



HAVi Newsletters

#08 - August 2017

TEC2014-53176-R HAVideo (2015-2017)

High Availability Video Analysis for People Behaviour Understanding

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

Eight four-months period progress report

The project has been working on plan and five deliverables have been published on schedule:

- D1.1 "System Infrastructure" version 2 (June 2017)
- D1.3 "Simulator documentation: description, programmer manual and user guide" version 2 (June 2017)
- D2.3 "Online quality analysis of people behaviour understanding" version 2 (June 2017)
- D4.1 "Evaluation methodology and datasets" version 3 (June 2017)
- D4.3 "Results Report" version 5 (June 2017)

Due to the delays in WP4 with respect to application developments, we are working in a request to the Ministry for the extension of the project during part of 2018.

The main achievements of the different workpackages are described below.

WP1 "Video Analysis Framework"

Portable equipment (smartphones) have been acquired and used the design of camera networks using these devices as well as for monitoring consumption in this scenario. The virtual camera network simulator has been improved with a new graphics engine and a client-server architecture.

WP2 "Video analysis tools, models and performance indicators"

The main achievements of WP2 are:

- Development of a new descriptor to fight camouflage.
- Use of region-constrained descriptors to identify objects in 2.5D scenarios.
- Use of region-and-motion-constrained descriptors to detect events of interest in video sequences.
- Combination of semantic segmentations in multi-camera scenarios

- Complete and rigorous study of quality estimators for foreground segmentation maps.
- Development of new techniques to improve precision of foreground segmentation maps. This completed work is part of the ongoing work in performance improvement.
- Estimation of detection quality in multi-camera networks.
- Estimation of detection quality for multiple people detectors.

WP3 “Self-configurable approaches for long-term analysis”

For WP3, the main achievements are:

- Development of Context-based adaptation of people detectors.
- Adaptation of people detection thresholds during prediction time.
- Development of abandoned object detection robust to illumination changes.
- Development of a video tracker based on dual RGB-D models.
- Development of a people detector based on adaptive scale selection.
- Development of approaches for collaborative multi-camera tracking via modelling resource usage.
- Development of approaches for collaborative multi-camera multi-target tracking via detection quality estimation.
- Development of a collaborative multi-algorithm tracker based on quality estimation.
- Study for performance of video trackers for long-term operation.
- Development of approaches for detection threshold adaptation during runtime for shadow detection.

WP4 “Evaluation framework, demonstrators and dissemination”

For this workpackage, the main achievements are:

- Development of a Parking Lot dataset (PLDs). It was recorded due to the lack of public parking car datasets, including a multicamera environment.
- Planning of the first set of applications.
- First versions of some applications (on-line contextual updating, abandoned object detection and people tracking).
- Participation in the VOT2017 challenges (results pending).
- 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.

Eight four-months period results

Datasets

Parking Lot dataset – PLds

This dataset 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.



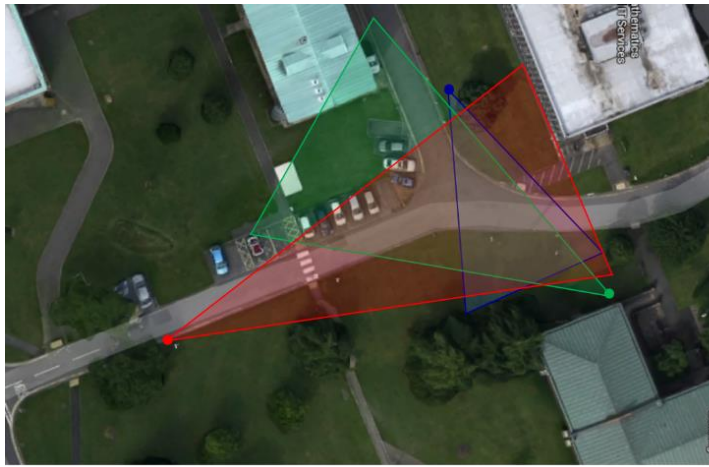
In addition to generating the images, the vehicles of all images have been manually annotated. The training images have been annotated for its use in the generation of the parked vehicle model, and the test images for the evaluation of the system. In the case of the synchronized multicamera set, the vehicle occupancy matrix has been manually generated. This dataset is freely available for research purposes and can be downloaded following the Content Sets link at the HAVideo web site.

