

# MPEG-4 PC: Authoring and Playing of MPEG-4 Content

P. GERKEN<sup>1</sup>, St. SCHULTZ<sup>1</sup>, G. KNABE<sup>1</sup>, F. CASALINO<sup>2</sup>, G. DI CAGNO<sup>2</sup>, M. QUAGLIA<sup>2</sup>, J.-C. DUFOURD<sup>3</sup>, S. BOUGHOUFALAH<sup>3</sup>, F. BOUILHAGUET<sup>3</sup>, M. STEPPING<sup>4</sup>, Th. BONSE<sup>4</sup>, U. MAYER<sup>5</sup>, J. DEICKE<sup>5</sup>, and M. GLESNER<sup>5</sup>

<sup>1</sup>*Q-Team Dr. Knabe Gesellschaft für Informations- und Qualifikationssysteme mbH, Brauereistr. 11, D-41352 Korschenbroich, Germany*

*Tel: +49 2161 6181 0; Fax: +49 2161 6181 81; E-mail: info@q-team.de*

*URL: <http://www.q-team.de>*

<sup>2</sup>*CSELT, Torino, Italy*

<sup>3</sup>*ENST, Paris, France*

<sup>4</sup>*FernUni Hagen/FTK, Hagen/Dortmund, Germany*

<sup>5</sup>*TU Darmstadt, Darmstadt, Germany*

**Abstract.** Goals and objectives of the European project “MPEG-4 PC - MPEG-4 System Implementation and Tools for Personal Computers (ESPRIT #23191) are presented. An overview of the project’s developments for authoring and playing, both from local and remote sources, of content that conforms to the new multimedia standard ISO/IEC:14496, nicknamed MPEG-4, is given. Example applications for proving the functionalities and the powerfulness of the MPEG-4 standard, that is expected to become of great importance in the media world, are presented. A new milestone of the project’s developments is planned to be demonstrated during the EMMSEC99 Conference.

## 1. Introduction

A new standard, ISO/IEC:14496, nicknamed MPEG-4, [1] for the representation of audio-visual information in multimedia applications is being developed by the International Organisation for Standardisation, ISO, and the International Electrotechnical Committee, IEC. It considers the growing-together of the three market areas TV, computer and telecommunication and will support e.g. future ways of interactive TV. Benefits and the impact of the new standard are reported e.g. in [2]. MPEG-4 specifies essentially:

- (1) the coded representation of audio-visual information as visual and audio objects rather than as rectangular image frames and mixed sound signals, respectively.
- (2) a binary format for scene descriptions (BIFS) in the kind of computer programs.
- (3) a network-transparent transport mechanism in the context of the so-called MPEG-4 Delivery Multimedia Integration Framework, DMIF.
- (4) an interface for intellectual properties management and protection (IPMP) tools.

With these specifications the end-user will have, in contrast to conventional TV applications, the possibility to interact with the content, to the extent enabled by the author. Not only the author but also the end-user can have access to the single audio and visual objects rather than to the scene at large. The objects can even be called from different, local and remote, sources. Visual objects, that can even have arbitrary shapes, are e.g. the 2-D projections of persons in the scene. Audio objects are e.g. the voice of a single person or the sound of a single instrument in an orchestra. The IPMP interface makes the new standard especially attractive for authors. This way a high acceptance of MPEG-4 can be expected not only for

compression purposes but also, and particularly, in the production area. Thus, MPEG-4 will provide the means for a new generation of (interactive) content. Before MPEG-4, every new media standard “merely“ provided a new technology, e.g. digital TV; there was no need but also no possibility to build up a novel kind of content. MPEG-4 however provides a new philosophy for content creation. The new functionalities give novel features for the author and the end-user to interact with the content. In order to exploit these features, novel tools for authoring and playing this new kind of content are needed.

