EURASIP Journal on Applied Signal Processing 2003:9, 861–877

MPEG-4 Authoring Tool Using Moving Object Segmentation and Tracking in Video Shots

Petros Daras

Electrical and Computer Engineering Department, The Polytechnic Faculty, Aristotle University of Thessaloniki, Gr-54124 Thessaloniki, Greece

Informatics and Telematics Institute (ITI), 1st Km Thermi-Panorama Road, Gr-57001 Thermi-Thessaloniki, P.O. Box 361, Greece
Email: daras@iti.gr

Ioannis Kompatsiaris

Informatics and Telematics Institute (ITI), 1st Km Thermi-Panorama Road, Gr-57001 Thermi-Thessaloniki, P.O. Box 361, Greece
Email: ikom@iti.gr

Ilias Grinias

Computer Science Department, University of Crete, Gr-71409 Heraklion, P.O. Box 2208, Greece
Email: grinias@csd.uch.gr

Georgios Akrivas

School of Electrical and Computer Engineering, National Technical University of Athens, Gr-15773 Athens, Greece
Email: gakrivas@image.ntua.gr

Georgios Tziritas

Computer Science Department, University of Crete, Gr-71409 Heraklion, P.O. Box 2208, Greece
Email: tziritas@csd.uoc.gr

Stefanos Kollias

School of Electrical and Computer Engineering, National Technical University of Athens, Gr-15773 Athens, Greece
Email: stefanos@softlab.ntua.gr

Michael G. Strintzis

Electrical and Computer Engineering Department, Aristotle University of Thessaloniki, Gr-54124 Thessaloniki, Greece
Email: strintzi@eng.auth.gr

Informatics and Telematics Institute (ITI), 1st Km Thermi-Panorama Road, GR-57001 Thermi-Thessaloniki, P.O. Box 361, Greece

Received 29 April 2002 and in revised form 22 November 2002

An Authoring tool for the MPEG-4 multimedia standard integrated with image sequence analysis algorithms is described. MPEG-4 offers numerous capabilities and is expected to be the future standard for multimedia applications. However, the implementation of these capabilities requires a complex authoring process, employing many different competencies from image sequence analysis and encoding of audio/visual/BIFS to the implementation of different delivery scenarios: local access on CD/DVD-ROM, Internet, or broadcast. However powerful the technologies underlying multimedia computing are, the success of these systems depends on their ease of authoring. In this paper, a novel Authoring tool fully exploiting the object-based coding and 3D synthetic functionalities of the MPEG-4 standard is described. It is based upon an open and modular architecture able to progress with MPEG-4 versions and it is easily adaptable to newly emerging better and higher-level authoring and image sequence analysis features.

Keywords and phrases: MPEG-4, Authoring tools, image sequence analysis.

1. INTRODUCTION

MPEG-4 is the next generation compression standard following MPEG-1 and MPEG-2. Whereas the former two MPEG standards dealt with coding of general audio and video streams, MPEG-4 specifies a standard mechanism for coding of audio-visual objects. MPEG-4 builds on the proven success of three fields [1, 2, 3]: digital television, interactive graphics applications (synthetic content), and interactive multimedia (worldwide web, distribution of and access to content). Apart from natural objects, MPEG-4 also allows the coding of two-dimensional and three-dimensional, synthetic and hybrid, audio and visual objects. Coding of objects enables content-based interactivity and scalability [4]. It also improves coding and reusability of content (Figure 1).

Far from the past "simplicity" of MPEG-2 one-video-plus-two-audio streams, MPEG-4 allows the content creator to compose scenes combining, spatially and temporally, large numbers of objects of many different types: rectangular video, arbitrarily shaped video, still image, speech synthesis, voice, music, text, 2D graphics, 3D, and more. However, the implementation of these capabilities requires a complex authoring process, employing many different competencies from image sequence analysis and encoding of audio/visual/BIFS to the implementation of different delivery scenarios: local access on CD/DVD-ROM, Internet, or broadcast. As multimedia system history teaches, however powerful the technologies underlying multimedia computing, the success of these systems ultimately depends on their ease of authoring.

In [5], the most well-known MPEG-4 Authoring tool (MPEG-Pro) was presented. This includes a graphical user interface, BIFS update, and a timeline, but it can only handle 2D scenes and it is not integrated with any image sequence analysis algorithms. In [6], an MPEG-4 compliant Authoring tool was presented, which, however, is capable only of the composition of 2D scenes. In other articles [7, 8, 9, 10], MPEG-4 related algorithms are presented for the segmentation and generation of video objects which, however, do not provide a complete MPEG-4 authoring suite. Commercial multimedia Authoring tools, such as IBM Hot-Media (http://www-4.ibm.com/software/net.media/) and Veon (http://www.veon.com), are based on their proprietary formats rather than widely acceptable standards. Other commercial solutions based on MPEG-4 like application suites with authoring, server, and client capabilities from iVAST (http://www.ivast.com) and Envivio (http://www.envivio.com) are still under development. In [11, 12], an Authoring tool with 3D functionalities was presented but it did not include any support for image sequence analysis procedures.

Although the MPEG-4 standard and powerful MPEG-4 compliant Authoring tools will provide the needed functionalities in order to compose, manipulate, and transmit the "object-based" information, the production of these objects is out of the scope of the standards and is left to the content developer. Thus, the success of any object-based authoring, coding, and presentation approach depends largely on the segmentation of the scene based on its image contents. Usually, segmentation of image sequences is a two-step process: first scene detection is performed, followed by moving object segmentation and tracking.

Scene detection can be considered as the first stage of a nonsequential (hierarchical) video representation [13]. This is due to the fact that a scene corresponds to a continuous action captured by a single camera. Therefore, application of a scene detection algorithm will partition the video into "meaningful" video segments. Scene detection is useful for coding purposes since different coding approaches can be used according to the shot content. For this reason, scene detection algorithms have attracted a great research interest recently, especially in the framework of the MPEG-4 and MPEG-7 standards; and several algorithms have been reported in the literature dealing with the detection of cut, fading, or dissolve changes either in the compressed or uncompressed domain. A shot is the part of the video that is captured by the camera between a record and a stop operation [14], or by video editing operations. The boundaries between shots are called shot changes, and the action of extracting the shot changes is called shot detection. A shot change can be abrupt or gradual. Examples of gradual changes are mixing, fade-in, and fade-out. During mixing, both shots are shown for a short time (a few seconds). For fade-in and fade-out, the first and the second shots, respectively, are the blank shot.

After shot detection, motion segmentation is a key step in image sequence analysis and its results are extensively used for determining motion features of scene objects as well as for coding purposes to reduce storage requirements [15]. In the past, various approaches have been proposed for motion or spatiotemporal segmentation. A recent survey of these techniques can be found in [16]. In these approaches, a 2D motion or optical flow field is taken as input and a segmentation map is produced, where each region undergoes a movement described by a small number of parameters. There are top-down techniques which rely on the outlier rejection starting from the dominant motion, usually that of the background. Other techniques are bottom-up starting from an initial segmentation and merging regions until the final partition emerges [17, 18]. Direct methods are reported too [19, 20, 21]. All these techniques could be considered automatic since only some tuning parameters are fixed by the user. Grinias and Tziritas [22] proposed a semiautomatic segmentation technique which is suitable for video object extraction for postproduction purposes and object scalable coding such as that introduced in the MPEG-4 standard.

In this paper, an Authoring tool[1] for the MPEG-4 multimedia standard integrated with image sequence analysis algorithms is described. The tool handles the authoring process from the end-user interface specification phase to the cross-platform MP4 file. It fully exploits the object-based coding and 3D synthetic functionalities of the MPEG-4 standard. More specifically, the user can insert basic 3D objects

[1]The Authoring tool is available at http://uranus.ee.auth.gr/pened99/Demos/Authoring_Tool/authoring_tool.html.

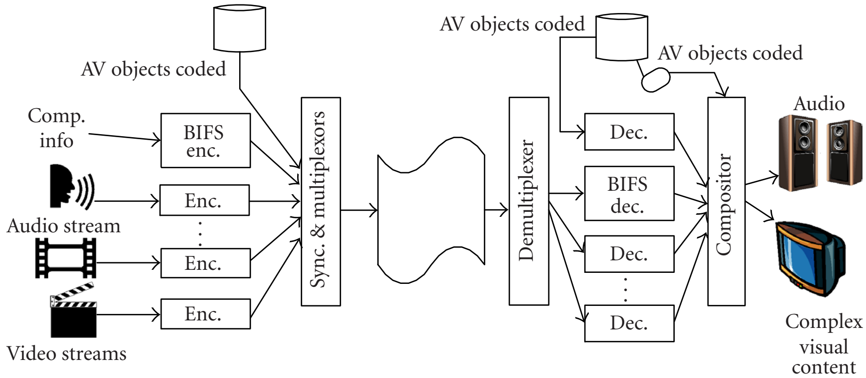


Figure 1: Overview of MPEG-4 systems.

(e.g., boxes, spheres, cones, cylinders) and text and can modify their attributes. Generic 3D models can be created or inserted and modified using the *IndexedFaceSet* node. Furthermore, the behavior of the objects can be controlled by various sensors (time, touch, cylinder, sphere, plane) and interpolators (color, position, orientation). Arbitrarily shaped static images and video can be texture mapped on the 3D objects. These objects are generated by using image sequence analysis integrated with the developed Authoring tool. For the shot detection phase, the algorithm presented in [14] is used. It is based on a method for the extraction of the DC coefficients from MPEG-1 encoded video. After the shots have been detected in an image sequence, they are segmented and the extracted objects are tracked through time using a moving object segmentation and tracking algorithm. The algorithm is based on the motion segmentation technique proposed in [22]. The scheme incorporates an active user who delineates approximately the initial locations in a selected frame and specifies the depth ordering of the objects to be tracked. The segmentation tasks rely on a seeded region growing (SRG) algorithm, initially proposed in [23] and modified to suit our purposes. First, colour-based static segmentation is obtained for a selected frame through the application of a region growing algorithm. Then, the extracted partition map is sequentially tracked from frame to frame using motion compensation and location prediction, as described in [22].

The user can modify the temporal behavior of the scene by adding, deleting, and/or replacing nodes over time using the *Update* commands. Synthetic faces can also be added using the *Face* node and their associated facial animation parameters (FAPs) files. It is shown that our choice of an open and modular architecture of the MPEG-4 authoring system endows it with the ability to easily integrate new modules.

MPEG-4 provides a large and rich set of tools for the coding of audio-visual objects [24]. In order to allow effective implementations of the standard, subsets of the MPEG-4 systems, visual, and audio tool sets that can be used for specific applications have been identified. These subsets, called *Profiles*, limit the tool set a decoder has to implement. For each of these profiles, one or more levels have been set, restricting the computational complexity. Profiles exist for various types of media content (audio, visual, and graphics) and for scene descriptions. The Authoring tool presented here is compliant with the following types of profiles: *The Simple Facial Animation* Visual Profile, *The Scalable Texture* Visual Profile, *The Hybrid* Visual Profile, *The Natural Audio* Profile, *The Complete Graphics* Profile, *The Complete Scene Graph* Profile, and *The Object Descriptor* Profile which includes the object descriptor (OD) tool.

The paper is organized as follows. In Sections 2 and 3, the image sequence analysis algorithms used in the authoring process are presented. In Section 4, MPEG-4 BIFS are presented and the classes of nodes in an MPEG-4 scene are defined. In Section 5, an overview of the Authoring tool architecture and the graphical user interface is given. In Section 6, experiments demonstrate 3D scenes composed by the Authoring tool. Finally, conclusions are drawn in Section 7.

2. SHOT DETECTION

The shot detection algorithm used in the authoring process is an adaptation of the method presented originally by Yeo and Liu [14]. The basic tenet is that the DC coefficients of the blocks from an MPEG-1 encoded video contain enough information for the purpose of shot detection. In addition, as shown in [14], the use of this spatially reduced image (DC image), due to its smoothing effect, can reduce the effects of motion and increase the overall efficiency of the method. Computing the DC coefficient for P- and B-frames would be computationally complex because it requires motion compensation. The DC coefficient is therefore approximated as a weighted average of the DC coefficients of the four neighboring blocks of the previous frame according to the motion vector. The weights of the averaging operation are proportional to the surface of the overlap between the current block and the respective block of the previous frame. By using this approximation and comparing each two subsequent images, using an appropriate metric as described in the sequel, a sequence of differences between subsequent frames is produced. Abrupt scene changes manifest themselves as sharp peaks at the sequence of differences. The algorithm must detect these peaks among the signal noise.

In the proposed procedure, the video is not available in MPEG format, therefore the aforementioned method is

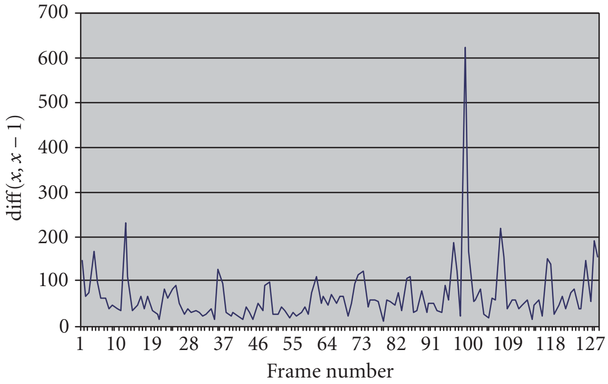


FIGURE 2: Absolute difference of consecutive DC images.

applied to YUV raw video after a lowpass filtering, which effectively reduces each frame to a DC image.

