3D-TV R&D Activities in Europe

Oliver Grau, *Member, IEEE*, Thierry Borel, Peter Kauff, Aljoscha Smolic, and Ralf Tanger, *Member, IEEE*

***Abstract*—3D-TV is a topic that has been studied for many years in Europe. Through the research frameworks of the European Commission in particular, a number of long-term issues have been addressed to overcome limitations of the traditional two-view stereoscopy. This article gives a brief overview of the goals and achievements of some completed European projects starting in the 1990s. It then reviews the topics related to 3D-TV in recent European research. Finally an overview with a selection of recent projects is presented.**

***Index Terms*—Digital video broadcasting, multimedia systems, stereo vision, TV broadcasting.**

I. Introduction

EUROPE has a long history in three-dimensional television (3D-TV), starting from the first demonstration of stereoscopic television by Baird in 1928. Although the normal, two-dimensional television has been broadcast for decades, again pioneered in Europe, 3D-TV did not become a regular service in the 20th century. One of the main reasons was immature technology, which then led to a number of research activities to overcome these inhibiting factors. The contributions from Europe in this global effort are the subject of this paper.

The prime concept of 3D-TV is to add a stereoscopic sensation to the television experience. The technological way to implement this is to provide different views to each eye of the viewer. This can be achieved in many ways. The simplest concept is still based on principles of stereo photography developed in the 19th century, in which two views of the scene with two horizontally offset cameras are captured and then presented individually to each eye of the viewer. This two-view stereo forms the basis for current implementation of 3D in cinema and recently emerging TV services.

Although two-view stereo is likely to be the representation of choice of the industry for some time to come, it has a number of limitations. These arise mainly out of the fact that the parameters of a two-camera capture rig have to be fixed during the capture and cannot be changed either in post-production or at the end-device. This limitation led to a number of European research initiatives in the 1990s that looked into topics including depth-based and model-based stereo representations. These representations allow for adjustments in post-production and at the user side. Moreover, they enable usage of more advanced display techniques, in particular auto-stereoscopic and holographic displays.

A. Scope

This paper aims to give an overview of some of the research and development (R&D) activities on subjects related to 3D-TV in Europe. Again, Europe has a long history of both research and implementation of these results by industry. However, such an overview can never be complete. We aim to give a brief overview of past research and a snap-shot of recent activities. The projects or working groups mentioned represent only a sample of the overall R&D effort and further literature is referenced where known to the authors of this contribution.

The paper tries to cover the main aspects of 3D-TV, including production, post-production, coding and distribution, end devices and human factors. Since this article focuses on the European point of view, it does not include detailed background on all these topics. Readers interested in more background information are referred to state-of-the-art overview articles, e.g. in this special issue [23].

The term 3D-TV was defined by broadcast industry, but there are aspects of 3D-imaging that are worked on by other industries and research disciplines. The one closest and with most influence is the cinema industry. Due to the change to digital production methods there is an increasing degree of convergence between the cinema and broadcast industries. In particular high-end HD[1]-productions and (low-end) cinema productions increasingly share the same equipment and production methods. This trend is expected to continue in the context of 3D-productions.

Moreover, recent 3D-cinema productions are expected to be an important source content for 3D-TV services and therefore are seen as an important driver for the successful introduction of 3D-TV. For this reason we will cover some aspects of 3D-cinema production in this article.

Another related area is 3D-imaging in virtual reality. There is a certain overlap between VR[2] techniques and 3D-TV. For example, techniques to combine real and virtual scenes have been applied to TV and film productions in a number of projects and would be or have already been demonstrated to be applicable to 3D-TV as well.

B. European Economy + Research Community

Europe has a developed economy and research landscape. Research and industrial activities are considering all aspects of media production in general and broadcasting in particular.

[1]High Definition.

[2]Viritual Reality.

Manuscript received October 18, 2010; revised January 24, 2011; accepted January 28, 2011. Date of publication April 07, 2011; date of current version May 25, 2011.

