: People Behaviour understanding in single and multiple camera settings
  - Automatic classification of videos using features descriptors [42].
  - Trafic Signs Detection and Recognition [65].
  - People Detection using Convolutional Neural Networks [70].
  - Object-Background Segmentation [71].
  - Platform for People Detection Algorithms Evaluation [68].
- Task 2.2: Contextual modelling and extraction
  - Scene-based automatic initialization of region seeds for skin lesion segregation and characterization (publication pending).
  - Automatic aggregation and refinement of contextual information in scenarios recorded with multiple moving cameras (publication pending).
- Task 2.3: Quality Analysis
  - Development of a Background Estimation algorithm for video sequences based on the temporal median [69].
  - Publication of the study of quality estimators for foreground segmentation maps [9] done in previous semesters.
- Task 2.4: Exploration and viability studies
  - Continuation of the visual materials identification work, presented in the previous chapter, considering more advanced techniques and a more elaborated processing.
  - Creation of a naïve prototype of augmented reality.

For Workpackage WP.3: Self-configurable approaches for long-term analysis, the main achievements are:

- Task 3.1 Online adaptive people behaviour understanding based on contextual and quality information
  - Automatic Adaptation of People Detection to the Scene [67].
  - Detection quality estimation for multi-camera setups [21]
  - Probabilistic multi-target multi-camera tracking (publication pending)
  - Using scene context to inhibit and improve people detection, updated results (publication pending).
- Task 3.2
  - Adaptation of people detection thresholds during prediction time [20]
  - Collaborative multi-camera location estimation.
  - Collaborative multi-camera analysis based on quality: a survey (publication pending)
  - Guiding video object segmentation with external results (publication pending)
  - Improvement of background subtraction algorithms based on quality information (publication pending).

For workpackage WP.4: Evaluation framework, demonstrators and dissemination, the main achievements are the implementation of the following applications/demonstrators:

- A multi-camera pedestrian detector with semantic constraining demonstrator [39].
- A complete abandoned object detection (AOD) system demonstrator [40].
- A long-term tracking demonstrator [41].