Two metrics were proposed for comparing frames, that of the absolute difference and that of the difference of the respective histograms. The first method, which was chosen by the authors of this paper for its computational efficiency, directly uses the absolute difference of the DC images [25]:

$$\operatorname{diff}(X, Y) = \frac{1}{M \times N} \sum_{i,j} |x_{i,j} - y_{i,j}|, \quad (1)$$

where M and N are the dimensions of the frame and $x_{i,j}$ and $y_{i,j}$ represent two subsequent frames. As Yeo and Liu [14] note, this is not efficient in the case of full frames because of the sensitivity of this metric to motion, but the smoothing effect of the DC coefficient estimation can compensate that to a large extent. The second metric compares the histograms of the DC images. This method is insensitive to motion [14], and, most often, the number of bins b used to form the histograms is in the range 4–6.

Once the difference sequence is computed (Figure 2), a set of two rules is applied to detect the peaks. First, the peak must have the maximum value in an interval with a width of m frames, centered at the peak. Secondly, the peak must be n times greater than the second largest value of the second interval. This rule enforces the sharpness of the peak.

When just the two aforementioned rules were used, the system seemed to erroneously detect low-valued peaks which originated from errors related to P- and B-frames. These short peaks can be seen in Figure 2. Therefore, we introduced a third rule, that of an absolute threshold, which excludes these short peaks. The threshold equals $d \times M \times N$, where M and N are the dimensions of the frame and d is a real parameter. In the case of histograms, the threshold is also proportional to 2^b.

In our experiments, good results, in terms of shot recall and precision, were obtained with $m = 3$–5, $n = 1.5$–2.0, and $d \approx 0.0015$. A more thorough discussion on the topic of the choice of parameters can be found in [25].

Another issue is the relative importance of chrominance in peak detection. In particular, the formula $d = (1 - c)d_L + cd_C$ was applied. Using the value $c = 0.4$–0.7 gives good results, but acceptable results (about 30% inferior) are obtained with other values of this parameter as well.

3. MOVING-OBJECT SEGMENTATION AND TRACKING

3.1. *Overall structure of video segmentation algorithms*

After shot detection, a common requirement in image sequence analysis is the extraction of a small number of moving objects from the background. The presence of a human operator, called here the *user* of the Authoring tool, can greatly facilitate the segmentation work for obtaining a semantically interpretable result. The proposed algorithm incorporates an active user for segmenting the first frame and for subsequently dealing with occlusions during the moving object tracking.

For each object, including the background, the user draws a closed contour entirely contained within the corresponding object. Then, a region growing algorithm expands the initial objects to their actual boundaries. Unlike [22], where the segmentation of the first frame is mainly based on the motion information, the region growing is based on the color of the objects and is done in a way that overcomes their color inhomogeneity. Having obtained the segmentation of the first frame, the tracking of any moving object is done automatically, as described in [22]. Only the layered representation of the scene is needed by the user in order to correctly handle overlaps. We assume that each moving region undergoes a simple translational planar motion represented by a two-dimensional velocity vector, and we reestimate an update for this vector from frame to frame using a region matching (RM) technique, which is an extension of block matching to regions of any shape and provides the required computational robustness. This motion estimation is performed after shrinking the objects in order to ensure that object contours lie within the objects. The "shrunken" objects are projected onto their predicted position in the next frame using motion compensation and the region growing algorithm is applied from that position.

In Section 3.2, the SRG algorithm is presented. In Section 3.3, the initial segmentation is described, as well as the modifications applied to SRG, in order to cope with the color inhomogeneity of objects. Section 3.4 presents, in summary, how the SRG algorithm is used for the temporal tracking of the initial segmentation.

3.2. *The SRG algorithm*

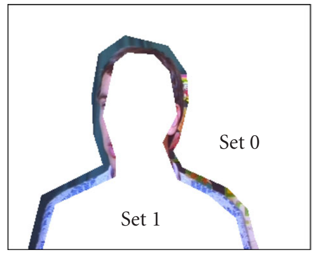
Segmentation is carried out by an SRG algorithm which was initially proposed for static image segmentation using a homogeneity measure on the intensity function [23]. It is a sequential labelling technique in which each step of the algorithm labels exactly one pixel, that with the lowest dissimilarity. Letting n be the number of objects (classes), an initial set of connected components $A_1^0, A_2^0, \ldots, A_n^0$ is required. At each step m of the algorithm, let B^{m-1} be the set of all yet unlabelled points which have at least one immediate neighbor already labelled, that is, belonging to one of the partially completed connected components $\{A_1^{m-1}, A_2^{m-1}, \ldots, A_n^{m-1}\}$. In this paper, 8-connection neighborhoods are considered. For each pixel $p \in B^{m-1}$, we denote by $i(p) \in \{1, 2, \ldots, n\}$ the index of the set A_i^{m-1} that p adjoins and by $\delta(p, A_{i(p)}^{m-1})$ the dissimilarity measure between p and $A_{i(p)}^{m-1}$, which depends on the segmentation features used. If the characterization of the sets is not updated during the sequential labelling process, the dissimilarity will be $\delta(p, A_{i(p)}^{0})$. If p adjoins two or more of the sets A_i^{m-1}, we define $i(p)$ to be the index of the set that minimizes the criterion $\delta(p, A_j^{m-1})$ over all neighboring sets A_j^{m-1}. In addition, we can distinguish a set F of boundary pixels and add p to F when p borders more than one set. In our implementation, boundary pixels p are flagged as belonging to F and, at the same time, they are associated with the set that minimizes the dissimilarity criterion over all sets on whose boundary they lie. The set of boundary points F is useful for boundary operations, as we will see in Section 3.4. Then we choose among the points in B^{m-1} one satisfying the relation

$$z = \arg\min_{p \in B^{m-1}} \left\{ \delta\left(p, A_{i(p)}^{m-1}\right) \right\} \tag{2}$$

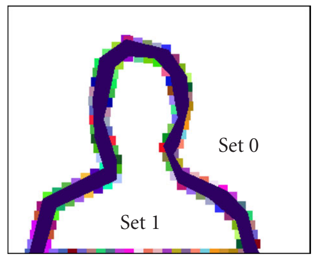
and append z to $A_{i(z)}^{m-1}$, resulting in $A_{i(z)}^{m}$. This completes one step of the algorithm and finally, when the border set becomes empty after a number of steps equal to the number of initially unlabelled pixels, a segmentation map $(R_1, R_2, \ldots, R_n)$ is obtained with $A_i^m \subseteq R_i$ (for all i, m) and $R_i \cap R_j = \varnothing$ $(i \neq j)$, where $\cup_{i=1}^{n} R_i = \Omega$ is the whole image.

For the implementation of the SRG algorithm, a list that keeps its members (pixels) ordered according to the criterion value $\delta(\cdot, \cdot)$ is used, traditionally referred to as sequentially sorted list (SSL).

(a)



(b)



(c)

Figure 3: User provided input of initial sets (b) and automatically extracted representative points (c) for *Erik*'s frame 0 (a).

3.3. Object initialization and static segmentation

The initial regions required by the region growing algorithm must be provided by the user. A tool has been built for drawing a rectangle or a polygon inside any object. Then points which are included within these boundaries define the initial sets of object points. This concept is illustrated in Figure 3b, where the input of initial sets for the frame 0 of the sequence *Erik* is shown. The user provides an approximate pattern for each object in the image that is to be extracted and tracked.

The color segmentation of the first frame is carried out by a variation of SRG. Since the initial sets may be characterized by color inhomogeneity, on the boundary of all sets we place representative points for which we compute the locally average color vector in the lab system. In Figure 3c, the small square areas correspond to the regions of points that participate to the computation of the average color vector for each such representative point. The dissimilarity of the candidate for labelling and region growing point z of (2) from the labelled regions that adjoins is determined using this feature and the Euclidean distance, which may be possibly combined with the meter of the color gradient of z. After the labelling of z, the corresponding feature is updated. Therefore, we search for sequential spatial segmentation based on color homogeneity, knowing that the objects may be globally inhomogeneous, but presenting local color similarities sufficient for their discrimination.

When the static color segmentation is completed, every pixel p is assigned a label $i(p) \in \{1, 2, \ldots, n\}$ while boundary information is maintained in set F. Thus, the set map i is the first segmentation map i_0, which is going to be tracked using the method that has been presented in [22] in detail and is described shortly in Section 3.4.

3.4. Tracking

We now briefly describe how the result of the initial segmentation (set map i_0) is tracked over a number of consecutive frames. We assume that the result has been tracked up to frame $k-1$ (set map i_{k-1}) and we now wish to obtain the set map i_k corresponding to frame k (partition of frame k). The initial sets for the segmentation of frame k are provided by the set map i_{k-1}. The description of the tracking algorithm follows, while the motivations of the algorithm have already been presented in Section 3.1.

For the purpose of tracking, a layered representation of the sets, rather than the planar one implied by SRG, is introduced in order to be able to cope with real world sequences which contain multiple motions, occlusions, or a moving background. Thus, we assume that sets are ordered according to their distance from the camera as follows:

$$\forall i, j \in \{1, 2, \ldots, n\}, \quad R_i \text{ moves behind } R_j \text{ iff } i < j. \tag{3}$$

In this way, set R_1 refers to the background, set R_2 moves in front of set R_1 and behind the other sets, and so forth. The user is asked to provide this set ordering in the stage of objects initialization.

Having this set ordering available, for each set $R \in \{R_2, R_3, \ldots, R_n\}$ of set map i_{k-1}, the following operations are applied in order of proximity, beginning with the most distant.

(i) The border of R is dilated for obtaining the set of seeds A of R, which are required as input by SRG.

(ii) The velocity vector of R is reestimated assuming that it remains almost constant over time. The estimation is done using RM (with subpixel accuracy) on the points of A.

(iii) The "shrunken" subset A of region R is translated from image $k-1$ to image k according to the estimated displacement.

The last step, before applying the motion-based SRG, is the estimation of the background velocity vector. Then, SRG is applied to points that remain unlabelled after the above operations, as described in [22].

Furthermore, two boundary regularization operations are proposed in [22] to stabilize object boundaries over time. The first one smooths the boundary of the objects, while the second computes an average shape using the information of a number of previously extracted segmentation maps.

3.5. System description

The proposed algorithm was designed for semiautomatic segmentation requiring an initial user input (the user must draw a rough boundary of the desired object), therefore it is suited for an Authoring tool where user interaction is expected. The spatiotemporal algorithm is a separate module developed in Java, integrated with the Authoring tool, which was developed in C++ for Windows (Borland Builder C++5) and OpenGL interfaced with the "core" module and the tools of the IM1 (MPEG-4 implementation group) software platform. The IM1 3D player is a software implementation of an MPEG-4 systems player [26]. The player is built on top of the core framework, which also includes tools to encode and multiplex test scenes. It aims to be compliant with the complete 3D profile [1]. This shows the flexibility of the architecture of the presented Authoring tool to efficiently combine different modules and integrate the results in the same MPEG-4 compatible scene. As can be seen for the experimental results, the SRG algorithm was shown to be very efficient. In case the tracking fails, the user can select a more appropriate boundary for the desired object, else the tracking process may be restarted from the frame where the tracking failed.

4. BIFS SCENE DESCRIPTION FEATURES

The image sequence analysis algorithms described above are going to be integrated with an MPEG-4 Authoring tool providing a mapping of BIFS nodes and syntax to user-friendly windows and controls. The BIFS description language [27] has been designed as an extension of the VRML 2.0 [28] file format for describing interactive 3D objects and worlds. VRML is designed to be used on the Internet, intranets, and local client systems. VRML is also intended to be a universal interchange format for integrated 3D graphics and multimedia. The BIFS version 2 is a superset of VRML and can be used as an effective tool for compressing VRML scenes. BIFS is a compact binary format representing a predefined set of scene objects and behaviors along with their spatiotemporal relationships. In particular, BIFS contains the following four types of information:

(i) the attributes of media objects which define their audio-visual properties;

(ii) the structure of the scene graph which contains these objects;

(iii) the predefined spatiotemporal changes of these objects, independent of user input;

(iv) the spatiotemporal changes triggered by user interaction.

The scene description follows a hierarchical structure that can be represented as a tree (Figures 4 and 5). Each node of the tree is an audio-visual object. Complex objects are constructed by using appropriate scene description nodes. The tree structure is not necessarily static. The relationships can evolve in time and nodes may be deleted, added, or modified. Individual scene description nodes expose a set of parameters through which several aspects of their behavior can be controlled. Examples include the pitch of a sound, the color of a synthetic visual object, or the speed at which a video sequence is to be played. There is a clear distinction between the audio-visual object itself, the attributes that enable the control of its position and behavior, and any elementary streams that contain coded information representing attributes of the object.

The proposed MPEG-4 Authoring tool implements the BIFS nodes graph structure allowing authors to take full advantage of MPEG-4 nodes functionalities in a friendly graphical user interface.