Journals

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](https://doi.org/10.1109/MC.2017.213))

Abstract: Smart-camera networks enable a wide range of services for vehicular ad-hoc networks, smart cities, home automation, wide-area surveillance, and search and rescue operations. These networks of cameras with inbuilt processing and communication capabilities generate large volumes of data, share high-data-rate messages and generally operate with limited resources. To design and test new applications for smart-camera networks a suitable simulator is needed to support

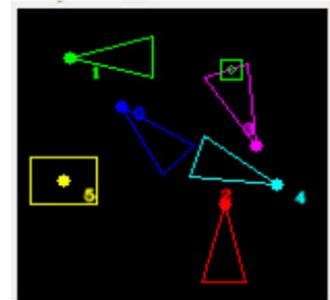
the development and accurately predict the performance of vision algorithms before deployment.



(b)



(a)



(c)

Conferences

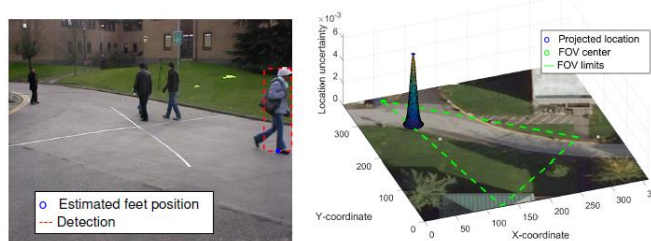
Juan Carlos San Miguel, Andrea Cavallaro, "**Efficient estimation of target detection quality**", Proc. of 2017 IEEE International Conference on Image Processing – ICIP 2017, Beijing, China, Sept. 2017 (in press).

Abstract: The quality of a target detection indicates the confidence for a (smart) camera to view and locate an object. The capability to determine the quality of target detections is pivotal in many smart-camera applications such as autonomous vehicles and surveillance. We propose to estimate the quality of target detections by integrating the target location uncertainty over polygonal domains which abstract each camera's field of view. As this problem is often intractable analytically, we define a framework based on numerical integration that easily accommodates multiple models for uncertainty and fields of view. We perform quadrature-based integration combined with importance sampling to provide accurate

quality estimations while reducing the associated computational cost. The proposed method outperforms alternative

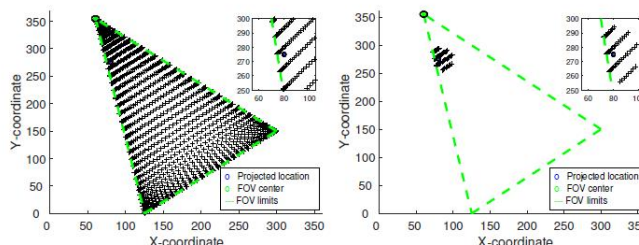
approaches in terms of estimation accuracy and execution time. We validate the proposed approach with a recent distributed multi-camera multi-

target tracker which we improved to consider realistic fields of view. Results demonstrate the effectiveness of the proposed method in decreasing the state estimation error.



(a) Target detection

(b) Ground-plane projection



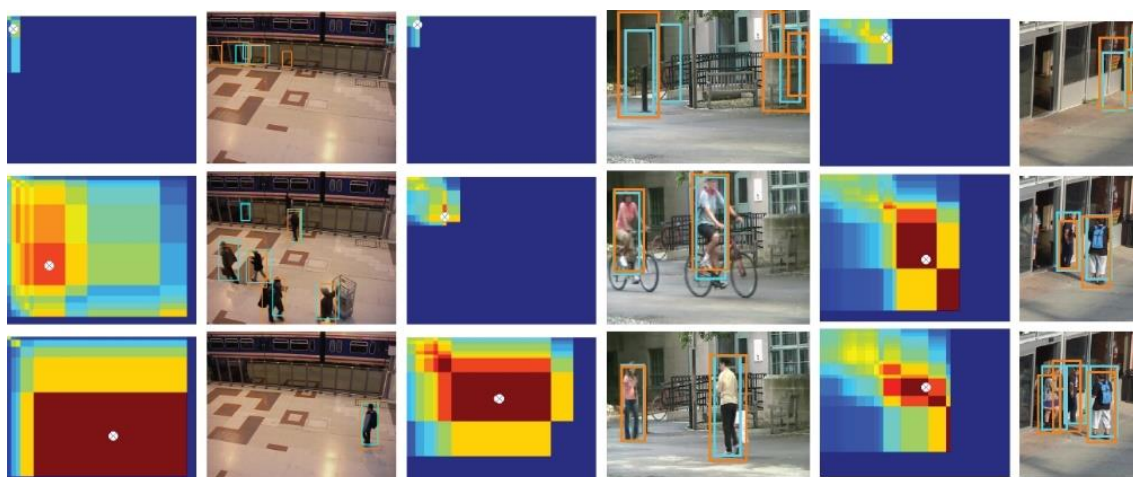
(c) Quadrature decomposition

(d) Proposal

Álvaro García-Martín, Juan Carlos San Miguel, "**Adaptive people detection based on cross-correlation maximization**", Proc. of 2017 IEEE International Conference on Image Processing – ICIP 2017, Beijing, China, Sept. 2017 (in press).

