

Semantic characterization, indexing and universal customized access to audiovisual content under controlled environments TIN2004-07860-C02

Jesús Bescós*
Universidad Autónoma de Madrid

Fernando Jaureguizar**
Universidad Politécnica de Madrid

Abstract

The main objective of this project is to research in the design of algorithms and the development of tools both oriented to automatic feature extraction in image sequences of controlled environments, additionally the further indexing, storage and customized access to the previously extracted and properly coded information out of those sequences. The project consists of to main stages: Information acquisition and Management of that information to offer services to the final user. The information acquisition out of images from several cameras performs the automatic extraction of features of people present in those images, segmenting in real time the individual persons out of the background, while building a 3D model of the scene thanks to self-calibration techniques. Finally, this information is described according to the MPEG-7 standard to allow for efficient access and storage. The second stage is mainly concerned with the management of the obtained information for its use in the development of applications. Its design is based on the user and terminal/network profiles (MPEG-21 DIA) and on the management of alert levels/options for the delivery of information relevant to each user. According to these profiles, the available coders with the proper parameters are run for the purpose of adaptation and customization of the information during each session.

Keywords: Autocalibration, segmentation, video analysis, compressed domain analysis, content adaptation, 3D modeling, natural coding, synthetic and hybrid coding,

1 Project objectives

1.1 Introduction

This project proposes the joint development of series of research fields, independent but complementary, intended to promote technology for automatic indexing and personalized delivery and retrieval of audiovisual content.

In order to establish a ser of services which enable querying, search, retrieval, adapted transmission, distribution, personalization, etc., in an information system, with the aim of

* Email: j.bescos@uam.es

** Email: fjn@gti.ssr.upm.es

generating added value and maximize its fields of use, the availability the precise description of the available information, which is commonly known as meta-information or meta-data, play a fundamental role.

There are plenty of initiatives intended to automatically perform automatic description of audiovisual content. However, the high complexity of succeeding in the general case, motivates that current research is focused on specific situation or controlled application domains (e.g., news, soccer, etc.). This project targets a specific domain (static cameras, almost controlled illumination, predefined relevant objects, etc.) which is bounded enough to achieve successful results with practical application, and generic enough to export its achievements to a wide range of domains.

The particular research fields that have been considered to reach a common objective include autocalibration techniques, detailed analysis of image sequences, 3D modeling of such sequences, description and indexing of the content (MPEG7), personalized and adapted access to such information according to network, terminal and user profiles (MPEG21-DIA), and final coding and transmission (MPEG4, H.264).

1.2 Specific objectives

This subsection provides a detailed description of the list of planned objectives:

- 0 Enhancement of present autocalibration techniques, mainly focusing on their robustness and precision. Camera calibration is the base for 3D scene modeling and for splitting the background and the mobile objects, but allowing a later integration of these on the background via description of their geometric relationships.
- 0 Integration of segmentation information into the coding strategy: discriminating between background and relevant objects, scalable coding analysis in several coding schemes (hybrid, wavelets) and integration of developed strategies in video coding standards (JPEG-2000, MPEG-4, H.264).
- 0 Efficient and hierarchical coding of synthetic 3D scenes (previously modeled as a static background and several mobile objects), and integration of hybrid synthetic and natural contents.
- 0 Development of algorithms for real-time foreground/background segmentation, based on compressed domain analysis, as a first step for a later refinement of the segmented regions.
- 0 Real-time feature extraction from sequences captured from the cameras in order to obtain the basic information about the sequence contents. These are the starting point for more advanced feature extraction algorithms.
- 0 Description and indexing of the segmented elements vide meta-data coded according to the MPEG-7 standard.
- 0 Development of real-time transmission control algorithms for maximizing the quality of the decoded information in the user terminal.
- 0 Information management, personalization and adaptation: user, network and terminal profiles management, transcoding and transmoding management, translation from MPEG7 descriptions, coding management of the adapted media, handling of alert criteria, and content synthesis from content descriptions.

1.3 Chronogram and resources

The Project is structured into five work packages: the first focused on establishing the working framework, three more specifically targeting each of the general fields of research considered, and a last one aimed to show practical applications of the achievements. The following Table summarizes the list of activities, with an indication of their current status.

