Quarterly newsletter of the MUSCADE consortium

www.muscade.eu

MUSCADE Focus

Volume 1, Issue 2 APR-JUN 2010

Special points of interest:

- KUK records German Football team in 3D
- The position statement is on 3D video rendering challenges and techniques.
- MUSCDE is represented at several events including Dimension3 2010 and NAB 2010

Inside this issue:

3D Video Rendering Challenges and Techniques

Position Statement submitted by Laurent Blondé, Didier Doyen, and Thierry Borel of Technicolor Research & Innovation, Rennes, France

The advent of 3D to the home raises a number of challenges beyond the basic display of an available 3D video content. A number of them are technical challenges, as for example the adaptation of various 3D video content types to different display technologies. Other aspects relate to the Quality of Experience, involving observation conditions and observers.

Indeed, a correct 3D perception for everyone cannot be taken for granted for any of the current 3D display technologies. Fatigue, eye strain, or variability of perception are phenomena of importance and they cannot be neglected when developing a 3D chain.

"3D Rendering" is used in this position statement as short for "*Adapting decoded 3D content to each target 3D display technology — taking into account Video Quality and Human Factors*". The goal is to ensure the best possible 3D experience for a given observer on a targeted display.

3D Rendering can be implemented as a Set Top, Gateway or software processing block. It is getting its input signals from 3D decoder(s) and providing output signals to various 3D display technologies (Fig. 1).

Various processing blocks need to be developed and assembled with parameters depending on:

- Input video format
- Display characteristics and viewing conditions
- Observers' characteristics and commands.

In 3D rendering, however, challenges appear even more sophisticated, with on one side **technical challenges** and on the other side **perception challenges**:

- **Technical challenges** are linked to signal input, processing and adaptation to displays.
- **Perception challenges** relate to the viewers and to observation conditions.

The following sections will describe more in details these 3D challenges, and will give examples of new components considered to improve the 3D experience.

To evaluate performances, a new range of 3D Quality of Experience tools has to be developed supplementing the 2D quality tools available (see [1] for instance).

Continued on page 2...

3D Video Rendering Challenges and Techniques continued...

Figure 1. Rendering module: an element of the 3D video chain

TECHNICAL CHALLENGES

A number of modules compose together the 3D rendering chain. Fig. 2 represents a tentative processing structure combining these modules and the Perception management module described in the next section.

Input data management: in the current state, no standard is established for the representation of transmitted 3D-TV stereoscopic and multi-view content. It is the purpose of SMPTE and DVB to propose stereoscopic formats in the near future and in a second step multi-view formats.

Input data processing needs to be provisioned to adapt the video format and split video and depth/disparity information. This module may also include the correction of imperfect images, including for example random compression blocks. This type of defects creates severe visual annoyances in the form of 3D structures.

Signal enrichment module creates complementary signals of use in 3D rendering, such as exploitation of metadata sent with the content.

"3D Insertion" term is used here as an extension of the "Overlay" term of 2D video. In 2D video, a number of visual elements commonly complement the programs. Examples of these are subtitles, menus, score boxes or logos. Some, as subtitles or menus are dependent on the consumer choices – and thus need to be inserted locally – while score boxes or logos can be inserted at the post-production stage. In all cases, the issue is the correct insertion of these elements in the 3D content, to avoid visual annoyance when inserted elements' depth is too different from the scene depth and structure around it. Situation is even worse when masking by foreground scene elements happens.

Taking subtitles insertion as an example, the easiest option is to let subtitles be in the screen plane, which generates discomfort due to occlusion by the content. Placing the insertion in front of any object in a sequence may create in turn highly crossed disparities difficult to fuse. Using depth/disparity information, intermediate strategies can be developed such as depth dependent subtitle placement or depth adjustment of the whole scene.

The goal of the **output formatting** module is to convert this signal to the actual display connected, depending on its technology. Output

formatting processing can include a number of signal processing elementary functions as: 2D filtering, sub-sampling/oversampling, temporal processing, color processing or crosstalk pre-correction. Another major processing is the view interpolation to address correctly a multi-view display with the right number of views. Light-field displays will also require a specific rendering.