Abstract: Applying people detectors to unseen data is challenging since patterns distributions may significantly differ from the ones of the training dataset. In this paper, we propose a framework to adapt people detectors during runtime classification. Such adaptation takes advantage of multiple detectors to identify their best configurations (i.e. detection thresholds) without requiring manually labeled ground truth. We maximize the mutual information of detectors by pair-wise correlating their outputs to obtain a set of hypotheses for the detection thresholds.

These hypotheses are later combined by weighted voting to obtain a final decision for the detection threshold of each detector. The proposed approach does not require re-training detectors and uses standard people detector outputs, i.e., bounding boxes, therefore it can employ various types of detectors. The experimental results demonstrate that the proposed approach outperforms state-of-the-art detectors whose optimal configuration is learned from training data.



Master thesis

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.

Abstract: This project describes a system to perform pedestrian detection and semantic segmentation in a multi-camera recorded scenario. The fusion of both disciplines leads to contextually filtered pedestrian detection. The multi camera system is used to reproject detections from one camera to the others. As an use example, statistical data usage of specific semantic areas in the scene is also extracted. For user interaction a Graphical User Interface (GUI) based on a multithread application has been designed. The GUI allows the user to define the method setup as well as display results in execution time. In order to carry these tasks out a study of the state of the art has been done. Pedestrian detection approaches are reviewed emphasizing in those that rely on object proposals. Also, recent trends in the task of Pedestrian Detection are analyzed. In addition, current state of the art in the extraction of contextual information and –specifically– on the use of, semantic segmentation is studied. Finally, multi-camera scenarios are also described. A new system has been proposed in order to achieve the objectives and perform pedestrian detections under different filtering and fusion conditions. Detectors such as HOG, DPM, ACF, Fast-RCNN or

PSP-Net have been integrated and a complete semantic segmentation has been performed. Both information has been combined in a common developed frame. Finally, system performance has been tested in a generated dataset with manually annotated ground truth.

Abandoned Object Detection in 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.

Abstract: Due to recent events, global security concern is significantly increasing in our society. This area of concern is directly connected with the growing demand of video surveillance systems, mainly in public and crowded scenarios, such as railway stations, due to the potential risk they present. In order to avoid the arduous manual task of supervising a video surveillance system, automatic analysis and detection of this kind of events is the challenging task to be achieved. Although diverse systems trying to reach this goal have been proposed in the literature there is a lack of evaluation within this field. An end-to-end configurable system for abandoned and stolen object detection has been designed and developed by integrating available techniques. This system integrates several algorithms in each module of the system, thus it allows the evaluation of several state-of-the-art techniques combinations. An evaluation protocol considering short and long-term sequences has been designed by classifying available datasets and analysis evaluation metrics. A graphical user interface has been developed allowing the algorithms and parameters selection and adjustment for each stage of the system, as well as displaying the results. In addition, a new different system integrating recent and innovative state of the art proposals has been proposed.

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.

Abstract: The objective of this Master Thesis is to improve the performance of an existing tracker, called PKLTF. 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 final version. These improvements allow deal with scale changes and maintain the real-time performance. Finally, the proposed tracking algorithm is evaluated against a representative selection of trackers of the state-of-the-art. The new tracker improves the performance of the base tracker in the comparative evaluation.

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

Abstract: Video surveillance systems are becoming more and more indispensable as far as security is concerned. Thanks to the increase in the number of cameras and recording times, the need arises to create automatic recognition systems that allow to detect different situations or events that are significant from the point of view of security. This project focuses on the automatic detection of violence based on the unique analysis of the images (frames) of the video, that is, removing additional information such as audio or context information. Currently, there are several algorithms developed with this type of application in mind: MoSIFT, LaSIFT or STIP. All of them are based on the extraction and description of characteristics. The proposed algorithm, LaSP-SIFT, is based on the description of the appearance and movement of the video using the SP-SIFT descriptor. The extraction of the movement is done by calculating the paths of each of the pixels of the image during a certain number of frames. These trajectories are known as Lagrangian measurements. After describing both the appearance and the movement of all frames, the Bag-of-Words (BoW) technique is applied to extract the most representative descriptors of the complete set also called vocabulary or dictionary. Once this is done, a single frequency histogram per video is calculated using the vocabulary extracted previously and these histograms will be the input to the training of the final Vector Support Machine. Finally, we evaluate the trained classifiers and compare them with the state of the art algorithms.