O. Grau is with the BBC R&D Centre House, W12 7SB, U.K. (e-mail: Oliver.Grau@bbc.co.uk).

T. Borel is with the Technicolor Research & Innovation 1, 35576 Cesson-Sévigné, France (e-mail: thierry.borel@technicolor.com).

P. Kauff and R. Tanger are with the Fraunhofer-Institut für Nachrichtentechnik Heinrich-Hertz-Institut Einsteinufer, 37 10587 Berlin, Germany (e-mail: kauff@hhi.fraunhofer.de).

A. Smolic is with the Disney Research, Zurich Clausiunsstrasse, 49 8092 Zurich, Switzerland (e-mail: smolic@disneyresearch.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBC.2011.2126350

On the commercial or industrial side of the content production pipeline, which includes content capture, post-production, distribution and end devices, Europe has major commercial players in all fields. However, in recent decades some shifts in industries have taken place: The production of consumer end devices has, to a large extent, moved out of Europe to places like Asia, but specialized niches still exist.

In the areas of professional equipment manufacture, including software, Europe maintains a competitive industry. This includes the manufacture of capture hardware, including cameras, supporting devices and digital capture servers. The same applies for transmission technology, which has some overlap with the ICT[3] industry. Further, a number of software manufacturer produce products for production and post-production purposes.

The service sector of content production, including post-production for broadcast and film has an important role in Europe. It is believed that for the UK alone the film and broadcast industry contributed over £4.5 billion to the UK GDP[4] in 2009 and employment in this sector went up by 7% compared to 2007 [38]. Research in Europe is carried out in national government funded research bodies, including Universities and industrial research. An important funding source on an international level is provided by the European Commission. The funding is structured in the so-called Framework Program. The budget for ICT-related research in the recent Framework 7 (FP7) for the years 2007-2013 is over 9bn Euros [46].

The funding by the Commission on a European level is complemented by a number of national and regional funding bodies that usually aim to strengthen particular national or local interests or more fundamental research.

C. Overview of Paper

The remainder of this paper is organized in three main parts: The next section gives a brief overview of the history of 3D-TV in Europe and puts European key research initiatives into context.

Section III gives an overview of important recent research topics. Section IV complements this view with a description of selected European research projects and activities on international and national levels.

The paper finishes with some concluding remarks.

II. History of 3D-TV in Europe

It is an old dream of humans to reproduce 3D sensations in pictures, photographs and movies. Painters over the centuries achieved this by exploiting pictorial 3D effects like shading or perspective. In 1692, it was the French painter G. A. Bois-Clair who discovered for the first time that the 3D sensation can be enhanced considerably by creating paintings containing two distinct images– one for the left eye and one for the right eye—instead of presenting just one image. The separation between the two views was achieved by a grid of vertical laths in front of the painting—an antecessor of today's auto-stereoscopic displays using parallax barriers or lenticular lenses [1]. In 1838, Sir Charles Wheatstone presented the mirror stereoscope—an optical device by which two different perspective drawings could be viewed stereoscopically through two angled mirrors [1]. After the invention of photography in 1839, Sir David Brewster developed the first stereoscopic camera in 1849 and following that watching 3D photographs in stereoscopes became a very popular amusement in European museums and exhibitions [3].

A new era of 3D reproduction began with the advent of motion pictures. The history of 3D movies is indeed as long as the story of the cinema itself. The Lumière brothers were the first to show moving 3D pictures at the 1903 world exhibition in Paris. A full-length 3D movie was then presented in Los Angles in 1922, and, in 1928, John Logie Baird applied the principle of stereoscopy to an experimental 3D-TV set-up using Nipkow's perforated disc technology [4]. In spite of these early demonstrations and many subsequent efforts to establish 3D movies in cinema, especially in the 1950s, the commercial break-through failed to appear for many years. Technical deficiencies and insufficient quality hampered a successful introduction in the entertainment segment. Apart from some theme park installations and other niche market applications, the only notable exception was IMAX 3D with some commercial success since its introduction in 1986 and with a limited number of specialized 3D productions each year.