4.1. Scene structure

Every MPEG-4 scene is constructed as a direct acyclic graph of nodes. The following types of nodes may be defined.

(i) *Grouping nodes* construct the scene structure.

(ii) *Children nodes* are offsprings of grouping nodes representing the multimedia objects in the scene.

(iii) *Bindable children nodes* are the specific type of children nodes for which only one instance of the node

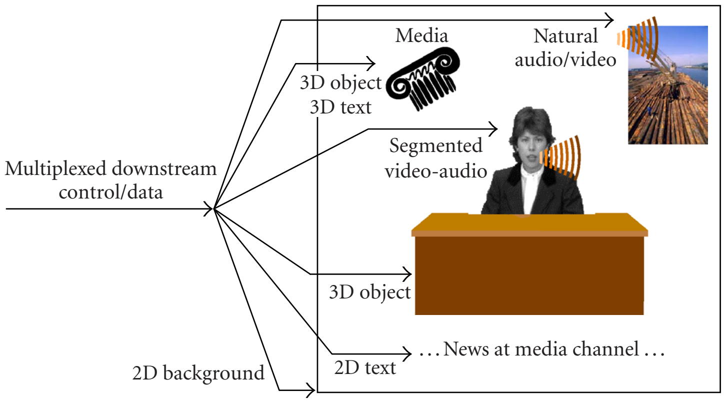


Figure 4: Example of an MPEG-4 scene.

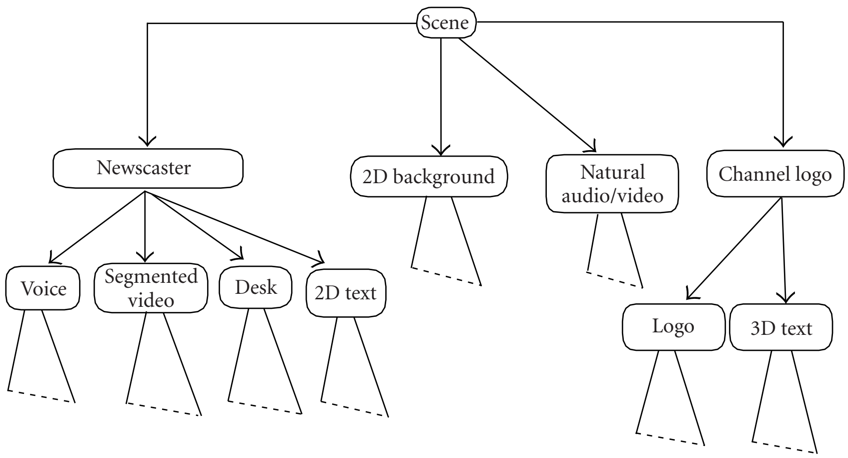


Figure 5: Corresponding scene tree.

type can be active at a time in the scene (a typical example of this is the viewpoint for a 3D scene; a 3D scene may contain multiple viewpoints or "cameras," but only one can be active at a time).

(iv) *Interpolator nodes* constitute another subtype of children nodes which represent interpolation data to perform key frame animation. These nodes generate a sequence of values as a function of time or other input parameters.

(v) *Sensor nodes* sense the user and environment changes for authoring interactive scenes.

4.2. Nodes and fields

BIFS and VRML scenes are both composed of collections of nodes arranged in hierarchical trees. Each node represents, groups, or transforms an object in the scene and consists of a list of fields that define the particular behavior of the node. For example, a Sphere node has a radius field that specifies the size of the sphere. MPEG-4 has roughly 100 nodes with 20 basic field types representing the basic field data types: boolean, integer, floating point, two- and three-dimensional vectors, time, normal vectors, rotations, colors, URLs, strings, images, and other more arcane data types such as scripts.

4.3. ROUTEs and dynamical behavior

The event model of BIFS uses the VRML concept of ROUTEs to propagate events between scene elements. ROUTEs are connections that assign the value of one field to another field. As is the case with nodes, ROUTEs can be assigned a "name" in order to be able to identify specific ROUTEs for modification or deletion. ROUTEs combined with interpolators can cause animation in a scene. For example, the value of an interpolator is ROUTEd to the rotation field in a transform node, causing the nodes in the transform node's children field to be rotated as the values in the corresponding field in the interpolator node change with time. This event model has been implemented in a graphical way, allowing users to add interactivity and animation to the scene (Figure 6).

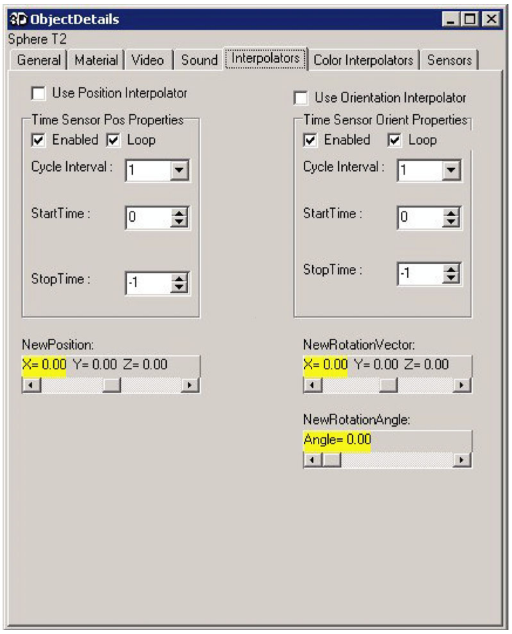


Figure 6: The interpolators panel.

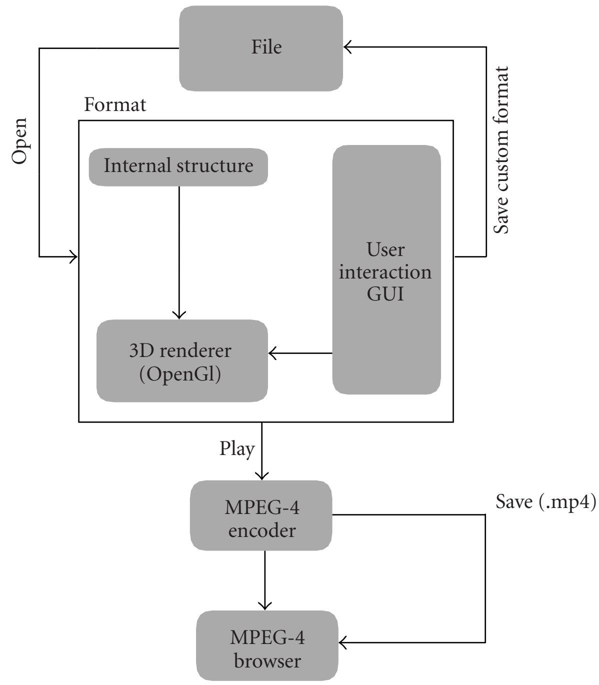


Figure 7: System architecture.

4.4. *Streaming scene description updates: BIFS command*

The mechanism with which BIFS information is provided to the receiver over time comprises the BIFS-Command protocol (also known as BIFS Update) and the elementary stream that carries it, thus called BIFS-command stream. The BIFS-Command protocol conveys commands for the replacement of a scene, addition or deletion of nodes, modification of fields, and so forth. For example, a "ReplaceScene" command becomes the entry (or random access) point for a BIFS stream, exactly in the same way as an Intraframe serves as a random access point for video. A BIFS-Command stream can be read from the web as any other scene, potentially containing only one "ReplaceScene" command, but it can also be broadcast as a "push" stream, or even exchanged in a communication or collaborative application. BIFS commands come in four main functionalities: scene replacement, node/field/route insertion, node/value/route deletion, and node/field/value/route replacement. The BIFS-Command protocol has been implemented so as to allow the user to temporarily modify the scene using the Authoring tool graphical user interface.

4.5. *Facial animation*

The facial and body animation nodes can be used to render an animated face. The shape, texture, and expressions of the face are controlled by the facial definition parameters (FDPs) and/or the FAPs. Upon construction, the face object contains a generic face with a neutral expression. This face can be rendered. It can also immediately receive the animation parameters from the bitstream, which will produce animation of the face: expressions, speech, and so forth. Meanwhile, definition parameters can be sent to change the appearance of the face from something generic to a particular face with its own shape and (optionally) texture. If so desired, a complete face model can be downloaded via the FDP set. The described application implements the Face node using the generic MPEG-4 3D face model, allowing the user to insert a synthetic 3D animated face.

5. MPEG-4 AUTHORING TOOL

5.1. *System architecture*

The process of creating MPEG-4 content can be characterized as a development cycle with four stages: Open, Format, Play, and Save (Figure 7). In this somewhat simplified model, the content creators can do the following.

(i) They can edit/format their own scenes inserting synthetic 3D objects, such as spheres, cones, cylinders, text, boxes, and background (Figure 8). They may also group objects, modify the attributes (3D position, color, texture, etc.) of the edited objects, or delete objects from the created content. The user can perform the image sequence analysis procedures described in Sections 2 and 3 in order to create arbitrarily shaped video objects and insert them into the scene. He can also insert sound and natural video streams, add interactivity to the scene using sensors and interpolators, and dynamically control the scene using an implementation of the BIFS-Command protocol. Generic 3D models can be created or inserted and modified using the IndexedFaceSet node. The user can insert a

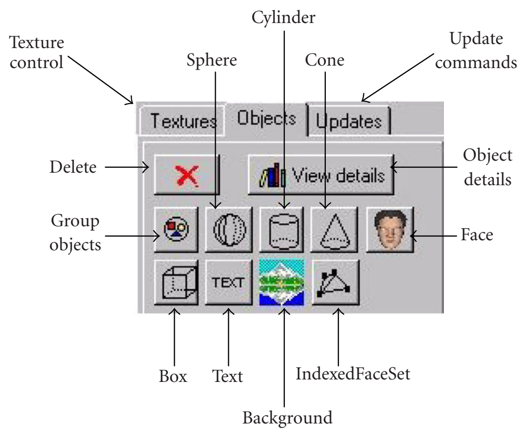


FIGURE 8: Authoring tool application toolbar.

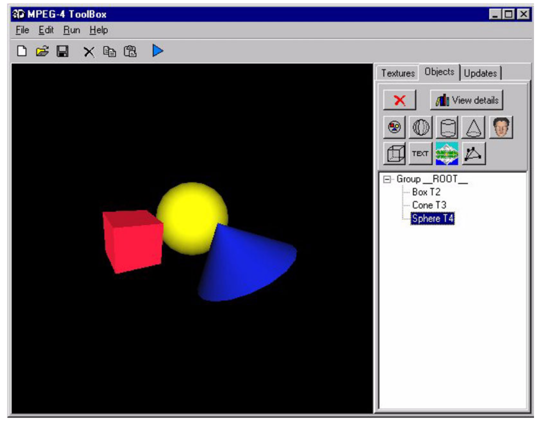


FIGURE 9: Main window indicating the different components of the user interface.

synthetic animated face using the implemented *Face* node. During these procedures, the attributes of the objects and the commands as defined in the MPEG-4 standard, and, more specifically, in BIFS, are stored in an internal program structure, which is continuously updated depending on the actions of the user. At the same time, the creator can see in real time a 3D preview of the scene on an integrated window using OpenGL tools (Figure 9).

(ii) They can present the created content by interpreting the commands issued by the edition phase and allowing the possibility of checking whether the current description is correct.

(iii) They can open an existing file.

(iv) They can save the file either in custom format or after encoding/multiplexing and packaging in an MP4 file [24], which is expected to be the standard MPEG-4 file format. The MP4 file format is designed to contain the media information of an MPEG-4 presentation in a flexible, extensible format which facilitates interchange, management, editing, and presentation of the media.

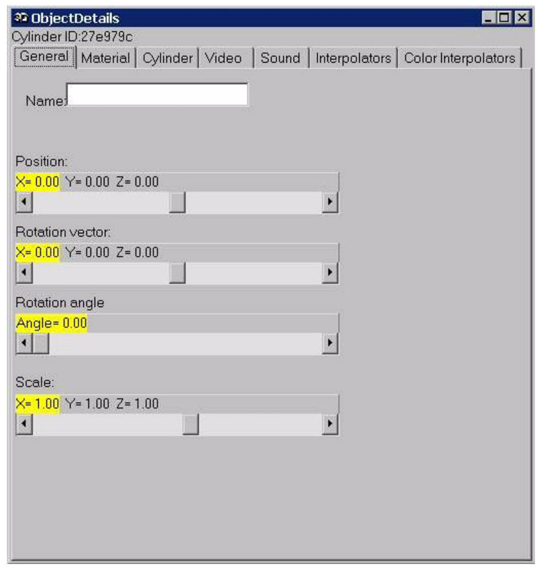


FIGURE 10: Object Details Window indicating the properties of the objects.