The European project *MPEG-4 PC* (ESPRIT#23191, full notation: “MPEG-4 System Implementation and Tools for Personal Computers“) [3] deals with the first implementation of such an MPEG-4 authoring/playing system on a PC. Goals of the project are:

- The development of an authoring system that supports the MPEG-4 functionalities, especially interactivity, access to and manipulation of audio-visual objects.
- The development of a playing station, a DMIF server/client chain and an application server for the playback of, local and remote, 2-dimensional MPEG-4 content.
- The promotion of the MPEG-4 standard and of the project achievements. For this purpose, demonstrations of prototype MPEG-4 applications are developed.
- Contribution to and even shaping of the MPEG-4 standard.

In the following sections, details about the authoring system (Section 2), the playing station including the broadcast modules (Section 3), and the demonstrator (Section 4) are reported. Short conclusions (Section 5) complete this paper.

## 2. The Authoring System

For reasons of conciseness, only principal elements are described here.

### 2.1 User Interface

Figure 1 shows a screenshot of the preliminary user interface with the following elements:

- The Editor with
  - The Object List  
This is a hierarchical listing of the objects in a scene. By clicking on an object icon certain operations can be performed: changing properties of an object, copying of an object together with all objects on lower hierarchy levels into a library (see below), renaming, removing, or adding behaviour (mouse sensitivity, draggability, visibility, blinking, movement) to an object. All objects are listed in the Object List, but only selected objects are displayable.
  - The Tree View  
This is a visualisation of the scene construction in form of a tree. Note that the specification of an MPEG-4 scene description is hierarchical so that a scene can be well represented by a tree. Clicking on the tree nodes, so-called BIFS nodes, allows various types of manipulation. By shifting nodes or subtrees, the arrangement of the whole scene can be manipulated.
  - The Routes Editor  
This editor defines the relations between the objects in a scene and the BIFS nodes according to the specification of the MPEG-4 scene description. With this editor, also various kinds of interactivity with an object can be defined.
  - The Exported Events Editor  
This editor sets a list of input/output events for an object.