"It is the purpose of SMPTE and DVB to propose stereoscopic formats in the near future and in a second step multi-view formats"

PERCEPTION CHALLENGES

Analyzing viewers' reports, some being enthusiastic and some others alarming, a correct 3D perception for everyone cannot be taken for granted for any of the current 3D display technologies. Fatigue, eye strain, or variability of perception are effects of importance and they cannot be neglected when developing a 3D chain.

Observation conditions are the first aspects to address. Viewing distances, field of view, and number of views are drastically dif-

Continued on page 3...

3D Video Rendering Challenges and Techniques continued...

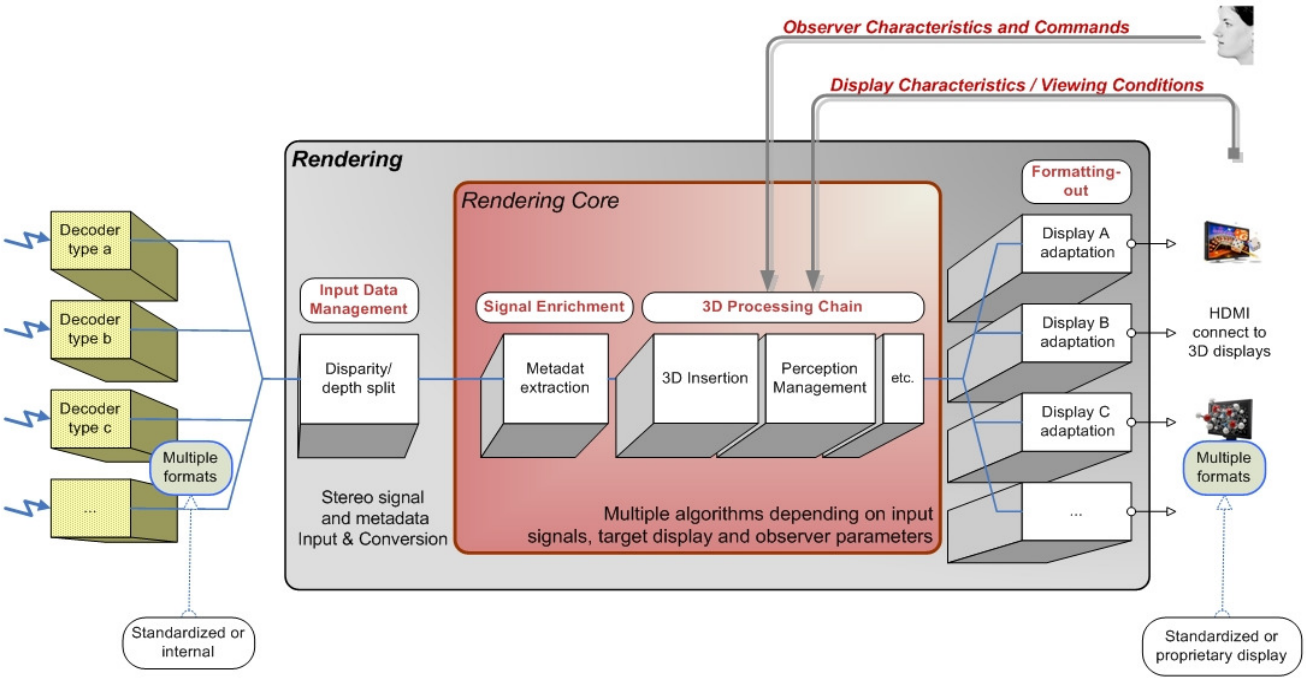

Figure 2. Rendering module: processing structure

ferent in a consumer environment from the movie theatre conditions. Accordingly, the content has to be adapted.

A widely discussed subject regarding comfort of 3D viewing is the *vergence-accommodation* conflict (*accommodation* is here the change of optical power to maintain focus on an object while *vergence* is the simultaneous movement of both eyes to adjust the gaze directions on a same object for binocular vision). Vergence vs. accommodation decoupling happens when displayed 3D objects are not portrayed in, or close to, the screen plane. The decoupling is mentioned as the main cause of visual discomfort and fatigue. The vergence-accommodation conflict is a well-identified problem of 3D vision on displays. It has been already described in different articles such as in [3].