5.2. *User interface*

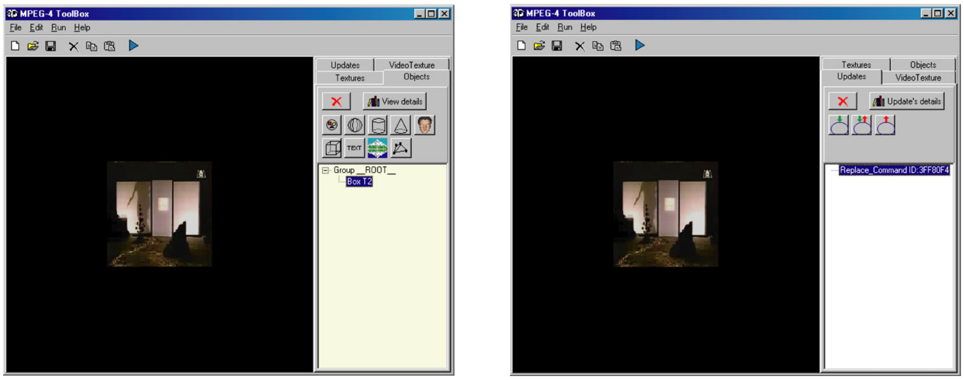
To improve the authoring process, powerful graphical tools must be provided to the author [29]. The temporal dependence and variability of multimedia applications hinder the author from obtaining a real perception of what he is editing. The creation of an environment with multiple synchronized views and the use of OpenGL were implemented to overcome this difficulty. The interface is composed of three main views, as shown in Figure 9.

Edit/Preview

By integrating the presentation and editing phases in the same view, the author is enabled to see a partial result of the created object on an OpenGL window. If any given object is inserted in the scene, it can be immediately seen on the presentation window (OpenGL window) located exactly in the given 3D position. The integration of the two views is very useful for the initial scene composition.

Scene Tree

This attribute provides a structural view of the scene as a tree (a BIFS scene is a graph, but for ease of presentation, the graph is reduced to a tree for display). Since the edit view cannot be used to display the behavior of the objects, the tree scene is used to provide more detailed information concerning them. The drag-n-drop and copy-paste modes can also be used in this view.



(a) (b)

FIGURE 11: Using Update commands in the Authoring tool.

Object Details

This window, shown in Figure 10, offers object properties that the author can use to assign values other than those given by default to the synthetic 3D objects. The user can perform the image sequence analysis procedures described in Sections 2 and 3 in order to create arbitrarily shaped video objects and insert them into the scene. This arbitrarily shaped video can be used as texture on every object. Other supported properties are 3D position, 3D rotation, 3D scale, color (diffuse, specular, emission), shine, texture, video stream, audio stream (the audio and video streams are transmitted as two separated elementary streams according to the OD mechanism), cylinder and cone radius and height, textstyle (plain, bold, italic, bolditalic) and fonts (serif, sans, typewriter), sky and ground background, texture for background, interpolators (color, position, orientation), and sensors (sphere, cylinder, plane, touch, time) for adding interactivity and animation to the scene. Furthermore, the author can insert, create, and manipulate generic 3D models using the *IndexedFaceSet* node. Simple VRML files can also be inserted in a straightforward manner. Synthetically animated 3D faces can be inserted by the *Face* node. The author must provide an FAP file [30] and the corresponding encoder parameter file (EPF), which is designed to give the FAP encoder all information related to the corresponding FAP file, like I and P frames, masks, frame rate, quantization scaling factor, and so on. Then, a bifa file (binary format for animation) is automatically created so as to be used in the scene description and OD files.

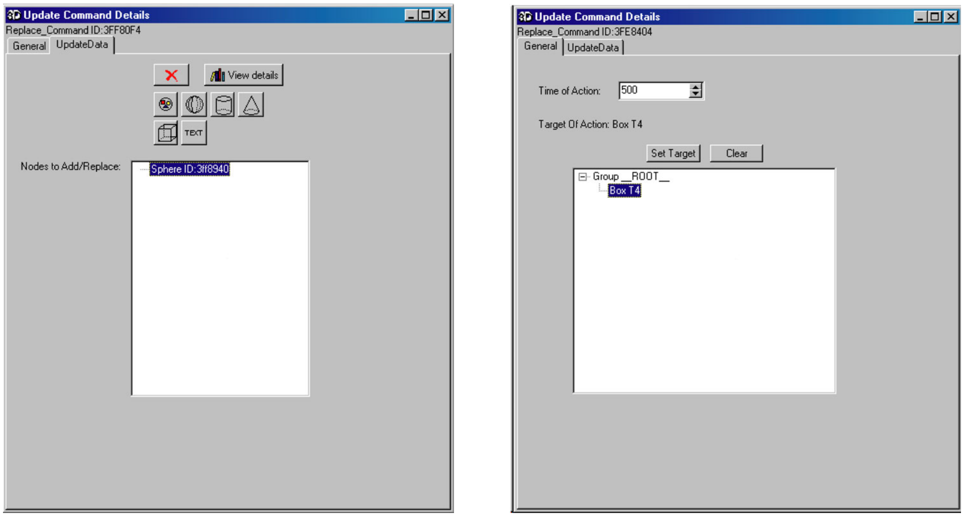
6. EXPERIMENTAL RESULTS

In this section, two examples are presented, describing the steps that lead to the creation of two MPEG-4 scenes.

The first example demonstrates the use of the BIFS commands (Update), which is used to give to the user a real perception about what he/she is editing in a temporal editing environment. In this scene, a textured box is first created and after a period of time is replaced by a textured sphere. The exact steps are the following: on the main window, a box with a video texture is created (Figure 11a). On the Updates tab (Figure 11b), the Replace command is selected ("Replace" button). On the Update Command Details panel (Figure 12a), in tab "UpdateData," a sphere with another video texture is selected. On the same panel, in tab "General," (Figure 12b), the box is specified ("Set Target" button) and also the time of action needed ("Time of Action" button) (e.g., 500 ms). Finally, by pressing the button "Play," the result is shown by the 3D MPEG-4 Player (Figures 13a and 13b).

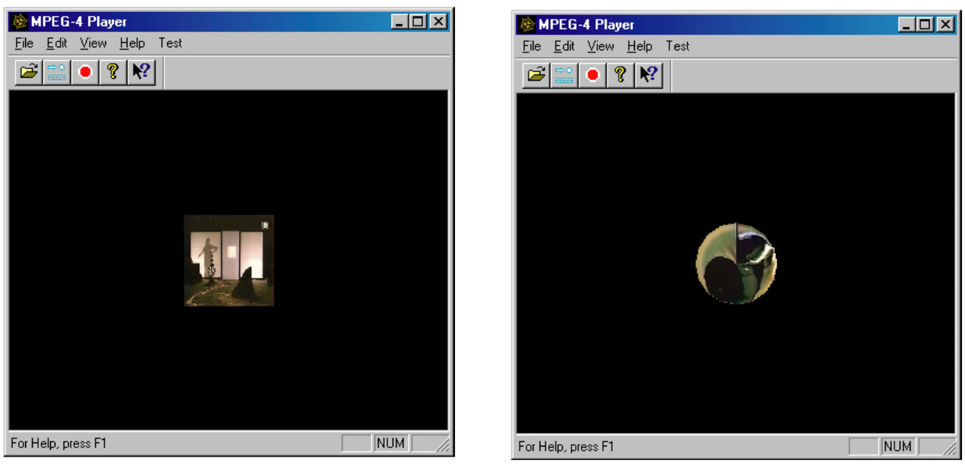
The second example leads to the creation of an MPEG-4 scene containing arbitrarily shaped video objects using the shot detection and object segmentation procedures. The scene represents a virtual studio. The scene contains several groups of synthetic objects including boxes with textures and text objects (Figure 20). The "logo" group which is located on the upper left corner of the studio is composed of a rotating box and a text object that describes the name of the channel. The background contains four boxes (left-right side, floor and back side) with image textures. The desk is created using two boxes. On the upper right corner of the scene, a box with natural video texture is presented. On this video box, relative videos are loaded according to the news. The newscaster (image sequence "Akiyo") is an arbitrarily shaped video object produced using the algorithms described in Sections 2 and 3. The virtual studio scence in the IM1 3D player can be seen in Figure 21.

In order to test the shot detection algorithm, a test sequence was created composed of the two image sequences "Akiyo" and "Eric." Using the user interface of the Authoring tool (Figure 14), the user can select a video for processing. The supporting formats are YUV color and gray scale at 176 × 144 pixels (QCIF) and 352 × 288 pixels (CIF). As soon as the user selects the video, the algorithm presented in Section 2 is applied. The result is the temporal segmentation of the image sequence into shots. After the shot detection procedure, the semiautomatic moving object segmentation procedure begins (Section 3). The user draws a rough



(a) (b)

Figure 12: Specifying the appropriate properties.



(a) (b)

Figure 13: The result of the Update commands as shown in the Authoring tool.

boundary around the moving foreground object (Figure 15) of each shot and the algorithm automatically performs the region growing and tracking procedures (Figure 16). The result is a set of segmentation masks for each shot of the image sequence (Figure 14). The user can easily select the objects from each shot that are to be included in the scene. Every selected mask (for every shot) is given as input to the MPEG-4 video reference software [31] which is used for encoding and decoding video sequences. After a transcoding procedure, the final result is an H.263 video which is compliant with the current MPEG-4 IM1 player implementation. This video can be used as texture on every object as shown in Figures 17 and 18. The block diagram of the complete procedure is presented in Figure 19.

The same scene can easily be modified so as to contain a synthetic newscaster (Figures 22 and 23). The body of the newscaster is an IndexedFaceSet imported from a VRML 3D model. The 3D face was inserted by using the corresponding button. After the selection of an FAP file and an audio stream (a saxophone appears on the upper left corner), the face animation is configured according to the selected FAP file. The video stream (H.263) and the audio stream (G.723) are transmitted as two separate elementary streams according to the OD mechanism. All animation (except the face animation) is implemented using interpolator nodes. Some major parts of the produced scene description file (.txt) are shown in Figure 24.

The AnimationStream node reads from an external source the selected FAP file. The Transform node inserted before the Face node controls the position of the animated

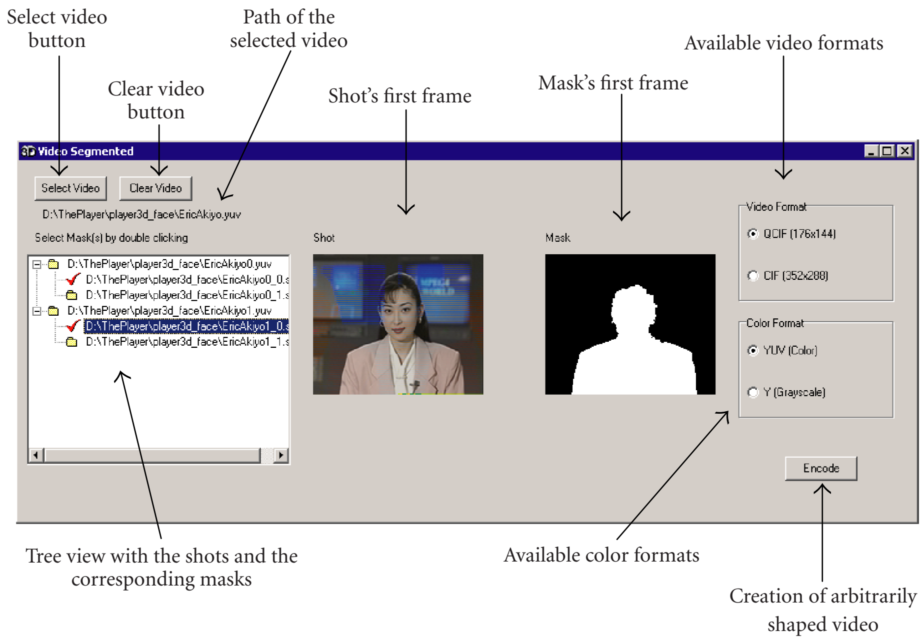


Figure 14: The segmentation form in the Authoring tool.

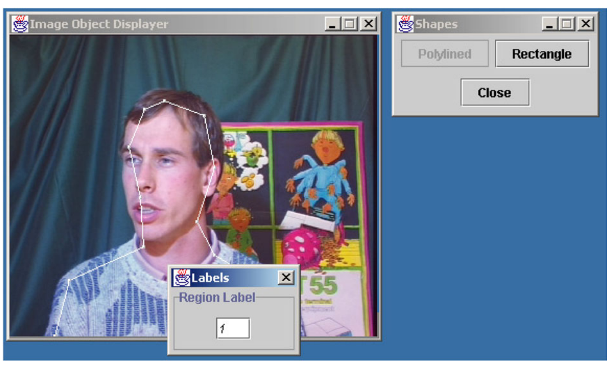


Figure 15: Rough boundary around the foreground object.

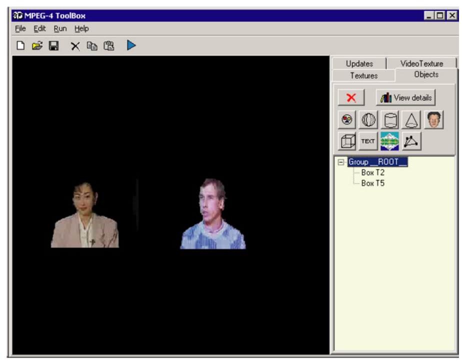


Figure 17: The result of the arbitrarily shaped video objects creation textured on two boxes as shown in the Authoring tool.