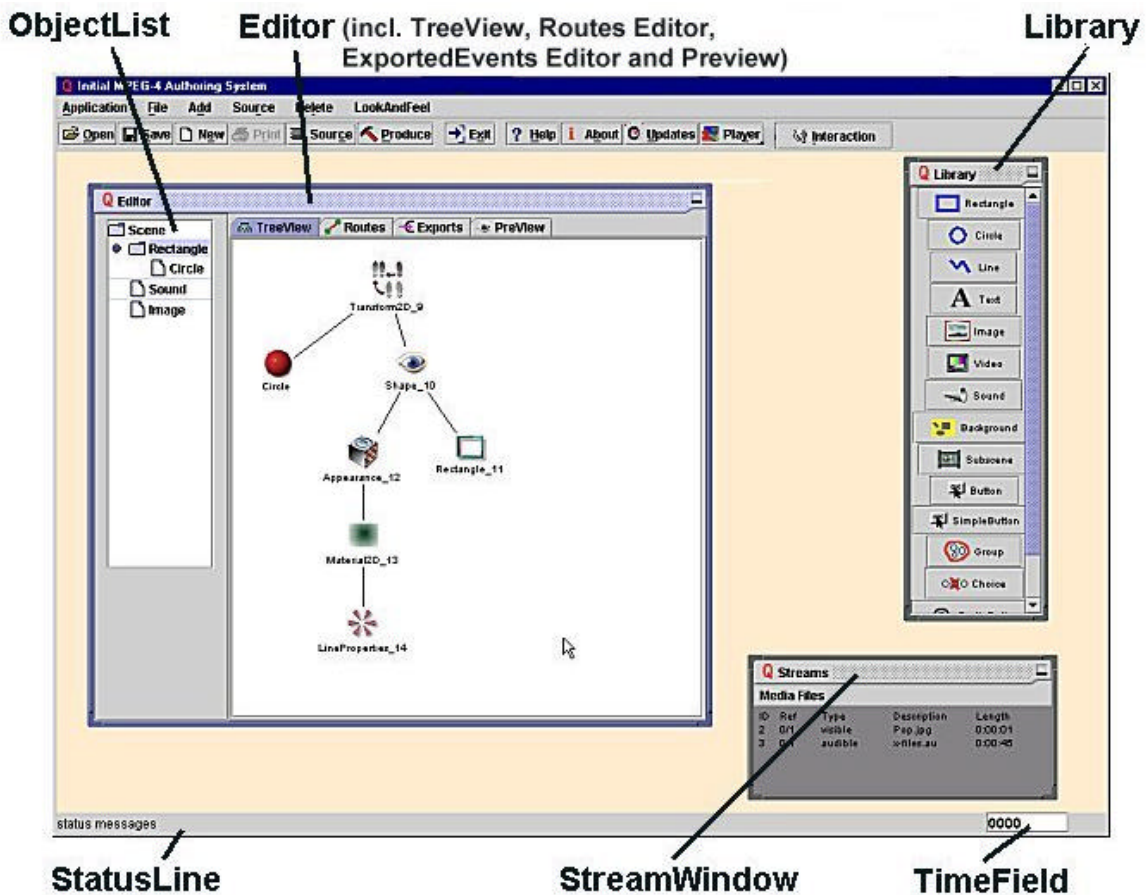


Figure 1: Screenshot of the authoring system's preliminary user interface

– The Preview

This window shows a snapshot of the created scene at the time set in the Time Field (see below). This gives the possibility of quick checks of (intermediate) authoring results.

- The Library

The Library contains pre-defined objects, e.g. graphical primitives or buttons. The properties of each object can be set in a corresponding dialogue window. Author-defined objects can be stored in the Library for later re-use.

- The Stream Window

Here all selected media streams are shown which the author may want to use during the authoring process. Setting of coding parameters is also done from within this window.

- The Status Line with the Time Field

The Status Line shows system messages. In the Time Field, a time can be set for which the scene being created can be viewed in the Preview window.

## 2.2 The Authoring Process

In this sub-section, the very principal steps of authoring with the presented MPEG-4 authoring system are described.

- (1) Starting an application. Either a previously completed application must be opened if an application shall be re-edited, or a new application must be opened as null content.
- (2) Introduction of objects. Some objects require the selection of media streams.

- (3) The editing.
  - a) The spatial and temporal composition of the scene has to be defined.
  - b) Properties of objects and nodes, routes, event dependencies, behaviours, have to be set. In this step, the elements described in Sub-Section 2.1 are extensively used.
  - c) Interaction mode for setting interactivity among objects and the time dependent insertion/deletion of objects has to be activated, as appropriate.
  - d) Preview of intermediate results.
- (4) Setting of coding parameters.
- (5) Production of an MPEG-4 application file.
- (6) Calling of the MPEG-4 Playing Station and viewing of the application file.
- (7) Saving the application for later re-use.

### 3. The Playing Station

The MPEG-4 PC project contributes with its playing station to the so-called “Reference Software“ of the Systems part of MPEG-4. The standard is divided into 6 parts: Systems, Visual, Audio, Conformance Testing, Reference Software, and DMIF. The Reference Software is a reference implementation of the textual specifications in the other parts. The Reference Software for the Systems part consists of 3 elements: the playing station provided by the MPEG-4 PC project for 2-dimensional content, a playing station for 3-dimensional content, and a common core that includes e.g. a demultiplexer and the synchronisation layer.

Figure 2 shows the architecture of an MPEG-4 terminal. The main components are the DMIF layer (see below), the demultiplexer in the Systems layer, the BIFS and media decoders, the executive, and the presenter which includes the audio and visual renderers.

Whenever an application stream arrives at the terminal an executive is started. All steps of demultiplexing, decoding etc. to presenting the application are under the control of the executive. The demultiplexer decomposes the incoming application stream into its elementary streams and sends each of them to the appropriate decoding component. The executive is also responsible for creating the correct number and types of decoders, along with setting up the data paths between the components. As soon as the BIFS decoder finishes decoding the BIFS stream, the scene graph is constructed. The presenter reads the scene graph and calls the visual and audio renderers to display the scene. The end-user can send events to the presenter in order to interact with the audio-visual objects or the scene composition.

The playing station supports JPEG, H.263, and MPEG-4 (including arbitrarily shaped objects) formats for video information, G.723 and AAC formats for audio information.