On the observers' side, each of them has a different physiology, a different vision history. He or she may have preferences regarding the way stereoscopic or multi-views images are presented, in correspondence with his or her own 3D perception.

Physiological and perceptive factors contributing to 3D perception are numerous. A basic one, often cited, is the inter-ocular distance defining the stereoscopic epipolar base. It is commonly used as a parameter in algorithms modeling the geometry of observation to transform the stereo pairs for varying viewing distances or conditions. However, 3D space perception goes far beyond exploiting just geometry. A quite successful model is the probabilistic combination of a variety of 3D cues as exposed in [4]. Complementing binocular disparity, binocular and monocular 3D cues include for instance accommodation, convergence, parallax, perspective, occlusion between object, relative and familiar size, etc.

"Fatigue, eye strain, or variability of perception are effects of importance and they cannot be neglected when developing a 3D chain"

While the content creator may play with these cues and their interpretation, collision of monocular or binocular cues will rapidly

Continued on page 4...

3D Video Rendering Challenges and Techniques continued...

create inconsistencies and annoyances. It is thus of interest to understand how strong these cues are relative to each other and compared to stereoscopic disparity. It is also of interest to evaluate the acceptance/annoyance in perception in case of conflicts.

All these effects, and more, create conflicts for the visual system and are source of discomfort. Apparently annoyance is at various levels for various observers, and effects perceived as being strong by some viewers are almost ignored by others.

Understanding, ranking the effects, and analyzing the cause of discomfort constitute the first step. Then video processing needs to be developed to better match the condition of acceptance for each potential viewer.

Observer variability needs to be studied and processing needs to be adapted to each type of vision/perception. Adaptation can be automated, and needs input parameters. A feedback/control from the viewer(s) towards 3D Rendering is thus very desirable, not feasible in a movie theatre, but more adapted to a small group or an individual viewer watching 3D-TV.

An important parameter not yet fully addressed is the specific source of discomfort we could get with multi-view displays. Their specific way of rendering 3D content has to be analyzed in terms of Quality of Experience and dedicated processing should be applied to improve it.

"There is a need to develop the adequate processing modules, and simultaneously to develop Quality of Experience procedures and tests for appropriate 3D quality assessments"

CONCLUSION

For a wide success, 3D-TV requires a large acceptance in the audience. Technical and perception challenges need to be addressed, taking into account the home visualization equipment as well as observation conditions, and giving options to the viewers to adapt 3D content to satisfy their perception. Without such module in the 3D chain, the risk is high of unsatisfied customers complaining from poor perceived quality, suffering headaches or eye pain, and potentially rejecting 3D-TV. The objective of the industry is thus to propose automatic or interactive features for the visualization system to perform at best, and for the consumer to enjoy 3D comfortably.

Facing this goal, there is a need to develop the adequate processing modules, and simultaneously to develop Quality of Experience procedures and tests for appropriate 3D quality assessments. We expect the relevant technical bodies to take this over, encompassing the challenges exposed here, and more.

Bibliography:

[1] K. Brunnström, D. Hands, F. Speranza, and A. Webster, "VQEG Validation and ITU Standardization of Objective Perceptual Video Quality Metrics", IEEE Signal Processing Magazine, May 2009, pp. 96-99.

[2] Q. Yang, L. Wang, R. Yang, H. Stewénius, and D. Nistér, "Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling" , IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31 No. 3, March 2009, pp. 492-504.

[3] D. M. Hoffman, A. R. Girshick, K. Akeley, and M.S. Banks. "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue", Journal of Vision, 8 (3), (2008), pp. 1-30.

[4] D. Kersten, P. Mamassian, and A. L. Yuille, "Object perception as Bayesian inference", Annual Review of Psychology 55, (2004) pp. 271-304.

Position Paper Submitted by: Technicolor Research & Innovation, Rennes, France

MUSCADE T1.2 Meeting took place at UNIS

The one-day MUSCADE T1.2 meeting took place at UNIS in Guildford, UK on 30 April 2010. Representatives from 5 MUSCADE institutions (UNIS, EBU, HHI, EADS Astrium, and KUK) attended this meeting, which included technical discussions on partners' respective contributions to deliverable D1.2.1 as well as on planning the specifics of the MUSCADE 3D content test shooting day to take place in July 2010. The attendees also visited the UNIS I-Lab facilities, where demonstrations particularly on the audio related research topics were given by the UNIS researchers. Furthermore, this meeting also comprised a session for presenting the latest QoE research activities performed under T4.1.