Activities/Tasks	Start	End	Partner	Status
P.1: Framework development				
<i>T.1.1:</i> Environment setup	m01	m06	ALL	Ongoing
<i>T.1.2:</i> Information model	m04	m06	ALL	Finished
P.2: Advanced image sequences processing				
<i>T.2.1:</i> Calibration and capturing system setup	m06	m18	UPM	Finished
<i>T.2.2:</i> Real-time segmentation in the transform domain	m08	m20	UAM	Finished
<i>T.2.3:</i> Segmented objects analysis in the transform domain	m20	m32	UAM	Ongoing
<i>T.2.4:</i> Segmentation in the pixel domain	m08	m20	UPM	Finished
<i>T.2.5:</i> Segmented objects analysis in the pixel domain	m08	m22	UPM	Finished
<i>T.2.6:</i> 3D scene modeling and management	m09	m27	UPM	Ongoing
P.3: Semantic indexing, storage, retrieval and personalization				
<i>T.3.1:</i> Segmented objects description (MPEG-7 + contents)	m06	m18	UAM	Ongoing
<i>T.3.2:</i> Efficient segmented objects storing	m18	m24	UAM	Ongoing
<i>T.3.3:</i> Information management, personalization and adaptation	m06	m30	UAM	Ongoing
P.4: Coding and transmission control for network and client adaptation				
<i>T.4.1:</i> Natural coding	m09	m30	UPM	Ongoing
<i>T.4.2:</i> Synthetic and hybrid coding	m09	m30	UPM	Ongoing
<i>T.4.3:</i> Transmission control	m10	m33	UPM	Ongoing
P.5: Application to remote monitoring services				
<i>T.5.1:</i> Design	m09	m27	ALL	Ongoing
<i>T.5.1:</i> Integration and trials	m24	m36	ALL	

The funding of the project has made available to the participant entities the necessary resources to conduct the research, mainly computers and video cameras have been installed in the participant entities, allowing for the developments of the members of the research team as well as of the students working on their master thesis and PhD Thesis integrated in the project.

2 Project achievements

2.1 Environment setup

At the UPM, initial system design of cameras in public indoor and outdoor places was not possible because the request for their installation was rejected due to legal issues about privacy. For the moment, a set of PTZ cameras (IP connected) has been installed in several parts of the ceiling of a students Laboratory (indoor semi controlled environment) in order to carry out the sequences acquisition for the project developments.

At the UAM the installation of the acquisition systems has been scheduled in three phases, corresponding to the three years of the project, in order to benefit from technology advances in the area. The first year installation comprises three high resolution (1024x768 YUV 4:2:2 at 15 fps) color cameras located in two access entrances to the EPS Building; they provide raw video which is delivered via optic fiber to our laboratory. The second year installation comprises three PTZ IP cameras, all located in the EPS Hall.

A main part of the environment consists of a framework for distributed real-time processing of video information, mainly designed to facilitate the testing of interrelated algorithms.

2.2 Advanced image sequences processing

2.2.1 Calibration and capturing system setup

The research conducted in autocalibration has been concentrated in two areas: the Absolute Line Quadric based algorithms and two cameras based algorithms.

In the first area the concept of Absolute Line Quadric has been introduced. This is the geometrical object that represents the set of lines intersecting the absolute conic, and allows for the obtainment of new autocalibration algorithms, in the same way as it was formerly done with the Absolute Quadric. The new algorithms [1][2][4][5] provide the direct solution of the autocalibration problem for ten or more images taken with square-pixel cameras with arbitrarily varying internal and external parameters.

In the second area the case of two square-pixel cameras with constant parameters has been studied, producing new results of practical interest [3], such as a explicit parameterization of the set of solutions that allows for the fast computation of the Euclidean 3D reconstruction using a single piece of knowledge of the scene, such as the parallelism or the orthogonality of two lines.

2.2.2 Segmentation and object analysis in the compressed domain

These activities confront a same objective with two different approaches, both based on compressed domain analysis. The first starts from uncompressed image sequences and performs a block-DCT image-by-image transform prior to applying conveniently modified typical segmentation approaches. The second starts from compressed video sequences (currently MPEG1 and MPEG2, and now including MPEG4); some advantages are the direct availability of estimators for features that are hard to extract at a pixel level (e.g., the motion field), and a dramatic reduction of the dimensionality (e.g., by working over DC images); the main drawback is that these techniques are highly dependent on the coding standard and on the coding parameters.