MPEG-4 also defines a network-transparent transport mechanism (Delivery Multimedia Integration Framework, DMIF) for MPEG-4 data streams. The idea behind it is that applications can always use the same interface, the so-called DMIF Application Interface DAI, no matter which delivery service is used. All services are transparent and the actual source of the MPEG-4 content doesn't need to be visible. The DAI cares for the correct adaptation to the particular network service. Within the MPEG-4 PC project a prototype of a DMIF server/client delivery mechanism for broadcast scenarios is developed. It provides reliable, real-time, point-to-multipoint connections with a defined Quality of Service. Because of the lack of bi-directional data flow for signalling in unidirectional broadcast scenarios the project implements carousel techniques in order to provide the end-user the feature of a late tuning-in. Also part of the MPEG-4 PC project is the development of a broadcast application server that prepares MPEG-4 content for broadcast transmission.

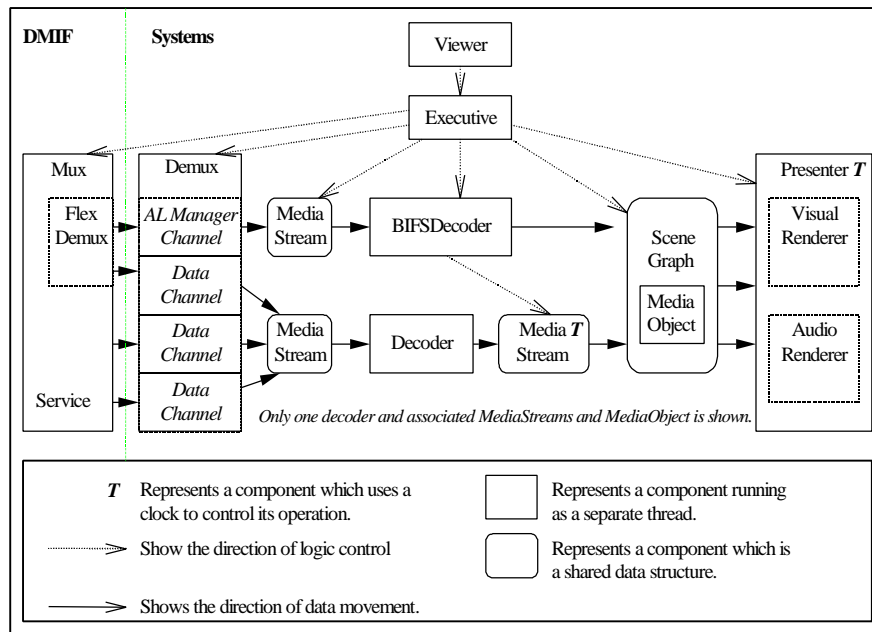


Figure 2: Architecture of an MPEG-4 terminal

## 4. Demonstrator

The aim of the MPEG-4 standard is to become THE future generic standard for all multimedia applications. For the design of prototype applications it is necessary to emphasise just the specific functionalities of this standard. One of its characteristics is to offer easy interactivity. We are presenting two examples here, that address applications of future-oriented, interactive, TV: the “Tour de France“ and the “Election Night“.

### 4.1 “Tour de France“ Application

This application consists of a main choice screen and four screen layouts with different levels of interactivity:

- (1) a regular MPEG-2-like TV screen with no interactivity,
- (2) a see-all screen for the race against the clock, where the screen is divided into quarter-sized parts for four channels and the viewer can zoom in on any,
- (3) a multiple-channels-in-one screen with a command bar and buttons, where e.g. by clicking on a group of racers in the command bar the view on that particular group can be selected and additional information such as geographical maps, altitude profiles, or information about a racer can be called,
- (4) an advanced screen with multiple channels where each can be chosen as the main channel while, in addition to layout #3, smaller versions of the other channels can be viewed in a bar on the right of the screen in order to be able to switch to the one which is most attractive for the end-user at a certain time.

Figure 3 gives screenshots of the choice screen and the layouts #2 through #4. The end-user can select among the different layouts and, within any layout except #1, among different views and can call additional information in a hypermedia fashion.