Graduate thesis

Análisis de consumo energético para aplicaciones de visión artificial en Android (**Energy consumption analysis for computer vision applications under Android**), 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.

Abstract: Although the first smartphone appeared in 1994 with the so-called IBM Simon, smartphones had not really become popular until the first iPhone was released. Since then, smartphones have changed our needs and way of communicating each other. In fact, according to a survey from SurveyMonkey1 which was conducted in 2015 to more than 1000 people, smartphones batteries autonomy is the most requested feature to get improved, with the 33% of respondents. Therefore, it is necessary to carry out an analysis about which components or modules of a mobile phone have a higher energy expenditure.

This is the Final Degree Thesis main objective. Firstly, a smartphone has been divided into three modules: images sensing, image processing and communication between devices. For each one of the modules, different parameters will be varied in order to see how these parameters influence the power consumption. To do that, three Android applications (one per each module) have been developed to measure the power consumption by obtaining different battery statistics. Finally, based on the statistics from the mobile device, different models have been obtained through the Matlab tool, which will be used to compare the results.

Diseño de Redes de Cámaras Inteligentes utilizando Smartphones (Design of Smart Cameras 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.

Abstract: This final degree project presents a software tool to remotely control a smart multicamera network using Android smartphones. This network is composed by a homemade video-surveillance system which consists of three different modules: smartphones (cameras), server and client applications. This system will be implemented by taking advantage of Android possibilities when controlling different camera capturing configuration and sending images. The processing images tool OpenCV will also be used to visualize these pictures in the client application and apply a people detection algorithm when it's requested. The complex managing of the camera and user network must be treated differently, it will be done by a server, which will be the system center in charge of the communication protocol between users and cameras. Once the development is finished, experimental performance tests will be done with different camera capture configurations and varying the number of simultaneous cameras to come to the conclusion of the proposed architecture's limitations for this video-surveillance system.

Detección y Reconocimiento de Señales Viales (Traffic 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.

Abstract: Lately, a huge evolution has taken place within the automotive world. We can find many improvements in this area, being the search of the autonomous driving the most important one. Despite the fact that there still remain a lot of years of progress to encounter autonomous driving in the street, more and more vehicles are now equipped with systems providing some useful information to the driver. These pieces of information are taken from the road itself. Some new companies and specific departments within car brands have arisen in search of this very general target, creating systems that, in the ordinary usage, are able to perform autonomous driving. Besides, these systems are starting to be sold as ADAS (Advanced Driver Assistance Systems), allowing their testing and

upgrading in a real driving environment. One of the above said systems, called TSR (Traffic Sign Recognition), consists in the recognition of the detected traffic sign. This Bachelor Thesis aims to obtain a first implementation of an own algorithm of vertical traffic signs detection and recognition. This detection is not easy due to the changing conditions of the road and the circulation, such as weather or light changes or the difficulty to receive traffic signs from a moving car. Although other detection and recognition systems use more than one camera or sensor, we will be making use of just one camera, trying to create a system whose best quality be its adaptive capacity to the environment. This way, it could be included in a modern car software or in a Smartphone and then used from a car dashboard or a motorbike wind deflector. As I am aware of my limited experience concerning image processing, the following system aspires, not to stand out among the ones described in the State of the art section, but to propose a simple solution to this existent question.

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.

Abstract: The objective of this Final Degree Thesis is to create a system of augmented reality. To this aim, we combine the tools and concepts of computer vision, Machine Learning and the use of graphics. Together, the shape an application that can detect an object and generate augmented reality on it. After studying the different possibilities for the creation of applications of augmented reality, we have selected a technique without markers (Markerless). In order to fulfil our objectives, the design of the application has been arranged on a per stage basis: First, the detection of points of interest for the detection and description of objects. Second, the training of a support vector machine. Third, the use of computer vision techniques. Fourth, the use of a graphic engine. The model has been evaluated and the results are, in general, promising. However, we have found limitations when detecting the object when it's subjected to heavy rotations. To conclude, we believe that, in spite of limitations, the results of the developed applications are positive and motivate further research.

