3D Television Production Based on MPEG-4 Principles

Marco Rittermann
Institute of Media Technology
Technische Universität Ilmenau
Postfach 10 05 65
Deutschland, 98684 Ilmenau
marco.rittermann@tu-ilmenau.de

Michael Schuldt
Institute of Media Technology
Technische Universität Ilmenau
Postfach 10 05 65
Deutschland, 98684 Ilmenau
michael.schuldt@tu-ilmenau.de

ABSTRACT

The processing and usage of 3D graphics and video have been becoming very important and popular in many areas. At present, some of these areas seem to grow together. 3D television may become one result of this development. In this paper the ideas of the project "Interactive Audiovisual Application Systems IAVAS" will be presented and it will be explained how the project's results can be one way getting content for 3D television. In the IAVAS project object based processing of audiovisual content is widely used. For this, the coding methods defined in MPEG-4 are explored and evaluated.

Keywords

3D Television, MPEG-4, Interactive Applications

1 INTRODUCTION

The success of 3D television will depend on both, the availability of 3D displays and the availability and attraction of content as well.

The development of 3D displays has been going on for several years. There are some approaches which are promising, e. g. by the Heinrich-Hertz-Institute, Berlin or 4D-Vision, Jena, Germany. Solutions for cinemas have been used for some years. The production of content ("3D films") is very specialized and expensive. Most productions are meant for 3D cinemas which are customary in all big cities. Naturally, 3D films will be only one kind of content for 3D television.

Last but not least the attraction of the content of 3D television will decide its success. It will not be sufficient enough showing conventional films in 3D. Therefore, novel benefits of 3D television are necessary. Interactivity will be one of these benefits. One example is the possibility given to the user choosing an own viewpoint.

Interactivity at a high level needs much information about the content. An object related description allows the access to single objects. A complete 3D description of the content allows a free viewpoint selection by the user. These principles are some of the aims targeted by MPEG-4. That's why, MPEG-4 and its targets should be checked. [Dru01], [Rit01]

2 MPEG-4

MPEG-4 passed in 1998 as a standard for the coding of audiovisual objects. The "Moving Pictures Expert Group" is a working group of the International Organization for Standardization (ISO). In the early 1990s, MPEG-1 and MPEG-2 became two very successful standards for coding video with associated audio. MPEG-4 is not only an improvement of MPEG-2. Novel approaches for coding aural and visual content are defined. The way is to code it as aural, visual, or audiovisual objects in a 2D or 3D scene. The coding of these objects relates to their origin. This means that e. g. a chair can either be a real object pictured by a camera or a synthetic one modeled on a graphics computer. Because of these many different processes defined in MPEG-4, it has become a very extensive norm.

WSCG POSTERS proceedings
WSCG'2003, February 3-7, 2003, Plzen, Czech Republic.

Due to the number of different coding processes and approaches the MPEG-4 standard is often called as a tool box. A successful rule of MPEG-4 is defining only the coded signal and how to decode it but not the way to encode it. Thus the standard can be up-to-date for a long time because the encoder technology can be improved while the standard remains altered. The description of the spatial and temporal relationships of these objects is laid down in Part 1 "Systems" of MPEG-4. [Dru02b], [ISO02], [Rit01]

3 WHY USING MPEG-4 FOR 3D TELEVISION?

3.1 The Benefits of MPEG-4 Application Systems

As described in chapter 2, MPEG-4 offers many methods for describing audiovisual content very close to its origin. Obviously, this requires much expenditure. So the question is: Is such a specialized processing of the content necessary? Yes, such a processing is very useful: The generation of the content has to be done only once for the utilization on different platforms, e. g. WWW applications, DVB, Cinema, and self-evident 3D television. Therefore, a complete digital media chain is necessary for such application systems. The main idea is the content adaptive and object based description beginning at the creation and ending at the presentation of the content in a 3D environment. [Kuh01], [Jag02]

3.2 3D Support of MPEG-4

MPEG-4 is already supporting the third dimension for many media objects. Basically, the scene description is similar to conventional syntax for 3D computer graphics. It was derived from VRML97 standard. Thus, synthetic 3D scenes and elements can be coded.