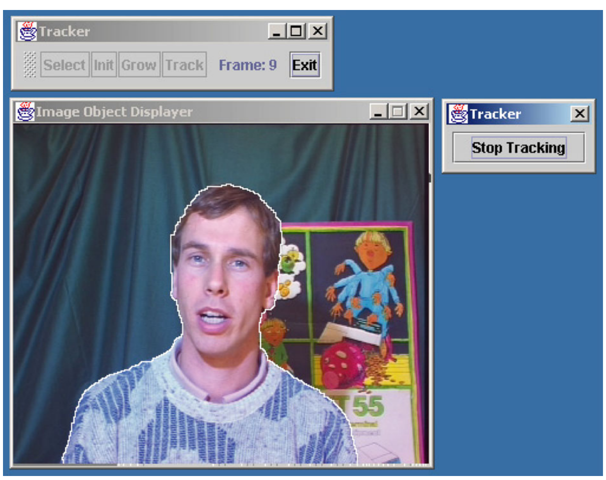


Figure 16: Snapshot of the tracking procedure.

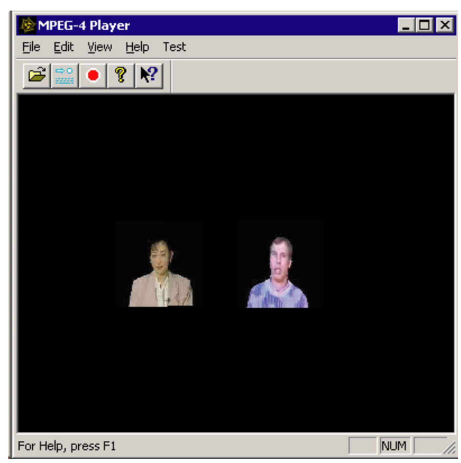


Figure 18: The result of the shot detection and segmentation procedures as shown in the player.

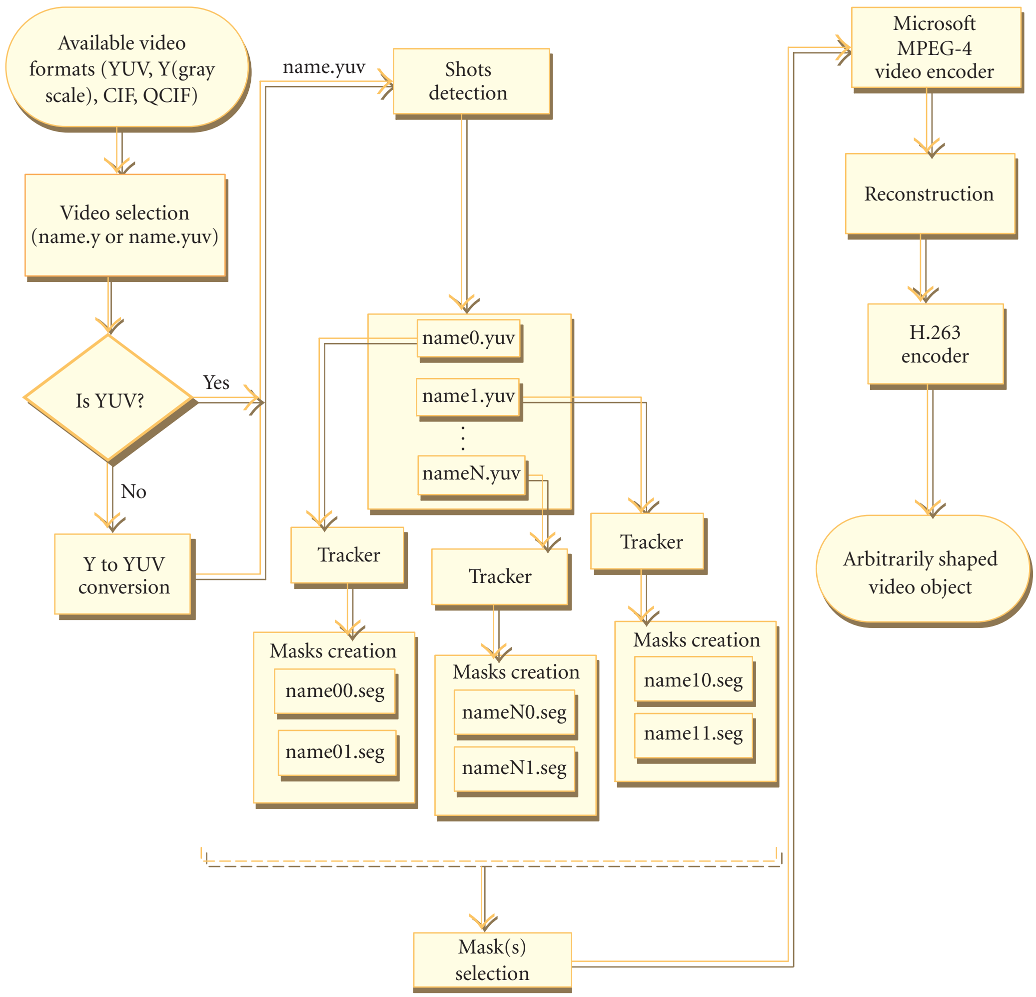


FIGURE 19: Segmentation procedure for the creation of H.263 video format.

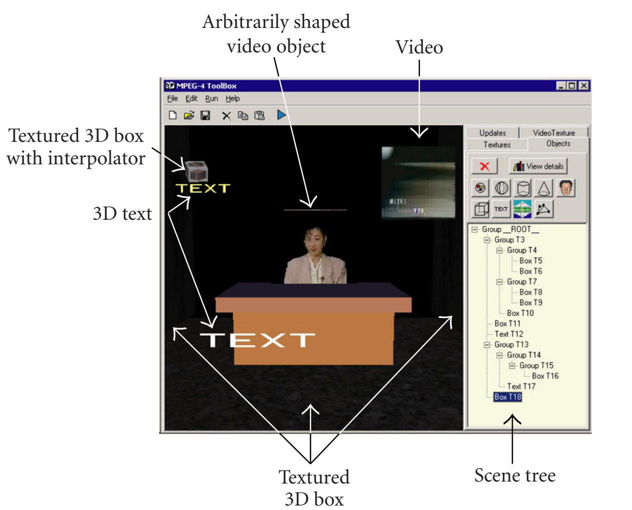


FIGURE 20: The virtual studio scene using arbitrarily shaped video objects in the Authoring tool.

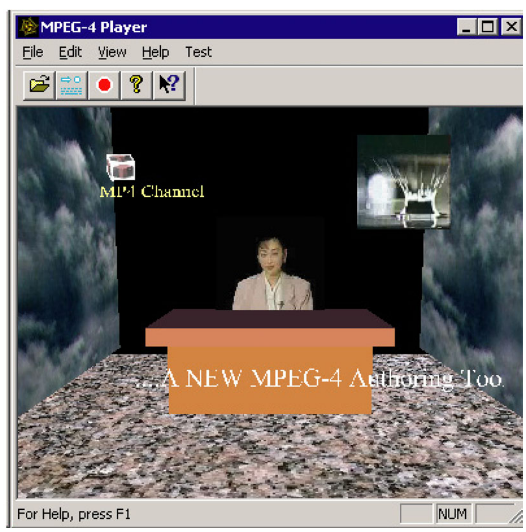


FIGURE 21: The virtual studio scene in the IM1 3D player.

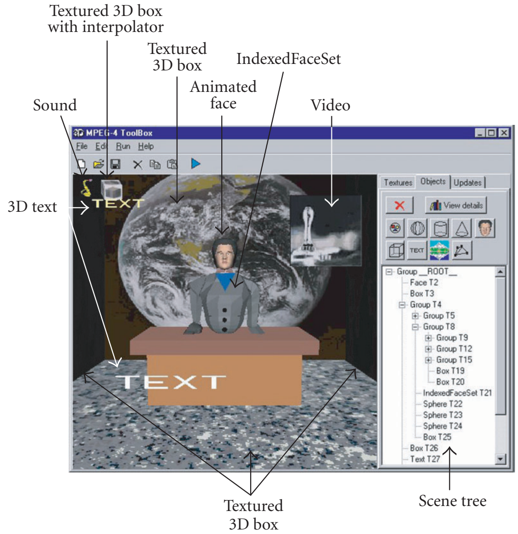


Figure 22: The virtual studio scene in the Authoring tool.

Figure 23: The virtual studio scene in the IM1 3D player.

face in the scene. The Face node inserts the animated face and connects it with the FAP file defined earlier. The following group creates the "logo" which is located on the upper left corner and more specifically, the textured rotating box. First the position of the box (Transform node) and then the image to be applied as texture (appearance and texture fields) are defined. Finally, the geometry and the dimensions of the object are defined (geometry node). In our case, the object is a box. The final part contains the necessary nodes for creating the rotating motion. First, the period of the motion is defined (how fast the box will be rotated) and whether the rotation speed will be constant. This is controlled by the TimeSensor node and the loop and cycleInterval fields. The OrientationInterpolator node defines the intermediate positions of the motion. Finally, the ROUTE nodes connect the defined parameters of the movement to the textured object. The objects are uniquely characterized by the DEF nodes. For example, the texture box is object T120661744.

As can be seen from the above, the text-based description format for MPEG-4 is very complicated. It is almost impossible to develop an MPEG-4 scene from scratch using text only. The user should be aware of a complicated syntax and a great number of MPEG-4 BIFS node names and at the same time keep track of all the defined object names. The presented Authoring tool allows nonexpert MPEG-4 users to create complicated scenes by converting this text-based description to a more native, graphical description.

7. CONCLUSIONS

In this paper, an Authoring tool for the MPEG-4 multimedia standard integrated with image sequence analysis algorithms was presented. The tool maps BIFS features and functionalities to common Window controls allowing users to efficiently create or edit and finally play MPEG-4 compliant scenes using an external MPEG-4 player. The Authoring tool is integrated with a shot detection algorithm along with a semiautomatic method for moving object segmentation and tracking. The user can perform these image sequence analysis procedures in order to create arbitrarily shaped video objects and insert them into the scene. Experimental results demonstrated that it is possible to create complex scenes using unique MPEG-4 features such as object-based coding, updates, and facial animation. The image sequence analysis algorithms were integrated as separate modules. This shows the flexibility of the architecture of the presented Authoring tool to efficiently combine different modules and integrate the results in the same MPEG-4 compatible scene.

The presented parts of the corresponding Text Description files show that it is almost impossible for the nonexpert to build even simple MPEG-4 scenes from scratch using only text. We found that while content developers were satisfied with the efficiency and the effectiveness of the system, those that were not familiar with the MPEG-4 standard had problems in understanding the terminology used. Thus, further development and refinement is needed before the tool can be useful for large-scale deployment.

Another important feature of the Authoring tool is that it produces scenes which are totally MPEG-4 compliant. These scenes can be visualized using the IM1-3D player developed by the MPEG-4 group without any modifications. Thus, the tool may be used to create MPEG-4 compliant applications without introducing proprietary features.

The presented paper also highlights and exemplifies the manner in which nonexpert MPEG-4 users may create and manipulate MPEG-4 content. Specifically, the tool developed is intended to help MPEG-4 algorithm and system developers integrate their algorithms and make them available through a user-friendly interface. It may also help as a beginning for the development of new tools. Finally, the tool may serve as a benchmark for the comparison of other, proprietary or not, Authoring tools to one with the capabilities of the MPEG-4 system.

```
    DEF ID_014   AnimationStream     #fap animation stream
    {
            url 50
    }
Transform {
        translation  0.000 1.529 1.690
        rotation  0.000 0.000 0.000 0.000
        scale  0.013 0.013 0.013
    Children Face                   #Face node
    {
        fap DEF ID_104 FAP{}
        renderedFace []
    }
}
. . .
DEF T120661744  Transform {
    translation  0.000 0.000 0.000
    rotation  1.786 1.014 0.000 0.911
    children Shape {
        appearance Appearance {
        texture ImageTexture {
            url 10
    }
        textureTransform TextureTransform {
    }
}
        geometry Box {        #box with image texture
            size 0.796 0.796 0.694
        }
    }
}

DEF OrientTS120658180      TimeSensor {
    stopTime -1
    startTime 0
    loop TRUE                   # time sensor for interpolation
                                #    purposes
    cycleInterval 15
}
DEF ORI120658180    OrientationInterpolator {
    key      [0, 1]
    keyValue [0.000 0.000 0.000 0.000 ,0.000 0.200 0.000 3.143 ]
}
. . .
ROUTE OrientTS120658180 .fraction_changed TO ORI120658180.set_fraction
ROUTE ORI120658180 .value_changed TO T120661744 .rotation
```

Figure 24

ACKNOWLEDGMENT

This work was supported by the PENED99 Project of the Greek Secretariat of Research and Technology and from the P2People EC IST Project.

REFERENCES

[1] "Tutorial issue on the MPEG-4 standard," *Signal Processing: Image Communication*, vol. 15, no. 4-5, pp. 269–478, 2000.
[2] R. Koenen, "MPEG-4 multimedia for our time," *IEEE Spectrum*, vol. 36, no. 2, pp. 26–33, 1999.
[3] L. Chiariglione, "MPEG and multimedia communications," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 5–18, 1997.
[4] F. Pereira, "MPEG-4: Why, what, how and when?," *Signal Processing: Image Communication*, vol. 15, no. 4-5, pp. 271–279, 2000.
[5] S. Boughoufalah, J.-C. Dufourd, and F. Bouilhaguet, "MPEG-Pro, an authoring system for MPEG-4," in *Proc. IEEE Interna-*

tional Symposium on Circuits and Systems (ISCAS 2000), pp. 465–468, Geneva, Switzerland, May 2000.