We currently apply the first approach just to static camera conditions. The method for foreground detection introduces a context-aware background subtraction method as its main innovation. In the current implementation, temporal change detection represents this context information. Assuming a non-complex background, common in many indoor surveillance applications, the method is able to work in real-time while adapting to smooth illumination changes. Results[7] show a good compromise between efficiency, accurate segmentation and robustness, being able to work with high and low frame rates while complying with real-time requirements (over 300 fps QCIF).

The second approach focuses more directly on a practical situation, as most commercial security cameras trend to output MPEG4 video. Part of the situation is similar to the previous approach, as DCT coefficients can be exactly obtained for any intra-coded macroblock. If not intra-coded, obtaining the exact value of a coefficient requires in general four block IDCTs and one block DCT; however, this can be avoided if we assume to work with an estimation of the coefficients which can be efficiently obtained. However, in the MPEG domain also a sparse motion field can be estimated from the motion estimation vectors, which are highly coding dependent. Our current effort[8][9] is directed to use this motion information to estimate the camera motion (from PTZ cameras) in order to compensate it and then apply segmentation techniques for static camera situations.

2.2.3 Segmentation and object analysis in the pixel domain

The work carried out addressed mainly the development of a set of tools which implement different approaches to the following problems: extraction of static backgrounds (BG), identifying moving objects, determination of foreground (FG) static elements of the scene, and tracking of moving objects for occlusion management and moving objects parts characterization. Researched technical solutions followed a complexity incremental approach, starting from the simplest ones to the more sophisticated –and accurate/robust- (although their real-time implementation may likely require the use of specific hardware processing elements).

Several strategies have been considered, implemented and compared for quasi-static BG/FG characterization and moving objects detection. Discrimination between non static and static (including some non-strictly static) elements has been based on a combination of several Gaussians for each pixel to model the distribution of its gray level value along the sequence carried out on sub-sampled versions of the filtered inputs. It offers a good trade-off between computational costs, adaptability to slow illumination changes, and robustness with respect to the noise. A new segmentation strategy [10] overcomes the problem of automatic detection of static elements in FG (moving object partially occluded by them). Tracking information through a Kalman filter is used to refine the different areas segmentation. Segmentation of different elements within the scene has been used to drive efficient coding strategies, based on JPEG2000 ROIs and on H.264 FMD [11][12].

Tracking is incorporated to the moving objects segmentation and static elements (BG and FG) identification to achieve: occlusion management at moving object level, and independent sub-region tracking for divided objects. Simplified implementations based on graphs theory were very efficient and could cope with simple partial occlusions when moving objects bounding boxes are tracked. Bounding box size evolution is also used to indicate possible separations or joins of independent moving objects which go together. Very efficient implementations have been carried out based on look-up tables. Kalman filtering have been incorporated to the developed techniques, but demonstrated that the simplification of using bounding boxes or centroids of the moving objects as tracking features, is only operative in those situations where the moving object size is small with respect to the observed area, and objects follow quasi-linear motion evolution. Other developed works incorporate relevant object information to the tracking: motion information (motion vectors), object gray level (color is not considered for the computational cost), and some texture indicators. Moving object regions are characterized by those parameters, which constitute representative feature vectors which are projected on the following image using Kalman filters. This complex strategy is only required when the graphs based approach do not produce an

adequate trajectory filtering, and introduces a sub-segmentation of the moving objects into joined regions, with independent motions, that need to be tracked. Therefore, the proposed final strategy is a strategy combining three subsystems: one based on graphs theory, another based on regions (including sub-segmentation of moving objects), and a third one which deals with partial observations and trajectory analysis.

2.2.4 3D scene modeling and management

MPEG-4 and, in particular, its Part 16, AFX (Animation Framework eXtension) for synthetic content, were chosen for 3D scene representation. They provide tools for the description of hierarchically structured, hybrid 3D scenes. Visual objects of an MPEG-4 scene may be grouped to form pyramidal structures to ease their hierarchical coding, transmission and rendering; and may have different nature and characteristics, as is the case of traditional, rectangular-shaped video coming from a natural source vs. synthetic 3D graphics, but also of passive/static BG vs. active/changing FG. These two characteristics, together with the scalable coding features of the MPEG-4 and AFX tools, enable flexible 3D scene modeling and management. More specifically, they provide the required handles for automatic content adaptation based on terminal, network and user profiles. Our work in the field of automatic content adaptation has led to several reports [13] [14].