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.

Abstract: Nowadays, people detection systems are widely used in video surveillance applications. In spite of the existence of a variety of studies about improving the obtained results, this area continues being of great interest due to its relevance and complexity. This work is focused on accomplishing the

premise that these security systems must be robust in order to be able to detect people in every scene. Occasionally, the scenarios we find at the time of the analysis are really complex for a proper detection. There might arise a number of real and adverse situations, both indoors and outdoors, of lighting and shadows alterations or people partially occluded by objects. The main idea is using the scene context of each video to adapt the model in pursuance of getting a more accurate detection for the majority of the cases. The work has consisted in implementing through Matlab an adaptation of the ACF's algorithm (Aggregate Channel Features) for people detection, as well as in the study of the system's variable parameters to achieve the best results. At last, an evaluation of the algorithm with its different variations about the reference sequences has been done. The obtained results demonstrate the improvement in the detections for the analyzed sequences.

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.

Abstract: The objective of the project has been to develop an automatic web platform for the evaluation of algorithms for detecting people. The detection algorithms are currently booming, in this case the topic of people detection has been chosen. First, it has been studied of the platforms ChangeDetection.net (CDNET) and People Detection Benchmark repository (PDbm) of the Video Processing Understanding Lab (VPULab). Both platforms are oriented to the evaluation of algorithms. After studying both platforms, it was decided to automate PDbm based on CDNET. The design of the system has been made in four functional blocks: Web server, database, evaluation application and mail manager. In addition, we adopt the layered programming model on the web server. During the development of the project, several specific technologies have been implemented for each block: Windows 7, Apache, MySQL, PHP, Matlab, Outlook. Subsequently, a control module is implemented for the platform administrator. Then, platform is tested in a local pre-production space and verified to be working properly. Finally, we analyze the possibility of continuing to develop the platform in a modular way due to the design of the system. In conclusion, we have worked with applications of different types and dedicated to different purposes. It has been possible to unite them in order to form a unique system.

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.

Abstract: In this Bachelor's thesis we have tackled the task of reconstructing or initializing a background image using the LabGen-P algorithm and some proposed variants. This algorithm reconstructs the background using the temporal median of the pixels with less motion in time. This algorithm has been selected due to the good results obtained in a recent background initialization competition. Initially, we will study the state-of-the-art to fully understand the different challenges that background initialization poses and the techniques used in the literature to address them. Subsequently, we will implement the LabGen-P algorithm, which has high quality results, in order to understand its strengths and weaknesses. Then, we will introduce variations in the algorithm to deal with its issues. Among these variations we can find the use of an improved motion estimation and the inclusion of spatial information based on segmented image regions sizes. Finally, we will evaluate the base algorithm and the proposed variations to determine the best configuration.

Detección de Personas mediante Redes Convolucionales (People Detection using Convolutional 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.

Abstract: Nowadays we live in a period in which the people detection is on the rise, being very important in some aspects of society like video security. As a matter of fact, many researchers create their own algorithms for object detection obtaining excellent results over databases chosen by themselves. But the problem arises, mainly, at the time of generating a model that includes as much characteristics as possible, to make it capable of overcoming changes in posture, movement, illumination, the interaction among people or changes in the point of view. To train the models, is becoming more popular the usage of a new kind of learning, known as Deep Learning, this handles huge amounts of data that solves those problems trying to simulate the behaviour of neurons of the human brain. New algorithms are based in Deep Learning for people detection, which are named as detectors based on convolutional networks. With all this being said, the aim of this final degree project is to offer a comparative between the traditional algorithms of people detection (Deformable Parts Model, DPM, and Aggregate Channel Features, ACF) and the modern ones based on convolutional networks (Faster Region-based Convolutional Network, FASTER R-CNN), studied in the state of the art. For the achievements of this, different models of people will be designed and implemented, and the obtained results will be evaluated by the chosen detectors over a database and metrics in common, in equal conditions.

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.

Abstract: This work aims to develop an object-background segmenter capable of generating segmentation masks where the images can be divided into the element to be segregated and the remaining ones. This has been achieved through the generalization of the previous work from which we proceed, the person-background segmenter, whose main feature is the use of detection in order to generate such masks. The difference to be highlighted between the two segmenters is the model they use for detection. In the person case, the model is always the same with a single pose, while the object-background segmenter can use any model with an indeterminate number of poses. To accomplish our system, two application procedures will be analyzed in this work: one method in which the independent parts of the object model can be detected, and another one in which these parts could be combined. Finally, the results of both methods will be put to the test, and the system will be assessed. To obtain the tests and validation of our algorithm it is necessary to conduct a selection of sequences of images or datasets and of a set of object models to detect whether they are useful for their application to our system or not.