[6] I. Kompatsiaris, V. K. Papastathis, and M. G. Strintzis, "An authoring tool for the composition of MPEG-4 audiovisual scenes," in *Proc. International Workshop on Synthetic-Natural Hybrid Coding and Three-Dimensional (3D) Imaging*, Santorini, Greece, September 1999.

[7] H. Luo and A. Eleftheriadis, "Designing an interactive tool for video object segmentation and annotation," in *Proc. ACM Multimedia '99*, pp. 265–269, Orlando, Fla, USA, October–November 1999.

[8] P. Correia and F. Pereira, "The role of analysis in content-based video coding and indexing," *Signal Processing*, vol. 66, no. 2, pp. 125–142, 1998.

[9] B. Erol and F. Kossentini, "Automatic key video object plane selection using the shape information in the MPEG-4 compressed domain," *IEEE Trans. Multimedia*, vol. 2, no. 2, pp. 129–138, 2000.

[10] S. Shirani, B. Erol, and F. Kossentini, "A concealment method for shape information in MPEG-4 coded video sequences," *IEEE Trans. Multimedia*, vol. 2, no. 3, pp. 185–190, 2000.

[11] P. Daras, I. Kompatsiaris, T. Raptis, and M. G. Strintzis, "MPEG-4 authoring tool for the composition of 3D audiovisual scenes," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS 2001)*, Sydney, Australia, May 2001.

[12] P. Daras, I. Kompatsiaris, T. Raptis, and M. G. Strintzis, "Authoring tool for the composition of 3D audiovisual scenes using the MPEG-4 standard," in *Proc. International Workshop on Digital and Computational Video (DCV '01)*, Tampa, Fla, USA, February 2001.

[13] Y. Avrithis, A. Doulamis, N. Doulamis, and S. Kollias, "A stochastic framework for optimal key frame extraction from MPEG video databases," *Computer Vision and Image Understanding*, vol. 75, no. 1-2, pp. 3–24, 1999.

[14] B.-L. Yeo and B. Liu, "Rapid scene analysis on compressed video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 533–544, 1995.

[15] G. Tziritas and C. Labit, *Motion Analysis for Image Sequence Coding*, Elsevier, Amsterdam, 1994.

[16] A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: a synopsis of current problems and methods," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.

[17] C. K. Cheong and K. Aizawa, "Structural motion segmentation based on probabilistic clustering," in *Proc. IEEE International Conference on Image Processing*, vol. I, pp. 505–508, Lausanne, Switzerland, September 1996.

[18] F. Moscheni, S. Bhattacharjee, and M. Kunt, "Spatiotemporal segmentation based on region merging," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 9, pp. 897–915, 1998.

[19] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 7, no. 4, pp. 384–401, 1985.

[20] J.-M. Odobez and P. Bouthemy, "Direct incremental model-based image motion segmentation for video analysis," *Signal Processing*, vol. 66, no. 2, pp. 143–155, 1998.

[21] J. Wang and E. Adelson, "Representing moving images with layers," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 625–638, 1994.

[22] I. Grinias and G. Tziritas, "A semi-automatic seeded region growing algorithm for video object localization and tracking," *Signal Processing: Image Communication*, vol. 16, no. 10, pp. 977–986, 2001.

[23] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[24] R. Koenen, "MPEG-4 Overview—(V.16 La BauleVersion)," ISO/IEC JTC1/SC29/WG11 N3747, October 2000.

[25] G. Akrivas, N. Doulamis, A. Doulamis, and S. Kollias, "Scene detection methods for MPEG-encoded video signals," in *Proc. MELECON 2000 Mediterranean Electrotechnical Conference*, Nicosia, Cyprus, May 2000.

[26] Z. Lifshitz, "Status of the systems version 1, 2, 3 software implementation," Tech. Rep. ISO/IEC JTC1/SC29/WG11 N3564, July 2000.

[27] J. Signès, Y. Fisher, and A. Eleftheriadis, "MPEG-4's binary format for scene description," *Signal Processing: Image Communication*, vol. 15, no. 4-5, pp. 321–345, 2000.

[28] ISO/IEC 14472-1, "The virtual reality modeling language," 1997, http://www.web3d.org/fs_specifications.htm.

[29] B. MacIntyre and S. Feiner, "Future multimedia user interfaces," *Multimedia Systems*, vol. 4, no. 5, pp. 250–268, 1996.

[30] ISO/IEC JTC1/SC29/WG11 N2502, FDIS of ISO/IEC 14496-2, "Generic coding of audio-visual object: Part 2—Visual," (MPEG-4), November 1998.

[31] ISO/IEC 14496-5, "ISO/IEC 14496-5 (MPEG-4) Video Reference Software," Tech. Rep., July 2001.

Petros Daras was born in Athens, Greece in 1974 and he is an Associate Researcher at the Informatics and Telematics Institute. He received the Diploma in electrical and computer engineering and the M.S. degree in medical informatics from the Aristotle University of Thessaloniki (AUTH), Greece, in 1999 and 2002, respectively. He is currently pursuing the Ph.D. degree in the area of multimedia from the AUTH. His main research interests include the MPEG-4 standard, streaming, interactive TV, mobile and wireless communication, and medical informatics applications. He has participated as a Researcher and Developer in European and national research projects. P. Daras is a member of the Technical Chamber of Greece.

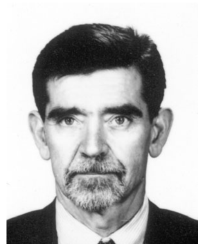
Ioannis Kompatsiaris received the Diploma degree in electrical engineering and the Ph.D. degree in 3D model-based image sequence coding from Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece in 1996 and 2001, respectively. He is a Senior Researcher (Researcher D) at the Informatics and Telematics Institute, Thessaloniki. Prior to his current position, he was a Leading Researcher on 2D and 3D imaging at AUTH. His research interests include computer vision, 3D model-based monoscopic and multiview image sequence analysis and coding, medical image processing, standards (MPEG-4, MPEG-7), and content-based indexing and retrieval. His involvement with those research areas has led to the coauthoring of 9 papers in refereed journals and more than 30 papers in international conferences. He has served as a regular reviewer for a number of international journals and conferences. Since 1996, he has been involved in more than 8 projects in Greece, funded by the EC and the Greek Ministry of Research and Technology. I. Kompatsiaris is an IEEE member and a member of the Technical Chamber of Greece.

Ilias Grinias was born in Larissa, Greece in 1973 and received the B.S. and the M.S. degrees in computer science from the University of Crete in 1997 and 1999, respectively. His research interests include image analysis, pattern recognition, data bases, information systems and telecommunications. His involvement with those research areas has led to the coauthoring of 3 papers in refereed journals and 4 papers in international conferences.

Georgios Akrivas received his Diploma from the Department of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Greece in 1999. He is currently working toward the Ph.D. degree in intelligent video retrieval in the Image, Video and Multimedia Systems Laboratory at NTUA. His research interests include multimedia content description, computer vision, intelligent retrieval, and fuzzy logic.

Georgios Tziritas received the Diploma of electrical engineering from the Technical University of Athens in 1977, the Diplome d'Etudes Approfondies (DEA), the Diplome de Docteur Ingenieur, and the Diplome de Docteur d'Etat from the Institut Polytechnique de Grenoble in 1978, 1981, and 1985, respectively. From 1982, he was a Researcher at the Centre National de la Recherche Scientifique, the Centre d'Etudes des Phenomenes Aleatoires et Geophysiques (CEPHAG) until 1985, the Institut National de Recherche en Informatique et Automatique (INRIA) until 1987, and the Laboratoire des Signaux et Systemes (LSS). From September 1992, he is an Associate Professor at the Department of Computer Science, University of Crete, teaching digital signal processing, digital image processing, digital video processing, and information and coding theory. G. Tziritas is the Coauthor (with C. Labit) of *Motion Analysis for Image Sequence Coding* (Elsevier, 1994), and of more than 70 journal and conference papers on signal and image processing, and image and video analysis. His research interests are in the areas of signal processing, image processing and analysis, computer vision, motion analysis, image and video indexing, and image and video communication.

Stefanos Kollias was born in Athens in 1956. He obtained his Diploma from National Technical University of Greece (NTUA) in 1979, his M.S. degree in communication engineering in 1980 from University of Manchester Institute of Science and Technology (UMIST) in England, and his Ph.D. degree in signal processing from the Computer Science Division of NTUA. He has been with the Electrical Engineering Department of NTUA since 1986 where he serves now as a Professor. Since 1990, he has been Director of the Image, Video and Multimedia Systems Laboratory of NTUA. He has published more than 120 papers in the above fields, 50 of which in international journals. He has been a member of the Technical or Advisory Committee or invited speaker in 40 international conferences. He is a Reviewer of 10 IEEE Transactions and 10 other journals. Ten graduate students have completed their Doctorate under his supervision, while other ten are currently performing their Ph.D. thesis. He and his team have been participating in 38 European and national projects.



Michael G. Strintzis received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1967, the M.A., and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1969 and 1970, respectively. He then joined the Electrical Engineering Department at the University of Pittsburgh, Pittsburgh, Pa, where he served as Assistant Professor (1970–1976) and an Associate Professor (1976–1980). Since 1980, he has been a Professor of electrical and computer engineering at the University of Thessaloniki, Thessaloniki, Greece, and since 1999, Director of the Informatics and Telematics Research Institute, Thessaloniki. His current research interests include 2D and 3D image coding, image processing, biomedical signal and image processing, and DVD and Internet data authentication and copy protection. Dr. Strintzis has served as Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology since 1999. In 1984, he was awarded one of the Centennial Medals of the IEEE.

Special Issue on Multirate Systems and Applications

Call for Papers

Filter banks for the application of subband coding of speech were introduced in the 1970s. Since then, filter banks and multirate systems have been studied extensively. There has been great success in applying multirate systems to many applications. The most notable of these applications include subband coding for audio, image, and video, signal analysis and representation using wavelets, subband denoising, and so forth. Different applications also call for different filter bank designs and the topic of designing one-dimensional and multidimentional filter banks for specific applications has been of great interest.

Recently there has been growing interest in applying multirate theories to the area of communication systems such as, transmultiplexers, filter bank transceivers, blind deconvolution, and precoded systems. There are strikingly many dualities and similarities between multirate systems and multicarrier communication systems. Many problems in multicarrier transmission can be solved by extending results from multirate systems and filter banks. This exciting research area is one that is of increasing importance.

The aim of this special issue is to bring forward recent developments on filter banks and the ever-expanding area of applications of multirate systems.

Topics of interest include (but are not limited to):

- Multirate signal processing for communications
- Filter bank transceivers
- One-dimensional and multidimensional filter bank designs for specific applications
- Denoising
- Adaptive filtering
- Subband coding
- Audio, image, and video compression
- Signal analysis and representation
- Feature extraction and classification
- Other applications

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	January 1, 2006
Acceptance Notification	May 1, 2006
Final Manuscript Due	August 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Yuan-Pei Lin, Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan; ypl@mail.nctu.edu.tw

See-May Phoong, Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan; smp@cc.ee.ntu.edu.tw

Ivan Selesnick, Department of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201, USA; selesi@poly.edu

Soontorn Oraintara, Department of Electrical Engineering, The University of Texas at Arlington, Arlington, TX 76010, USA; oraintar@uta.edu

Gerald Schuller, Fraunhofer Institute for Digital Media Technology (IDMT), Langewiesener Strasse 22, 98693 Ilmenau, Germany; shl@idmt.fraunhofer.de

Special Issue on

Multisensor Processing for Signal Extraction and Applications

Call for Papers

Source signal extraction from heterogeneous measurements has a wide range of applications in many scientific and technological fields, for example, telecommunications, speech and acoustic signal processing, and biomedical pattern analysis. Multiple signal reception through multisensor systems has become an effective means for signal extraction due to its superior performance over the monosensor mode. Despite the rapid progress made in multisensor-based techniques in the past few decades, they continue to evolve as key technologies in modern wireless communications and biomedical signal processing. This has led to an increased focus by the signal processing community on the advanced multisensor-based techniques which can offer robust high-quality signal extraction under realistic assumptions and with minimal computational complexity. However, many challenging tasks remain unresolved and merit further rigorous studies. Major efforts in developing advanced multisensor-based techniques may include high-quality signal extraction, realistic theoretical modeling of real-world problems, algorithm complexity reduction, and efficient real-time implementation.

The purpose of this special issue aims to present state-of-the-art multisensor signal extraction techniques and applications. Contributions in theoretical study, performance analysis, complexity reduction, computational advances, and real-world applications are strongly encouraged.