Concerning natural video, the question of the third dimension has to be answered regarding to the concrete case of usage: stereoscopic viewing, or free viewpoint viewing, or other kinds of viewing. For some types of objects, the coding is already standardized, for example shaped video. Shaped video can be used in some cases in a 3D environment if the third dimension of the object itself will not be perceptible much. A general solution for this problem would be very difficult to find.

The next chapter introduces the IAVAS project where such ideas will be realized.

4 THE IAVAS SCENARIO

The IAVAS ("Interactive Audiovisual Application Systems") project currently under development at the Institute of Media Technology at the Technische Universität Ilmenau (university of technology) aims at integrating and dominating the whole chain of production for interactive application systems, from recording and/or modeling the audiovisual content, storing and transmitting it, to the reproduction at the user's side. For the reproduction at the consumer's side the used display device plays an important role. Besides the normal TV screen, the using of other devices should be possible. The immersion of the user into a virtual scenery would be increased by presenting the virtual scenery using a 3D device. But this demands that during the production of audiovisual content information about the object attributes (e. g. geometry, dimension, texture) does not disappear. In case of synthetic visual objects (e. g. meshes) the creation process is a rather simple task. Synthetic visual objects can be produced by using modeling software, e. g. 3D computer animation, or CAD software.

During the capturing of natural video objects for 3D television the achieving of all object attributes is more difficult. Normally, natural video objects are captured using standard broadcast cameras. Due to the projection of 3D objects onto a 2D plane during the capturing process the third dimension will be lost. Preserving the third dimension of these objects is very difficult. At one side the third dimension is needed for the presentation of these objects at 3D devices. At the other side the scene realism can be increased by the support of 3D natural video. Normal shaped video does not allow a free choice of the viewpoint. There are many reasons for supporting 3D natural video in virtual environments. Possible solutions and requirements for 3D natural video will be discussed in the next chapter.

5 NATURAL 3D VIDEO

5.1 The World of 3D Video

There are many meanings of the term "3D video". This variety of different meanings is a result of many applications, many ideas, and many approaches achieving the third dimension of video.

5.2 3D View

First, there is a group of solutions which have one aim: giving the viewer the impression, that she or he is watching on 3D scenes instead of 2D moving pictures. Normally, the solution is a stereoscopic presentation needing a technical aid for the viewer to watch it. There are some methods using glasses for the viewer, e. g. using anaglyphs, or perpendicular polarizations. All kinds of glasses separate the view onto the left and onto the right eye. (see Fig. 1)

Corresponding to the kind of glasses, the film has to be shown with different pictures for the left and for the right eye, e. g. with perpendicular polarizations. These techniques can be used in cinemas or at

personal displays. Naturally, the wearing of glasses is uncomfortable. Other techniques are limited to only one viewer. Another kind of displays shows different pictures in different angles. Due to this independence of the viewpoint a three-dimensional impression can be achieved for some users.

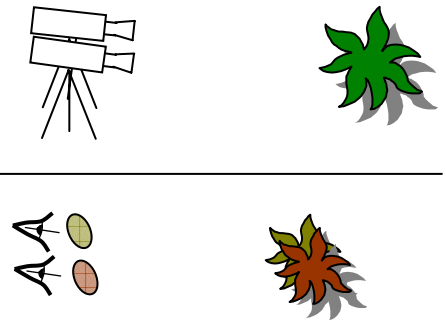

Figure 1. Stereoscopic recording and viewing

All this techniques use conventional films with the added feature 3D. The attraction of 3D films shown in 3D cinemas is the third dimension itself. The cinema spectators are interested in watching a 3D film. The story is quite unimportant. That's why, 3D cinemas will not play a more important role.