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.

Abstract: The objective of this TFG is how to use frequency information in a sequence of images to be able to automatically identify materials. In addition, it also seeks to study the impact of adding images with infrared light and depth information (other than RGB images) in the development of this task. To do this, we start by creating a new dataset with the help of the Kinect sensor, composed of RGB, depth and infrared images. The recorded sequences are divided into two parts: one part for training and another part for test. Those intended for the training of the model are manually labeled to separate the materials from each other. Once we have the materials separated, we generate descriptors of this information using the Fourier transform and create knowledge models from these data, using the probabilistic model Gaussian Mixture Models. In conclusion, we have a model for each material. In the evaluation part, we study the goodness of our models using the sequences assigned to the evaluation part. The Fourier transform of the complete sequence is performed and introduced in each of the generated models, so that it returns the probability that each pixel has to belong to an object. Possible modifications that can be made to descriptors to improve identification are also studied. Finally, we collect the conclusions of this work and mention the possible future work in relation to the chosen task.

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.

Abstract: The aim of this end-of-degree project is to obtain the segmentation of the contextual objects that compose a video. The general implementation will consist of transferring masks of contextual objects from a database to a not observed image. For this purpose, importance will be attached to the construction of data-sets that will be used in each of the different scenes to be analyzed. The dissertation will consist of studying all the information given by an image of entry through two distinguished models with the aim to extract every pixel's probability of belonging to one of the two labels that are established in this work: "context" or "not context". Through a model of energies, the final segmentation will be obtained, determining which of the pixels on the image correspond with a contextual object. Finally, the contextual segmentation for a video will be acquired by averaging the masks obtained for the different frames that compose it.

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.

Abstract: This final degree thesis has as main goal, the video-tracking of objects based in their color and shape features. This requires the use of a tracking algorithm, that locate the target over time, given his initial position, and estimates his later positions. In this FDT is going to be use the classic particle filter based in their color features of the RGB space. The first step, was to make the current algorithm compatible with another color space, and HSV was chosen as alternative. Subsequently, took place the implementation of two papers and in order to add an additional feature to the initial particle Filter, extracting more information about the target. One feature that complements color information very well is orientation, so we extract that information using gradient histograms. This FDT includes a novel technique that benefits from the position of the generated particles too. Apart from detecting the target, it extracts surrounding area information, discriminating against what corresponds to the target and what corresponds to the image background. Particles that provide information about the target will have more importance that those that provide information about the background. Once we have the result of the tracker, it will be compared with manual annotations about target's position (ground-truth) through evaluation metrics, where effectiveness is provided by the overlapping area between the tracker result and the ground-truth. Finally, the results thrown by the initial

particle Filter, based in color histograms; and the particle filter modified to extract additional information about target's shape and orientation, as well as its surrounding information, will be compared. It will be possible to see slight improvement in the results, because for the more information has the tracker about the target, the better the tracker will detect the target and estimate future positions.

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.

Abstract: This Bachelor Thesis aims to study and develop different techniques for de analysis of video sequences in a surveillance environment in order to detect anomalies present in the scene, and therefore build a complete system to carry out this task. Based on the literature and other papers about anomaly detection in video, the detector has been designed from blocks, and every block has been analyzed separately. In this design, there has been deepened specially in the feature choice and extraction, being decided in this thesis for the spatio-temporal features of the intensity gradients. From these, it has been established a training and learning method, and two algorithms for modeling the scene behavior have been implemented and analyzed, evolving from a distance based system to detect anomalies until one based on the amount of membership to each class which are K-Means and Fuzzy C-Means. The advantages and disadvantages of each method are analyzed and on adequate adjustment is found throughout a series of predefined tests. The analysis of results is done with real video surveillance sequences from different environments and situations, for every sequence and every model proposed, thus following a configuration from particularity to generality, motivated by the small similarity between each of the sequences. In order to finish the thesis, it is verified if the marked objectives have been reached and it is tried to extract conclusions about the features and algorithms used and implemented, as well as propose some possible future improvements for the system.

The HAVideo (TEC2014-53176-R) project is supported by