Topics of interest include (but are not limited to):

- Multiantenna processing for radio signal extraction
- Multimicrophone speech recognition and enhancement
- Multisensor radar, sonar, navigation, and biomedical signal processing
- Blind techniques for multisensor signal extraction
- Computational advances in multisensor processing

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	January 1, 2006
Acceptance Notification	May 1, 2006
Final Manuscript Due	August 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Chong-Yung Chi, National Tsing Hua University, Taiwan; cychi@ee.nthu.edu.tw

Ta-Sung Lee, National Chiao Tung University, Taiwan; tslee@cc.nctu.edu.tw

Zhi-Quan Luo, University of Minnesota, USA; luozq@ece.umn.edu

Kung Yao, University of California, Los Angeles, USA; yao@ee.ucla.edu

Yue Wang, Virginia Polytechnic Institute and State University, USA; yuewang@vt.edu

Special Issue on

Search and Retrieval of 3D Content and Associated Knowledge Extraction and Propagation

Call for Papers

With the general availability of 3D digitizers, scanners, and the technology innovation in 3D graphics and computational equipment, large collections of 3D graphical models can be readily built up for different applications (e.g., in CAD/CAM, games design, computer animations, manufacturing and molecular biology). For such large databases, the method whereby 3D models are sought merits careful consideration. The simple and efficient query-by-content approach has, up to now, been almost universally adopted in the literature. Any such method, however, must first deal with the proper positioning of the 3D models. The two prevalent-in-the-literature methods for the solution to this problem seek either

- Pose Normalization: Models are first placed into a canonical coordinate frame (normalizing for translation, scaling, and rotation). Then, the best measure of similarity is found by comparing the extracted feature vectors, or
- Descriptor Invariance: Models are described in a transformation invariant manner, so that any transformation of a model will be described in the same way, and the best measure of similarity is obtained at any transformation.

The existing 3D retrieval systems allow the user to perform queries by example. The queried 3D model is then processed, low-level geometrical features are extracted, and similar objects are retrieved from a local database. A shortcoming of the methods that have been proposed so far regarding the 3D object retrieval, is that neither is the semantic information (high-level features) attached to the (low-level) geometric features of the 3D content, nor are the personalization options taken into account, which would significantly improve the retrieved results. Moreover, few systems exist so far to take into account *annotation* and *relevance feedback* techniques, which are very popular among the corresponding content-based image retrieval systems (CBIR).

Most existing CBIR systems using knowledge either annotate all the objects in the database (full annotation) or annotate a subset of the database manually selected (partial annotation). As the database becomes larger, full annotation is increasingly difficult because of the manual effort needed. Partial annotation is relatively affordable and trims down the heavy manual labor. Once the database is partially annotated, traditional image analysis methods are used to derive semantics of the objects not yet annotated. However, it is not clear "how much" annotation is sufficient for a specific database and what the best subset of objects to annotate is. In other words how the knowledge *will be propagated*. Such techniques have not been presented so far regarding the 3D case.

Relevance feedback was first proposed as an interactive tool in text-based retrieval. Since then it has been proven to be a powerful tool and has become a major focus of research in the area of content-based search and retrieval. In the traditional computer centric approaches, which have been proposed so far, the "best" representations and weights are fixed and they cannot effectively model high-level concepts and user's perception subjectivity. In order to overcome these limitations of the computer centric approach, techniques based on *relevant feedback*, in which the human and computer interact to refine high-level queries to representations based on low-level features, should be developed.

The aim of this special issue is to focus on recent developments in this expanding research area. The special issue will focus on novel approaches in 3D object retrieval, transforms and methods for efficient geometric feature extraction, annotation and relevance feedback techniques, knowledge propagation (e.g., using Bayesian networks), and their combinations so as to produce a single, powerful, and dominant solution.

Topics of interest include (but are not limited to):

- 3D content-based search and retrieval methods (volume/surface-based)
- Partial matching of 3D objects
- Rotation invariant feature extraction methods for 3D objects

- Graph-based and topology-based methods
- 3D data and knowledge representation
- Semantic and knowledge propagation over heterogeneous metadata types
- Annotation and relevance feedback techniques for 3D objects

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	February 1, 2006
Acceptance Notification	June 1, 2006
Final Manuscript Due	September 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Tsuhan Chen, Carnegie Mellon University, Pittsburgh, PA 15213, USA; tsuhan@cmu.edu

Ming Ouhyoung, National Taiwan University, Taipei 106, Taiwan; ming@csie.ntu.edu.tw

Petros Daras, Informatics and Telematics Institute, Centre for Research and Technology Hellas, 57001 Thermi, Thessaloniki, Greece; daras@iti.gr

Special Issue on Robust Speech Recognition

Call for Papers

Robustness can be defined as the ability of a system to maintain performance or degrade gracefully when exposed to conditions not well represented in the data used to develop the system. In automatic speech recognition (ASR), systems must be robust to many forms of signal degradation, including speaker characteristics (e.g., dialect and accent), ambient environment (e.g., cellular telephony), transmission channel (e.g., voice over IP), and language (e.g., new words, dialect switching). Robust ASR systems, which have been under development for the past 35 years, have made great progress over the years closing the gap between performance on pristine research tasks and noisy operational data.

However, in recent years, demand is emerging for a new class of systems that tolerate extreme and unpredictable variations in operating conditions. For example, in a cellular telephony environment, there are many nonstationary forms of noise (e.g., multiple speakers) and significant variations in microphone type, position, and placement. Harsh ambient conditions typical in automotive and mobile applications pose similar challenges. Development of systems in a language or dialect for which there is limited or no training data in a target language has become a critical issue for a new generation of voice mining applications. The existence of multiple conditions in a single stream, a situation common to broadcast news applications, and that often involves unpredictable changes in speaker, topic, dialect, or language, is another form of robustness that has gained attention in recent years.

Statistical methods have dominated the field since the early 1980s. Such systems tend to excel at learning the characteristics of large databases that represent good models of the operational conditions and do not generalize well to new environments.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Channel and microphone normalization
- Stationary and nonstationary noise modeling, compensation, and/or rejection
- Localization and separation of sound sources (including speaker segregation)
- Signal processing and feature extraction for applications involving hands-free microphones
- Noise robust speech modeling
- Adaptive training techniques
- Rapid adaptation and learning
- Integration of confidence scoring, metadata, and other alternative information sources
- Audio-visual fusion
- Assessment relative to human performance
- Machine learning algorithms for robustness
- Transmission robustness
- Pronunciation modeling

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	February 1, 2006
Acceptance Notification	June 1, 2006
Final Manuscript Due	September 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Herve Bourlard, IDIAP Research Institute, Swiss Federal Institute of Technology at Lausanne (EPFL), 1920 Martigny, Switzerland; herve.bourlard@idiap.ch

Mark Gales, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK; mjfg@eng.cam.ac.uk

Maurizio Omologo, ITC-IRST, 38050 Trento, Italy; omologo@itc.it

S. Parthasarathy, AT&T Labs - Research, NJ 07748, USA; sps@research.att.com

Joe Picone, Department of Electrical and Computer Engineering, Mississippi State University, MS 39762-9571, USA; picone@cavs.msstate.edu

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on

Signal Processing Technologies for Ambient Intelligence in Home-Care Applications

Call for Papers

The possibility of allowing elderly people with different kinds of disabilities to conduct a normal life at home and achieve a more effective inclusion in the society is attracting more and more interest from both industrial and governmental bodies (hospitals, healthcare institutions, and social institutions). Ambient intelligence technologies, supported by adequate networks of sensors and actuators, as well as by suitable processing and communication technologies, could enable such an ambitious objective.

Recent researches demonstrated the possibility of providing constant monitoring of environmental and biomedical parameters, and the possibility to autonomously originate alarms, provide primary healthcare services, activate emergency calls, and rescue operations through distributed assistance infrastructures. Nevertheless, several technological challenges are still connected with these applications, ranging from the development of enabling technologies (hardware and software), to the standardization of interfaces, the development of intuitive and ergonomic human-machine interfaces, and the integration of complex systems in a highly multidisciplinary environment.

The objective of this special issue is to collect the most significant contributions and visions coming from both academic and applied research bodies working in this stimulating research field. This is a highly interdisciplinary field comprising many areas, such as signal processing, image processing, computer vision, sensor fusion, machine learning, pattern recognition, biomedical signal processing, multimedia, human-computer interfaces, and networking.

The focus will be primarily on the presentation of original and unpublished works dealing with ambient intelligence and domotic technologies that can enable the provision of advanced homecare services.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Video-based monitoring of domestic environments and users
- Continuous versus event-driven monitoring
- Distributed information processing
- Data fusion techniques for event association and automatic alarm generation
- Modeling, detection, and learning of user habits for automatic detection of anomalous behaviors
- Integration of biomedical and behavioral data
- Posture and gait recognition and classification
- Interactive multimedia communications for remote assistance
- Content-based encoding of medical and behavioral data
- Networking support for remote healthcare
- Intelligent/natural man-machine interaction, personalization, and user acceptance

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	March 1, 2006
Acceptance Notification	July 1, 2006
Final Manuscript Due	October 1, 2006
Publication Date	1st Quarter, 2007

GUEST EDITORS:

Francesco G. B. De Natale, Department of Information and Communication Technology, University of Trento, Via Sommarive 14, 38050 Trento, Italy; denatale@ing.unitn.it

Aggelos K. Katsaggelos, Department of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3118, USA; aggk@ece.northwestern.edu

Oscar Mayora, Create-Net Association, Via Solteri 38, 38100 Trento, Italy; oscar.mayora@create-net.it

Ying Wu, Department of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3118, USA; yingwu@ece.northwestern.edu

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on Spatial Sound and Virtual Acoustics

Call for Papers

Spatial sound reproduction has become widespread in the form of multichannel audio, particularly through home theater systems. Reproduction systems from binaural (by headphones) to hundreds of loudspeaker channels (such as wave field synthesis) are entering practical use. The application potential of spatial sound is much wider than multichannel sound, however, and research in the field is active. Spatial sound covers for example the capturing, analysis, coding, synthesis, reproduction, and perception of spatial aspects in audio and acoustics.

In addition to the topics mentioned above, research in virtual acoustics broadens the field. Virtual acoustics includes techniques and methods to create realistic percepts of sound sources and acoustic environments that do not exist naturally but are rendered by advanced reproduction systems using loudspeakers or headphones. Augmented acoustic and audio environments contain both real and virtual acoustic components.

Spatial sound and virtual acoustics are among the major research and application areas in audio signal processing. Topics of active study range from new basic research ideas to improvement of existing applications. Understanding of spatial sound perception by humans is also an important area, in fact a prerequisite to advanced forms of spatial sound and virtual acoustics technology.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Multichannel reproduction
- Wave field synthesis
- Binaural reproduction
- Format conversion and enhancement of spatial sound
- Spatial sound recording
- Analysis, synthesis, and coding of spatial sound
- Spatial sound perception and auditory modeling
- Simulation and modeling of room acoustics
- Auralization techniques
- Beamforming and sound source localization
- Acoustic and auditory scene analysis
- Augmented reality audio
- Virtual acoustics (sound environments and sources)
- Intelligent audio environments
- Loudspeaker-room interaction and equalization
- Applications

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	May 1, 2006
Acceptance Notification	September 1, 2006
Final Manuscript Due	December 1, 2006
Publication Date	1st Quarter, 2007

GUEST EDITORS:

Ville Pulkki, Helsinki University of Technology, Espoo, Finland; ville@acoustics.hut.fi

Christof Faller, EPFL, Lausanne, Switzerland; christof.faller@epfl.ch

Aki Harma, Philips Research Labs, Eindhoven, The Netherlands; aki.harma@philips.com

Tapio Lokki, Helsinki University of Technology, Espoo, Finland; ktlokki@cc.hut.fi

Werner de Bruijn, Philips Research Labs, Eindhoven, The Netherlands; werner.de.bruijn@philips.com

Special Issue on

Advances in Electrocardiogram Signal Processing and Analysis

Call for Papers

Since its invention in the 19th century when it was little more than a scientific curiosity, the electrocardiogram (ECG) has developed into one of the most important and widely used quantitative diagnostic tools in medicine. It is essential for the identification of disorders of the cardiac rhythm, extremely useful for the diagnosis and management of heart abnormalities such as myocardial infarction (heart attack), and offers helpful clues to the presence of generalised disorders that affect the rest of the body, such as electrolyte disturbances and drug intoxication.

Recording and analysis of the ECG now involves a considerable amount of signal processing for S/N enhancement, beat detection, automated classification, and compression. These involve a whole variety of innovative signal processing methods, including adaptive techniques, time-frequency and time-scale procedures, artificial neural networks and fuzzy logic, higher-order statistics and nonlinear schemes, fractals, hierarchical trees, Bayesian approaches, and parametric models, amongst others.