Submitted by: UNIS

MUSCADE features at NAB 2010

"Content Creation in 2020 "

NAB 2010, Digital Cinema Summit, April 10-15, 2010, Las Vegas, USA. 3D: Cinema and Home:

Link: http://expo.nabshow.com/annual10/public/Content.aspx?ID=282&sortMenu=105008&MainMenuID=272

Dr Aljoscha Smolic of Disney Research Zurich gave a talk entitled "Content Creation in 2020" at the Digital Cinema Summit of NAB. He presented an overview of the state-of-the-art in the area and outlined challenges over the next decade. He also introduced MUSCADE as a major European project working on the solution of those challenges.

The session was organized by Hans Hoffman of EBU, and Tibor Balogh of Holografika was also among the speakers. The Digital Cinema Summit is a highly important event in the broadcast, cinema and media industry in general. The talks gave the MUSCADE project and the developed technology significant visibility among highly important players.

"The talks gave the MUSCADE project and the developed technology significant visibility among highly important players"

Submitted by: DRZ

MUSCADE T1.4 Meeting took place in Budapest

The one-day MUSCADE T1.4 meeting took place at Holografika in Budapest, Hungary on 25 May 2010. Representatives from 5 MUSCADE institutions (Holografika, Fraunhofer HHI, Technicolor, University of Surrey and Astrium) attended this meeting, which included technical discussions about the architecture of the rendering system, refinement of requirements for the two phases of the project, and interfacing issues between the decoders and renderers. The partners also discussed several alternatives regarding the rendering framework, synchronization of audio and video, and the connection between video and interactive content. The discussions served as a basis for the deliverable D1.4.1. During the day, a demonstration of a large-scale HoloVizio display and interfacing capabilities was also performed.

Submitted by: Holografika

Astrium Represents MUSCADE at the Toulouse Space Show in France

Figure 3. EADS Astrium Booth at the Toulouse Space Show

"The system architecture and the key transmission technologies of MUSCADE were presented at the Astrium Booth"

Link: http://www.toulousespaceshow.eu/

MUSCADE was represented by Astrium at the Toulouse Space Show, which took place in Toulouse, France on 8-11 June 2010. As the event featured a session on 3D-TV, the system architecture and the key transmission technologies of MUSCADE were presented at the Astrium Booth (Fig. 3).

The event was inaugurated by Valérie Pécresse, the French Minister for Higher Education and Research. Jointly organized by the French space agency, CNES, and the regional authorities, the Toulouse Space Show played host to over 1,000 attendees from 45 countries and more than 80 exhibitors.

Submitted by: EADS Astrium

SES ASTRA and SAMSUNG to promote 3D

Link: http://www.ses-astra.com/business/en/news-events/news-latest/index.php?pressRelease=/pressReleases/pressReleaseList/10-04-28/index.php

SES ASTRA, an SES company (Euronext Paris and Luxembourg Stock Exchange: SESG) and Samsung Electronics, announced that they will jointly promote 3D television to the market. Under the terms of the agreement, Samsung will support SES ASTRA with 3D television content for its new 3D demo channel to be launched on 4 May 2010. The demo channel will be broadcast free-to-air via ASTRA's orbital position 23.5 degrees East, and is intended to help retailers promote 3D television to end-consumers at the point of sale.

Submitted by:
SES ASTRA TECHCOM

MUSCADE addresses Dimension3 2010

"Approaches towards Multi-View-Content "

Link: http://www.dimension3-expo.com/

MUSCADE researcher Ralf Tanger of Fraunhofer HHI attended this year's Dimension 3: International S-3D and New Images Forum. During this event, he made a presentation entitled "Approaches towards Multi-View-Content". The presentation provided an overview of the HHI stereo to multi-view conversion technology which will be used and improved within MUSCADE. In addition, the MUSCADE 4-camera setup that will be used to capture the 3D audiovisual project content was explained and proposed as a next step for very high quality multi-view content creation. This presentation provided a good opportunity to promote the project directions and capture technologies, which have been considered in the consortium, among a wider audience.