5.3 Video Using an Omni Directional View

The term 3D video is used in another context, too. Assuming a fixed point of view, this technique allows views in many directions. (see Fig. 2)

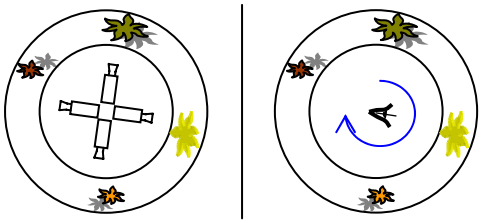

Figure 2. Omni directional recording and viewing

Normally, these are all angles in one given plane or in one hemisphere. Such panoramic views are used interactively in most cases. The user can turn around and look at the surroundings.

5.4 Video Using Multiple Views

Another technique provides different views at one scene, e. g. a soccer game. First realizations do not offer a free viewpoint choice to the user. Instead of this, e. g. a certain camera flight is prepared at the TV station to show details of a goal. For this purpose, a large number of cameras is required. In future applications may all the camera signals be transmitted and the TV viewer can choose an own viewpoint. [Ran98]

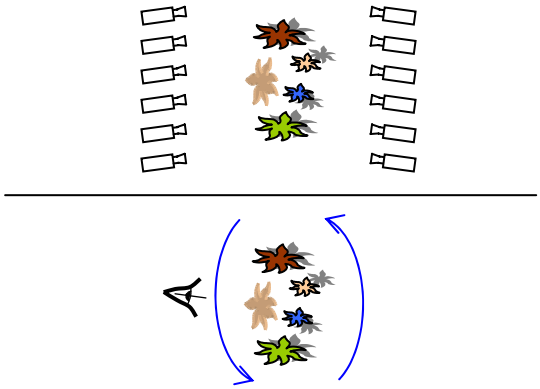

Figure 3. Recording multiple view and using of multiple view

5.5 3D Video Objects and Scenes

What does the term "3D video object" mean? A simple explanation could be: a single object which depends on time and which can be viewed from any viewpoint. The detailed requirements to a 3D video object can differ. This depends on the application it is needed for. One typical usage would be in a 3D scene (e. g. consisting also of many static 3D objects) with some 3D video objects where a free viewpoint for the user is allowed.

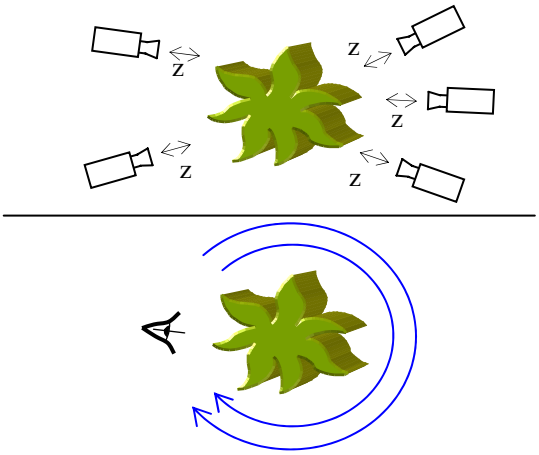

Figure 4. 3D video object recording and free viewpoint viewing

Obviously, a complete free choice will not be realizable. Limitations will ever exist. Regarding 3D objects, a complete 3D geometry description with its high defined texture would be one entire description. In order to get an approximation, disparity and depth information can be used. [Ohm00], [Gra01], [Ohm99]

Another approach describing 3D video objects and scenes is to code the emitted light of every point in every direction. This may similar to the principles of light field mapping. [Che02]

6 DEMONSTRATIONS

In preparation of the IAVAS project some experiments were done in order to demonstrate the look-

and-feel of the main ideas explained in MPEG-4. One outcome is a demonstrator based on a VRML scene. Some natural video objects are placed in a simple 3D environment. These are represented by sequences of still images. The user can navigate through the scene and can interact with any object. These experiences of a MPEG-4 feeling gave a way for the further steps.