Besides, researching on synthetic content generation from high-level descriptions or by means of procedural techniques has been conducted[15][16]. The expected outcome is a very efficient method for describing “fractalish” objects (textures, mountains, trees, etc.) in the scene BG, as well as for creating or modifying some active objects, such as virtual humanoids.

2.3 *Semantic indexation, storage, retrieval and personalization*

This task has two main outputs: generation and creation of different representations of a video surveillance scene with different levels of detail, and a framework for content adaptation. On one side, the application of technologies of generation and transmission of descriptions, allow to reduce transmission and storage requirements as well as to provide descriptions capable of helping the automatic interpretation and triggering of alarms. On the other side, the proposed content adaptation framework provides a generic and extensible adaptation engine working with MPEG-21 and MPEG-7 descriptions for both usage context description and content description.

2.3.1 Segmented objects description (MPEG-7 + content)

This subtask focuses on the generation of descriptions (following the MPEG-7 standard) in order to provide on-line adaptive transmission of video sequence allowing to send the most relevant information and conditions of use (in exchange for greater processing resources both in the sender and receiver terminals, for analysis and synthesis, respectively).

The input to this subtask is a segmentation mask coming from the analysis modules. This segmentation masks indicate the moving regions detected. The original image is also need for further description extraction.

The descriptions of moving objects can be generated at different levels of detail, ranging from the trajectory of the centroid of the moving region to the description of trajectory of the exact object shape, and including bounding boxes trajectory and associated textures, among others. Depending

The work in this subtask has focused on the creation of a generic and extensible content adaptation engine named CAIN (Content Adaptation INtegrator)[18][19]. CAIN constitutes a complete system for multimedia content adaptation. CAIN Works with MPEG-7 descriptions for content description and MPEG-21 DIA for usage environment description. Fig.1 depicts the CAIN architecture that, besides the input/output facilities and additional support modules, can be divided in three main functional blocks.

- o The Decision Module (DM)[21] is on charge of the making the necessary decisions for performing the adaptations. It returns the Content Adaptation Tool (CAT) to be executed and the input parameters to use in the call. For getting these results the DM will take as inputs the different CATs' capabilities, the network and terminal restrictions, and the user's preferences. This module could be considered as the one which contribution is the intelligence of the system.
- o The Execution Module (EM) is a pretty simple module which aim is to execute the method responsible of the adaptation, in the indicated CAT and with the received input parameters. The decision has already been taken by the DM. In future version this module may be more complex, when dynamic adaptation (based on changing usage environment conditions) is introduced.
- o CATs "battery", that is, the set of CATs made available to CAIN thanks to the definition of an extensibility mechanism[20].

Currently we have integrated CATs belonging to the four categories considered in the design of CAIN: Transcoder CATs, Scalable Content CATs, Real-time Content Driven CATs and Transcoding CATs.

2.4 Coding and transmission control for network and client adaptation

2.4.1 Natural coding

Improvement of coding efficiency to achieve real-time coding strategies based on the development of Fast Mode Decision (FMD) strategies in the framework of H.264 is successful in the project. Two different research approaches have been followed to drive FMD: incorporation of the segmentation process results, and incorporation of thresholded distortion measures regardless of the contents of the video sequences.

The former is useful in our case because an analysis in terms of object segmentation is already available. Its target is to exploit the knowledge on the scene to reduce motion estimation computation to achieve real time coding for transmission and efficient storage. The knowledge of the moving object areas and BG areas allows to reduce the huge amount of modes permitted by H.264. A sub-set of modes is selected to obtain a more accurately motion estimation for active objects and less intensive for quasi-static BG, reducing the number of comparisons needed to perform inter prediction by 60%-70%, while keeping the same image quality and bit-rate. This work was successfully published in [11]. If BG/moving object segmentation is not available, a fast segmentation strategy aiming to detect activity regions in the images was used to drive the mode selection process. This research work [12] complements the previous one, while still exploiting some knowledge about the motion and activity present in the sequences.

Other developed alternatives have as main target to keep rate-distortion measures very close to those obtained with the full exploration of all the possible modes considered in H.264, while reducing dramatically the number of modes to be computed. These works explored predictions in terms of distortion and rate-distortion values computing adaptive thresholds to stop mode selection when quality improvement vs. computational cost is considered perceptually non relevant. Developed work [22][23] successfully improve available FMD strategies in H.264.