This special issue will review the current status of ECG signal processing and analysis, with particular regard to recent innovations. It will report major achievements of academic and commercial research institutions and individuals, and provide an insight into future developments within this exciting and challenging area.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Beat (QRS complex) detection
- ECG compression
- Denoising of ECG signals
- Morphological studies and classification
- ECG modeling techniques
- Expert systems and automated diagnosis
- QT interval measurement and heart-rate variability
- Arrhythmia and ischemia detection and analysis
- Interaction between cardiovascular signals (ECG, blood pressure, respiration, etc.)
- Intracardiac ECG analysis (implantable cardiovascular devices, and pacemakers)
- ECGs and sleep apnoea
- Real-time processing and instrumentation
- ECG telemedicine and e-medicine
- Fetal ECG detection and analysis
- Computational tools and databases for ECG education and research

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	May 1, 2006
Acceptance Notification	September 1, 2006
Final Manuscript Due	December 1, 2006
Publication Date	1st Quarter, 2007

GUEST EDITORS:

William Sandham, Scotsig, Glasgow G12 9pf, UK; w.sandham@scotsig.co.uk

David Hamilton, Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, UK; d.hamilton@eee.strath.ac.uk

Pablo Laguna Lasaosa, Departmento de Ingeniería Electrónica y Communicaciones, Universidad de Zaragoza, 50015 Zaragoza, Spain; laguna@unizar.es

Maurice Cohen, University of California, San Francisco, USA; mcohen@fresno.ucsf.edu

Special Issue on

Emerging Signal Processing Techniques for Power Quality Applications

Call for Papers

Recently, end users and utility companies are increasingly concerned with perturbations originated from electrical power quality variations. Investigations are being carried out to completely characterize not only the old traditional type of problems, but also new ones that have arisen as a result of massive use of nonlinear loads and electronics-based equipment in residences, commercial centers, and industrial plants. These nonlinear load effects are aggravated by massive power system interconnections, increasing number of different power sources, and climatic changes.

In order to improve the capability of equipments applied to monitoring the power quality of transmission and distribution power lines, power systems have been facing new analysis and synthesis paradigms, mostly supported by signal processing techniques. The analysis and synthesis of emerging power quality and power system problems led to new research frontiers for the signal processing community, focused on the development and combination of computational intelligence, source coding, pattern recognition, multirate systems, statistical estimation, adaptive signal processing, and other digital processing techniques, implemented in either DSP-based, PC-based, or FPGA-based solutions.

The goal of this proposal is to introduce powerful and efficient real-time or almost-real-time signal processing tools for dealing with the emerging power quality problems. These techniques take into account power-line signals and complementary information, such as climatic changes.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Detection of transients
- Classification of multiple events
- Identification of isolated and multiple disturbance sources
- Compression of voltage and current data signals
- Location of disturbance sources
- Prediction of transmission and distribution systems failures
- Demand forecasting
- Parameters estimation for fundamental, harmonics, and interharmonics

Digital signal processing techniques applied to power quality applications are a very attractive and stimulating area of research. Its results will provide, in the near future, new standards for the decentralized and real-time monitoring of transmission and distribution systems, allowing to closely follow and predict power system performance. As a result, the power systems will be more easily planned, expanded, controlled, managed, and supervised.

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	May 1, 2006
Acceptance Notification	September 1, 2006
Final Manuscript Due	December 1, 2006
Publication Date	2nd Quarter, 2007

GUEST EDITORS:

Moisés Vidal Ribeiro, Department of Electrical Circuit, Federal University of Juiz de Fora, CEP 36036-330, Juiz de Fora, Brazil; mribeiro@ieee.org

Jacques Szczupack, Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, CEP 22453-900, Rio de Janeiro, Brazil; jacques@ele.puc-rio.br

M. Reza Iravani, The Edward S. Rogers SR., Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada M5S 3G4; aki.iravani@ecf.utoronto.ca

Irene Yu-Hua Gu, Department of Signals and Systems, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden; irenegu@s2.chalmers.se

Pradipta Kishore Dash, C. V. Raman, College of Engineering Bhubaneswar, Khurda-752054, Orissa, India; pkdash_india@yahoo.com

Alexander Mamishev, Department of Electrical Engineering, University of Washington, WA 98195-2500, Seattle, USA; mamishev@ee.washington.edu

Hindawi Publishing Corporation
http://www.hindawi.com

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on Super-resolution Enhancement of Digital Video

Call for Papers

When designing a system for image acquisition, there is generally a desire for high spatial resolution and a wide field-of-view. To achieve this, a camera system must typically employ small f-number optics. This produces an image with very high spatial-frequency bandwidth at the focal plane. To avoid aliasing caused by undersampling, the corresponding focal plane array (FPA) must be sufficiently dense. However, cost and fabrication complexities may make this impractical. More fundamentally, smaller detectors capture fewer photons, which can lead to potentially severe noise levels in the acquired imagery. Considering these factors, one may choose to accept a certain level of undersampling or to sacrifice some optical resolution and/or field-of-view.

In image super-resolution (SR), postprocessing is used to obtain images with resolutions that go beyond the conventional limits of the uncompensated imaging system. In some systems, the primary limiting factor is the optical resolution of the image in the focal plane as defined by the cut-off frequency of the optics. We use the term "optical SR" to refer to SR methods that aim to create an image with valid spatial-frequency content that goes beyond the cut-off frequency of the optics. Such techniques typically must rely on extensive a priori information. In other image acquisition systems, the limiting factor may be the density of the FPA, subsequent postprocessing requirements, or transmission bitrate constraints that require data compression. We refer to the process of overcoming the limitations of the FPA in order to obtain the full resolution afforded by the selected optics as "detector SR." Note that some methods may seek to perform both optical and detector SR.

Detector SR algorithms generally process a set of low-resolution aliased frames from a video sequence to produce a high-resolution frame. When subpixel relative motion is present between the objects in the scene and the detector array, a unique set of scene samples are acquired for each frame. This provides the mechanism for effectively increasing the spatial sampling rate of the imaging system without reducing the physical size of the detectors.

With increasing interest in surveillance and the proliferation of digital imaging and video, SR has become a rapidly growing field. Recent advances in SR include innovative algorithms, generalized methods, real-time implementations, and novel applications. The purpose of this special issue is to present leading research and development in the area of super-resolution for digital video. Topics of interest for this special issue include but are not limited to:

- Detector and optical SR algorithms for video
- Real-time or near-real-time SR implementations
- Innovative color SR processing
- Novel SR applications such as improved object detection, recognition, and tracking
- Super-resolution from compressed video
- Subpixel image registration and optical flow

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	September 1, 2006
Acceptance Notification	February 1, 2006
Final Manuscript Due	April 15, 2007
Publication Date	3rd Quarter, 2007

GUEST EDITORS:

Russell C. Hardie, Department of Electrical and Computer Engineering, University of Dayton, 300 College Park, Dayton, OH 45469-0026, USA; rhardie@udayton.edu

Richard R. Schultz, Department of Electrical Engineering, University of North Dakota, Upson II Room 160, P.O. Box 7165, Grand Forks, ND 58202-7165, USA; RichardSchultz@mail.und.nodak.edu

Kenneth E. Barner, Department of Electrical and Computer Engineering, University of Delaware, 140 Evans Hall, Newark, DE 19716-3130, USA; barner@ee.udel.edu

NEWS RELEASE

Nominations Invited for the Institute of Acoustics 2006 A B Wood Medal

The Institute of Acoustics, the UK's leading professional body for those working in acoustics, noise and vibration, is inviting nominations for its prestigious A B Wood Medal for the year 2006.

The A B Wood Medal and prize is presented to an individual, usually under the age of 35, for distinguished contributions to the application of underwater acoustics. The award is made annually, in even numbered years to a person from Europe and in odd numbered years to someone from the USA/Canada. The 2005 Medal was awarded to Dr A Thode from the USA for his innovative, interdisciplinary research in ocean and marine mammal acoustics.

Nominations should consist of the candidate's CV, clearly identifying peer reviewed publications, and a letter of endorsement from the nominator identifying the contribution the candidate has made to underwater acoustics. In addition, there should be a further reference from a person involved in underwater acoustics and not closely associated with the candidate. Nominees should be citizens of a European Union country for the 2006 Medal. Nominations should be marked confidential and addressed to the President of the Institute of Acoustics at 77A St Peter's Street, St. Albans, Herts, AL1 3BN. The deadline for receipt of nominations is **15 October 2005**.

Dr Tony Jones, President of the Institute of Acoustics, comments, "A B Wood was a modest man who took delight in helping his younger colleagues. It is therefore appropriate that this prestigious award should be designed to recognise the contributions of young acousticians."

Further information and an nomination form can be found on the Institute's website at www.ioa.org.uk.

A B Wood

Albert Beaumont Wood was born in Yorkshire in 1890 and graduated from Manchester University in 1912. He became one of the first two research scientists at the Admiralty to work on antisubmarine defence. He designed the first directional hydrophone and was well known for the many contributions he made to the science of underwater acoustics and for the help he gave to younger colleagues. The medal was instituted after his death by his many friends on both sides of the Atlantic and was administered by the Institute of Physics until the formation of the Institute of Acoustics in 1974.

PRESS CONTACT

Judy Edrich
Publicity & Information Manager, Institute of Acoustics
Tel: 01727 848195; E-mail: judy.edrich@ioa.org.uk

EDITORS NOTES

The Institute of Acoustics is the UK's professional body for those working in acoustics, noise and vibration. It was formed in 1974 from the amalgamation of the Acoustics Group of the Institute of Physics and the British Acoustical Society (a daughter society of the Institution of Mechanical Engineers). The Institute of Acoustics is a nominated body of the Engineering Council, offering registration at Chartered and Incorporated Engineer levels.

The Institute has some 2500 members from a rich diversity of backgrounds, with engineers, scientists, educators, lawyers, occupational hygienists, architects and environmental health officers among their number. This multidisciplinary culture provides a productive environment for cross-fertilisation of ideas and initiatives. The range of interests of members within the world of acoustics is equally wide, embracing such aspects as aerodynamics, architectural acoustics, building acoustics, electroacoustics, engineering dynamics, noise and vibration, hearing, speech, underwater acoustics, together with a variety of environmental aspects. The lively nature of the Institute is demonstrated by the breadth of its learned society programmes.

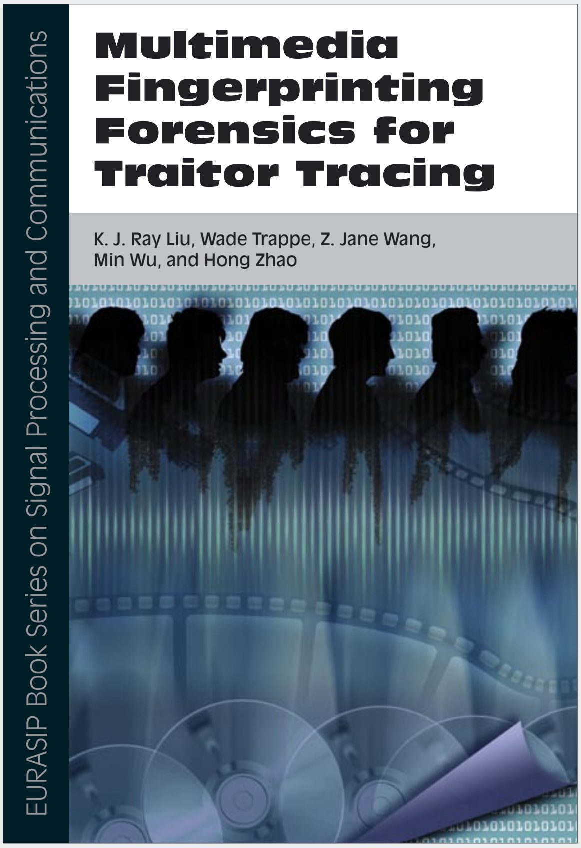
For more information please visit our site at www.ioa.org.uk.

EURASIP Book Series On Signal Processing and Communications

EURASIP Book Series on SP&C, Volume 4, ISBN 977-5945-18-6

MULTIMEDIA FINGERPRINTING FORENSICS FOR TRAITOR TRACING

Edited by: K. J. Ray Liu, Wade Trappe, Z. Jane Wang, Min Wu, and Hong Zhao



The popularity of multimedia content has led to the widespread distribution and consumption of digital multimedia data. As a result of the relative ease with which individuals may now alter and repackage digital content, ensuring that media content is employed by authorized users for its intended purpose is becoming an issue of eminent importance to both governmental security and commercial applications. Digital fingerprinting is a class of multimedia forensic technologies to track and identify entities involved in the illegal manipulation and unauthorized usage of multimedia content, thereby protecting the sensitive nature of multimedia data as well as its commercial value after the content has been delivered to a recipient.

"Multimedia Fingerprinting Forensics for Traitor Tracing" covers the essential aspects of research in this emerging technology, and explains the latest development in this field. It describes the framework of multimedia fingerprinting, discusses the challenges that may be faced when enforcing usage polices, and investigates the design of fingerprints that cope with new families of multiuser attacks that may be mounted against media fingerprints. The discussion provided in the book highlights challenging problems as well as future trends in this research field, providing readers with a broader view of the evolution of the young field of multimedia forensics.

Topics and features:

Comprehensive coverage of digital watermarking and fingerprinting in multimedia forensics for a number of media types; Detailed discussion on challenges in multimedia fingerprinting and analysis of effective multiuser collusion attacks on digital fingerprinting; Thorough investigation of fingerprint design and performance analysis for addressing different application concerns arising in multimedia fingerprinting; Well-organized explanation of problems and solutions, such as order-statistics-based nonlinear collusion attacks, efficient detection and identification of colluders, group-oriented fingerprint design, and anticollusion codes for multimedia fingerprinting.

For more information and online orders please visit http://www.hindawi.com/books/spc/volume-4/
For any inquires on how to order this title please contact books.orders@hindawi.com

The EURASIP Book Series on Signal Processing and Communications publishes monographs, edited volumes, and textbooks on Signal Processing and Communications. For more information about the series please visit: http://hindawi.com/books/spc/about.html