Submitted by: HHI

The German Football Team in 3D

KUK Film produced a short documentary about the training of the German football team for the 2010 FIFA World Cup.

SONY Germany contracted KUK Film to film at the team's training camp in Eppan, Italy. Two mirror rigs equipped with SONY HDC P1 and SONY HDC 1550 cameras were used for the 3D shooting. The recording was made with SONY HD CAM SRW1 recorders (Fig. 4).

Deutscher Fussball Bund (DFB) gave exclusive permission to KUK Film to place one 3D camera very close to the goal to get the best possible 3D effect. KUK Film installed their own 3D editing and compositing equipment on the site to produce a short trailer of the training match against the local FC Suedtirol within a three day-and-night editing session.

Submitted by: KUK

"DFB gave exclusive permission to KUK Film to place one 3D camera very close to the goal to get the best possible 3D effect"

Figure 4. KUK Film's 3D recording of the German football team

Publications Corner

This section features the academic publications that were done under the MUSCADE project during this quarter. The abstract of each publication is also given for each publication.

Predicting sensation of depth in 3D video

Authors: S.L.P. Yasakethu, D.V.S.X. De Silva, W.A.C. Fernando, and A. Kondoz

Affiliation: UNIS

Publication: IET Electronics Letters, Vol. 46, No. 12, pp. 837–839, June 2010

Link: http://www.theiet.org/publishing/journals/letters/el/index.cfm

Abstract:

A new quality assessment model is presented to predict sensation of depth in 3D video in the format of colour (monocular) video augmented by the grey scale depth map. The proposed technique is capable of evaluating both monoscopic and stereoscopic contributions towards depth perception. Results show that sensation of depth can be effectively modelled with the proposed model by combining visually important features to the brain.

3D video communication scheme for error prone environments based on motion vector sharing

Authors: D.V.S.X. De Silva, W.A.C. Fernando, and S.T. Worrall

Affiliation: UNIS

Publication: In Proceedings of 3DTV Conference, Tampere, Finland, 7-9 June 2010

Link: http://sp.cs.tut.fi/3dtv-con2010/

Abstract:

The emergence of three dimensional (3D) video applications, based on Depth Image Based Rendering (DIBR) has brought about new dimensions to the video transmission problem, due to the need to transmit additional depth information to the receiver. Until the transmission problem of 3D video is adequately addressed, consumer applications based on 3D video will not gain much popularity. Exploiting the unique correlations that exist between the color and their corresponding depth images, will lead to more error resilient video encoding schemes for 3D video. In this paper we present an error resilient 3D video communication scheme that exploits the correlation of motion vectors in color and depth video streams. The presented method achieves up to 0.8 dB gain for color sequences and up to 0.7 dB gain for depth sequences over error prone communication channels.

Continued on page 9...

Publications Corner continued..

Perceptual video quality metric for 3D video quality assessment

Authors: P. Joveluro, H. Malekmohamadi, W.A.C Fernando, and A.M. Kondoz

Affiliation: UNIS

Publication: In Proceedings of 3DTV Conference, Tampere, Finland, 7-9 June 2010

Link: http://sp.cs.tut.fi/3dtv-con2010/

Abstract:

One method of evaluating the quality of stereoscopic video is the use of conventional two dimensional (2D) objective metrics. Metrics with good representation of the Human Visual System (HVS) will present more accurate evaluation. In this paper we propose a perceptual based objective metric for 2D videos for 3D video quality evaluation. The proposed Perceptual Quality Metric (PQM) shows better results for 3D video quality evaluation and outperforms the Video Quality Metric (VQM); as it is sensitive to slight changes in image degradation and error quantification starts at pixel level right up to the sequence level. Verifications are done through series of subjective tests to show the level of correlation of PQM and user scores.

Project Management Office

EADS Astrium
Telecom Systems Department
31 rue des cosmonautes
31402 Toulouse
France
Phone: +33 562 196 364
Fax: +33 662 199 494
E-mail: greet.verelst@astrium.eads.net
Web: **www.muscade.eu**

Newsletter Editor

University of Surrey

MUSCADE:
MUltimedia SCAlable 3D for Europe