2.4.2 Synthetic and hybrid coding

MPEG-4 and AFX provide tools for the efficient hierarchical coding of 3D scenes, in that there are perfectly fit for streaming applications (unlike, e.g., VRML.97). A clear example of this is the WSS (Wavelet Subdivision Surface) tool of AFX, which permits to model, code and transmit a 3D shape in a progressive manner: a coarse base mesh is extracted from a given fine mesh with arbitrary connectivity, representing the target surface, thanks to an automatic simplification algorithm; this base mesh is taken as the lowest LOD (Level Of Detail) of a collection of hierarchically nested meshes, obtained by recursively applying a subdivision algorithm, and approximating at increasing LOD the target surface; finally, the 3D details necessary to go from one LOD to the next are efficiently coded thanks to wavelet techniques, to yield a fully embedded bit-stream.

Bit-streams of MPEG-4's WSS tool is only SNR-scalable, as opposed to also spatially scalable: it cannot be easily parsed according to a given maximum resolution (i.e., number of subdivisions or LODs) imposed by the decoder. This is not desirable in environments with mobile platforms (PDAs, cell phones), since they have much less resources (computational power and memory) than fixed ones (PCs). We have devised a novel wavelet-based method for the hierarchical coding of 3D surfaces which features both SNR and spatial scalability, and which has much better compression efficiency than MPEG-4's WSS tool. The work has led to a publication [24] and several proposals in MPEG, which will likely adopt our PLTW (Progressive Lower Trees of Wavelet coefficients) tool in a future edition of AFX. Another limitation of MPEG-4's WSS tool is that there are several restrictions for the attributes (normal vectors, colors, texture coordinates, etc.) that can be attached to the vertices or facets of a 3D mesh. We have proposed as an alternative the IWS (Indexed Wedge Set) format, which has also led to MPEG contributions.

2.4.3 Transmission control

The framework proposed to compute optimized rate control strategies is based on the analysis of the characteristics of the main elements involved in the control problem: the source of coded information and the transmission channel. Special attention has been given to wireless channels, since its inherent stochastic behavior poses difficulty to the transmission of real-time content.

On the one hand, different stochastic packet-loss models have been developed for UMTS and Wi-Fi channels. More specifically, for UMTS Gilbert-Elliott and Fritchman models have been analyzed and a new model based on Hidden Markov chains have been extended to capture the behavior of UMTS channels. Regarding Wi-Fi, Gilbert-Elliott model has been analyzed and new models based on both Markov chains and Hidden Markov chains have been designed and implemented.

With regard to source models, our previous work on R-D models (based on an H.263 codec and for I and P frames) has been extended to an MPEG-4 codec and for I, P and B frames. With these models, from a single rate measurement (at a particular quantizer) on a given frame, it is possible to estimate any other rate and distortion value (at any other quantizer).

Finally, a basic rate control strategy has been developed for the transmission of video content in a Wi-Fi network. This preliminary approach has been focused on increasing the robustness of video transmission in a multicast scenario.

3 Relevance indicators

3.1 *Educational Training*

3.1.1 Master thesis

- o UPM: Ainhoa Mendizábal (Jan. 2005), Laure-Hélène Caron and David Fougerousse (Feb. 2005), Alberto Castillo (Jul. 2005), Cristina Calleja (Sep. 2005), Sergio Sanz (Sep. 2005), José M^a Cobo (Sep. 2005), Sebastián Amengual (Oct. 2005), Laurent Isenegger (Nov. 2005), Marcos Nieto (Nov. 2005), Carlos Cuevas (Apr. 2006), Raquel Ribagorda (Jul. 2006).
- o UAM: J. Carlos San Miguel (Sep. 2006).

3.1.2 PhD. Students

- o With the Advanced Studies Diploma (DEA): Luis Herranz (UAM, Jul. 2005), Fernando López (UAM, Jul. 2006), Víctor Fernández-Carvajal (UAM, Jul. 2006), Enrique Estalayo (UPM, Sep. 2006), Víctor Valdés (UAM, Sep. 2006), Fabricio Tiburzi (UAM, Sep. 2006).
- o 2nd year finished: José Cubero (UPM), Javier Molina (UAM), Daniel Márquez (UAM).
- o 1st year finished: Carlos del Blanco (UPM), Marcos Nieto (UPM), Juan Carlos Plaza (UPM), Pablo García (UPM), Álvaro García (UAM).
- o Starting: Carlos Cuevas (UPM), J. Carlos San Miguel (UAM).