Figure 3: Screen layouts of the “Tour de France“ application; top left: the main choice screen, bottom left: the see-all screen, top right: the multiple-channels-in-one screen with an increased level of interactivity (see the command bar and buttons), bottom right: the advanced screen with additional channels in small windows

#### 4.2 “Election Night“ Application

In this application the user gets the possibility to select among election or poll results and a lot of additional information such as interviews of candidates.

For example, a map of France will be displayed with national results, and it will stay on the screen unless the viewer uses a pointer to select a particular town. In this case, a map of the region around that town is displayed with more detailed information. Maps of all regions of France are sent with the map of France, but only displayed on demand. The user will also be able to configure his/her terminal to automatically switch to the detailed view around his/her town upon reception of localised information (election results, weather forecast, ...). The user will also be able to get a split screen with the map of France presented in one part, and the map of a specific region in the other.

The essential difference to the “Tour de France“ application is in the complexity of the graphics content per voting site (face of candidates, image of site, bar charts) as well as the real-time updating of the figures.



## 5. Conclusions

Goals and objectives of the European project “MPEG-4 PC - MPEG-4 System Implementation and Tools for Personal Computers“ have been presented. The software under development will allow the authoring and playing of content that conforms to the specification of the standard ISO/IEC:14496, called MPEG-4. It is expected, e.g. by market analysts, media and technology experts, that MPEG-4 will have tremendous influence on the world of media, especially for content creators, art directors, and content providers and thus also for the consumers. The amount of money companies invest only in the specification of the standard is estimated at roughly 2 to 4 billion EURO, the highest amount ever paid for the specification of a multimedia standard. How much money companies have already invested in the development of MPEG-4 products and services is difficult to assess. Precise figures on that have not yet been published by market analysts; however a multiple of this sum, likely more than hundred times as much, can be expected. MPEG-4 covers many areas of multimedia, from hardware through software to services. Only for multimedia hardware and software, the revenue for the year 2000 is estimated by market analysts at more than 200 billion EURO with substantial increase in the following years. MPEG-4 products will occupy year by year more and more parts in the multimedia areas. This shows the importance of the MPEG-4 standard also from a market point of view.

Application areas of MPEG-4 are e.g. (interactive) TV, infotainment, edutainment, computer-based training, advertisement, games, movies, streaming video over Internet/Intranet, collaborative scene visualisation, virtual meetings, real-time communications, surveillance, content-based storage and retrieval, digital AM broadcasting. For almost all of the mentioned applications, content needs to be created. MPEG-4 provides the means for the creation of a novel kind of content with which the end-user can interact. MPEG-4 authoring and playing systems will therefore become of great importance in the near future.

The development performed within the MPEG-4 PC project is the first combined implementation of an MPEG-4 authoring/playing system, as known to the MPEG community. With this system it is possible to author and play the novel kind of (interactive) content on a regular PC; no specific hardware is needed. Thus, it is relatively easy to introduce it in the market, what shows the relevance of the project work. The project also contributes successfully to different parts of the standard, the Systems, the DMIF, the Reference Software, and the Conformance Testing specifications, as well as to the promotion of the new standard. The consortium has close contact and cooperation with other projects, companies, and institutions that either contribute with other elements to the standard, thus taking advantage e.g. by mutual use of software, or intend to develop MPEG-4 products or services.

Partial results of the project were demonstrated successfully to the press and the public during the first MPEG-4 (IP multicast) satellite transmission from the West Coast of the US to Atlantic City, NJ, at the East Coast in October 1998. Since then, the development has progressed, and a new milestone including full, i.e. unidirectional, broadcast support is expected just on time before the EMMSEC99 Conference.

Additional information about the project and MPEG-4 can be found in [3] and [4], resp..

## References

- [1] ISO/IEC, International Standard IS 14496, 1999.
- [2] Rob Koenen, MPEG-4 - Multimedia for our Time, *IEEE Spectrum*, **36** (1999) 26-33.
- [3] <http://www.q-team.de/mpeg4/mpeg4pc.htm>
- [4] <http://cselt.it/mpeg>