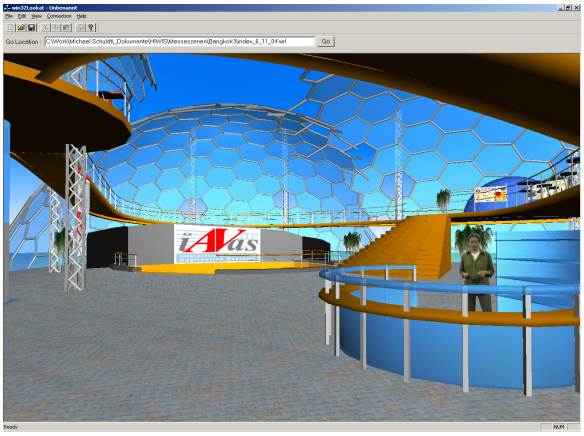

Figure 5. Application with MPEG-4 encoded shaped video objects

The integration of one or several shaped video objects into a 3D scene was realized by an early application. These objects are coded as MPEG-4 elementary streams. This application is the starting point for further examinations and developments. One goal e. g. is the realization of 3D video. [Schu02]

7 SUMMARY

3D television will need enough content and attraction in order to become successful. The growing together of different services and media can give these prerequisites for 3D television. Novel forms of uses the audiovisual content, like interactivity, are a chance for 3D television. The coding of objects in 3D environments like it is defined in MPEG-4 allows such features. These principles will only work in complete application systems as focused in the IAVAS project. The results of this project will be necessary to define the requirements and approaches e. g. for 3D video objects. 3D television would be one outcome of this concept.

8 ACKNOWLEDGMENTS

Our thanks to Mr. Prof. Dr.-Ing. Karlheinz Brandenburg, to the Technische Universität Ilmenau (university of technology), and to the Thuringian Ministry of Science, Research, and Arts.

9 REFERENCES

[Che02] Chen, W.-C.; Bouguet, J.-Y.; Chu, M.; Grzeszczuk, R., Light Field Mapping: Efficient Representation and Hardware Rendering of Surface Light Fields, To appear in the ACM Siggraph conference proceedings, 2002.

[Dru01] Drumm, H.; Rittermann, M., Authoring und Visualisierung von animierten Szenen auf Basis von realer und synthetischer Objekte, 46th International Scientific Colloquium, Technische Universität Ilmenau/Germany, 24th-27th September 2001.

[Dru02a] Drumm, H.; Kühhirt, U.; Rittermann, M., Anwendungssysteme für MPEG-4; 20. Jahrestagung der FKTG. Zürich/Switzerland, 10th-13th June 2002.

[Dru02b] Drumm, H.; Kühhirt, U.; Reiter, U.; Rittermann, M., Application Systems for MPEG-4; IEEE International Symposium on Consumer Electronics, Erfurt/Germany, 23rd-26th September 2002.

[Gra01] Grau, O.; Price, M.; Graham, A. T., Use of 3-D Techniques for Virtual Production, BBC R&D, Surrey/UK, 2001.

[ISO02] International Standardization Organization, ISO 14496 Information Technology – Generic, Coding of Audio-Visual Objects, 1998 pp.

[Jag02] Jäger, R.; Knorr, M.; Kühhirt, U.; Rittermann, M.; Weigel, C., Interactive Data Services in a Digital Video Broadcasting Environment; IEEE International Symposium on Consumer Electronics, Erfurt/Germany, 23rd-26th September 2002.

[Kuh01] Kühhirt, U. Processing of audiovisual objects and scenes using MPEG-4, Ilmenau Summer School 2001, Bangkok/Thailand, 17th – 21st December 2001

[Ohm99] Ohm, J.-R., Encoding and Reconstruction of Multiview Video Objects, IEEE Signal Processing Magazine, May 1999.

[Ohm00] Ohm, J.-R.; Müller, M., Incomplete 3D – Multiview Representation of Video Objects, Berlin/Germany, 2000

[Ran98] Rander, P., A Multi-Camara Method for 3D Digitization of Dynamic, Real-World Events, PhD Thesis, Carnegy Mellon University, Pittsburgh/USA, 1998

[Rit01] Rittermann, M., Overview of MPEG-4 – A format for multimedia representation, Ilmenau Summer School 2001, Bangkok/Thailand, 17th– 21st December 2001

[Schu02] Schuldt, M., Implementierung von shaped Video in 3D-Szenen, Diploma Thesis, Technische Universität Ilmenau, Ilmenau/Germany, 2002