3.2 *Publications*

3.2.1 Journal papers

- o On autocalibration: [1][2][3]
- o On automatic graphical content adaptation: [14]

3.2.2 Conference papers

- o On autocalibration: [4][5][6]
- o On segmentation and object analysis in the compressed domain: [7][8][9]
- o On segmentation and tracking in the pixel domain: [10][11][12]
- o On automatic graphical content adaptation: [13]
- o On content generation from high-level descriptions: [15][16]
- o On personalization and audiovisual content adaptation: [18][19][20][21]
- o On natural coding: [22][23]
- o On 3D surface coding with the PLTW technique: [24][25]

3.3 *Regular collaboration with companies and EU projects*

Regular contacts are maintained with companies in the sector such as Check-Out Control and Infoglobal (this via a long-term established collaboration), both interested in developments mainly related to their application for video-surveillance non-supervised systems.

Members of the research team at UPM have participated in the EU projects VISIRE (in calibration, tracking and 3D reconstruction) and OLGA (3D modeling and synthetic content and

coding); and in national project MOSSLADAR (in segmentation and tracking). As a result of part of the research work, the research team elaborated successful proposals for FP6, 36 month project I-WAY (in segmentation and tracking) and 27 month project SUIT (in video transmission and rate control) both started in February 2006. Additionally, the research team at UAM are currently involved in two FP6 EU projects (aceMedia and MESH), leading activities related to audiovisual content analysis and adaptation.

3.4 Collaboration with other research groups

Contacts have been established with national research groups in image processing/analysis/coding in Spain: GPI at Univ. Polit. Cataluña, CeTVD at Univ. Ramón Llull, GTS at Univ. Polit. Valencia, G2PI at Univ. Carlos III, GCT at Univ. Rey Juan Carlos. Due to the current research in EU projects, fluent relationship has also been established with European research groups in image processing/analysis/coding and transmission: IT at Univ. Aveiro (PT), Dep. of Computer Science at Univ. Ioannina (GR), CCSR at Univ. Surrey (UK), IBBT at Univ. Ghent (BE), ETRO at Vrije Universiteit Brussel (BE), DESICS Dept. at IMEC (BE), ARTEMIS Unit at INT (FR), MVLab at Queen Mary Univ. of London (UK), IVML at NT Uni. Athens (GR).

4 References

- [1] A. Valdés and J.I. Ronda, "Camera autocalibration and the calibration pencil", *Journal of Mathematical Imaging and Vision*, vol. 23, no. 2, pp. 167-174, Sep. 2005
- [2] A. Valdés, J.I. Ronda, and G. Gallego, "The absolute line quadric and camera autocalibration", *International Journal of Computer Vision*, vol. 66, no. 3, pp. 283-303, Mar. 2006.
- [3] J.I. Ronda and A. Valdés, "Conic geometry and autocalibration from two images", accepted conditionally to a minor revision by the *Journal of Mathematical Imaging and Vision*, 2007.
- [4] J.I. Ronda, A. Valdés, and G. Gallego, "Camera autocalibration using Plücker coordinates", *Proc. IEEE Int. Conf. on Image Processing ICIP 2005*, vol. 3, pp. 800-803, Genoa, Italy, Sept. 2005.
- [5] G. Gallego, J.I. Ronda, and A. Valdés, "Recursive camera autocalibration with the Kalman filter", submitted to *IEEE Int. Conf. on Image Processing ICIP 2007*.
- [6] P. Pérez and N. García, "Robust and accurate registration of images with unknown relative orientation and exposure", *Proc. IEEE Int. Conf. on Image Processing ICIP 2005*, vol. 3, pp. 1104-1107, Genoa, Italy, Sept. 2005.
- [7] A. García, J. Bescós, "Real-time video foreground extraction based on content-aware background subtraction", submitted to *IEEE Int. Conf. on Image Processing ICIP 2007*.
- [8] L. Herranz, J. Bescós, "Reliability based optical flow estimation from mpeg compressed data", *Proc. International Workshop on Very Low Bit-rate Video*, 4 pags. (CD-ROM), Sardinia (Italia), Sept. 2005.
- [9] F. Tiburci, J. Bescós, "Camera Motion Analysis in On-line MPEG Sequences", submitted to the *Int. Workshop on Image Analysis for Multimedia Interactive Services, WLAMIS 2007*.
- [10] L. Isenegger, L. Salgado, and N. García, "Moving objects segmentation based on automatic foreground/background identification of static elements", *Proc. Advanced Concepts for Intelligent Vision Systems ACIVS 2005*, vol. LNCS 3708, pp. 491-498, 2005.
- [11] M. Nieto, L. Salgado, and J. Cabrera, "Fast Mode Decision and motion estimation with object segmentation in H.264/AVC encoding", *Proc. Advanced Concepts for Intelligent Vision Systems ACIVS 2005*, vol. LNCS 3708, pp. 571-578, 2005.

- [12] M. Nieto, L. Salgado, and N. García, "Fast mode decision based on activity segmentation in H. 264/AVC encoding", *VLBV 2005*, vol. LNCS 3893, pp. 146-152, 2006.
- [13] L. Salgado, J. Bescós, F. Morán, J. Cabrera, E. Estalayo, and J. Cubero, "Efficient Architecture for Low-Cost Adaptive 3D Graphics-Based E-Learning Applications", *Proc. EuroGraphics EG 2005, Computer Graphics Forum*, vol. 24, no. 3, pp. 41-50, Dublin, IE, Aug. 2005.
- [14] F. Morán, M. Preda, G. Lafruit, P. Villegas, and R.-P. Berretty, "3D Game Content Distributed Adaptation in Heterogeneous Environments", accepted conditionally to a minor revision by the *EURASIP Journal on Applied Signal Processing* (special issue on Video Adaptation in Heterogeneous Environments).
- [15] V. Fernández-Carbajales, J.M. Martínez, and F. Morán, "Description-Driven Generation of 3D Humanoids Within Authoring 744", *Proc. IEE European Workshop on the Integration of knowledge, semantics and digital Media Technologies EWIMT 2005*, pp. 227-231, London, UK, Nov. 2005.
- [16] V. Fernández-Carbajales, J.M. Martínez, and F. Morán, "High-Level Description Tools for Humanoids", *Proc. intl. workshop on Multimedia content Representation, Classification and Security MCRS 2006*, vol. 4105, pp. 387-394, Istanbul, TR, Sep. 2006.
- [17] J.C. San Miguel, J. M. Martínez, "On-line adaptive video sequence transmission based on generation and transmission of descriptors", submitted to *IEEE Int. Conf. on Image Processing ICIP 2007*.
- [18] V. Valdés, J.M. Martínez, "Content Adaptation Tools in the CAIN framework", *Visual Content Processing and Representation*, Lecture Notes in Computer Science, Vol. 3893, 2006.
- [19] J.M. Martínez, V. Valdés, J. Bescós, L. Herranz, "Introducing CAIN: a Metadata-driven Content Adaptation Manager Integrating Heterogeneous Content Adaptation Tools", *Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services, WLAMIS'2005*, Apr. 2005
- [20] V. Valdés, J.M. Martínez, "Content Adaptation Capabilities Description Tool for Supporting Extensibility in the CAIN Framework", *Multimedia Content Representation, Classification and Security-MCRS2006*, Lecture Notes in Computer Science, 2006.
- [21] F. López, J.M. Martínez, V. Valdés, "Multimedia Content Adaptation within the CAIN framework via Constraints Satisfaction and Optimization", *Adaptive Multimedia Retrieval-AMR06*, Lecture Notes in Computer Science, 2006.
- [22] L. Salgado, and M. Nieto, "Sequence independent very fast mode decision algorithm on H.264/AVC baseline profile", *Proc. IEEE Int. Conf. on Image Processing ICIP 2006*, vol.1, pp. 41-44, Oct. 2006.
- [23] M. Nieto, L. Salgado, and J. Cabrera, "Fast Mode Decision on H.264/AVC main profile encoding based on PSNR predictions", *Proc. IEEE Int. Conf on Image Processing ICIP 2006*, pp. 49-52, Oct. 2006.
- [24] M. Avilés, F. Morán and N. García, "Progressive Lower Trees of Wavelet Coefficients: Efficient Spatial and SNR Scalable Coding of 3D Models", *Proc. Pacific-rim Conf. on Multimedia PCM 2005*, vol. LNCS 3767, pp. 61-72, Jeju, KR, Nov. 2005.
- [25] M. Avilés and F. Morán, "Spatial and SNR Scalable Coding of Wavelet Subdivision Surfaces", *MPEG contribution M12329*, Poznań, PL, Jul. 2005.