The situation started to change in the early 1990s with the advent of the transition from analog to digital TV services. Worldwide research activities were launched with the aim of developing standards, technologies and production facilities for 3D-TV. Key European projects like DISTIMA [6], [9] and MIRAGE [7] made some of the first experiments with electronic production of 3D content and laid the foundations for more European 3D-TV research. As an example, the MPEG-2 Multi-View-Profile (MVP), which was finally approved in 1996, is one result of this early research period [10].

Today this slow start in the digital 3D-TV era has been accelerated dramatically due to the recent introduction of digital 3D cinema. Hollywood is releasing more and more 3D productions with increasing revenues in recent years and the value-added chain of 3D cinema is under intensive worldwide discussion. Major events have been broadcast in 3D and the first commercial 3D-TV channels are on air now. However, almost all these recent 3D-TV activities rely on the straightforward concept of an end-to-end stereoscopic video chain as originally proposed by Wheatstone, i.e., on the capturing, transmission, and display of two separate video streams, one for the left and one for the right eye. Due to these restrictions, stereo capture had to fit to the display geometry and vice versa. Display properties and viewing conditions as well as stereoscopic 3D production rules have to be taken into account by the production staff from the beginning during shooting. As explained in more detail in Section III-A, an important example is the exploitation of an available depth budget to achieve a comfortable viewing experience for the viewer. Clearly, these constraints made stereoscopic 3D productions and re-purposing of stereoscopic 3D content for different application and viewing platforms extremely complicated.

In the late 1990s some researchers therefore came to the conclusion that it is essential to separate capture and display geometry by using emerging methods of computer vision, 3D video

[3]Information and communication technologies.

[4]Gross domestic product.

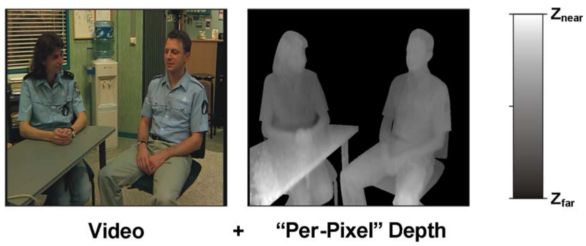

Fig. 1. 3D-TV data representation using video-plus-depth.

processing, and depth image based rendering (DIBR). The main idea was to derive an almost generic depth-based data representation from captured images to decouple camera and display geometry and, with it, to obtain the highest flexibility and adaptability at the display side. In principle, this can be achieved by estimating depth information from a given stereo or multi-view camera system and to use these depth data to re-calculate at the receiver side a virtual stereo pair that is perfectly adapted to display properties and related viewing conditions. In 1998, the European PANORAMA project [8], [9], [21] was one of the first research activities that demonstrated the feasibility and the potential of such a depth based process of stereo adaptation.

While the PANORAMA project was focused on stereoscopic videoconferencing, the European follow-up project ATTEST [12] took up this concept and applied it to the requirements of a 3D-TV processing chain. The ATTEST system is based on the transmission of regular video images enriched with depth maps providing a Z-value for each pixel—a data representation that is often called video-plus-depth (see Fig. 1). The final stereo images are then reconstructed at the receiver side by using DIBR techniques. This approach gives the ATTEST concept some crucial advantages over conventional 3D-TV proposals, such as back ward compatibility to existing 2D services for Digital Video Broadcast (DVB), efficient compression capabilities, and a high adaptability to 3D display properties, viewing conditions, and user preferences [13]. A Real-time transmission demonstrator of the ATTEST system was shown at IBC 2003. During this time, similar real-time 3D-TV demonstrators using video-plus-depth and multi-view representations were also shown at SIGGRAPH[5] 2004 in Los Angeles, and IFA 2005 [14], [15].

The capture and representation of a scene in video-plus-depth brings a number of challenges: When a scene is captured by a two-camera stereoscopic rig, it is fundamentally impossible to compute an absolutely correct depth map for each view, because of occlusions and other model violation (like the lack of texture). Different capture and reconstruction techniques have been suggested: The ATTEST project investigated the use of an active time-of-flight sensor co-axial to one main camera [12]. Another configuration is to use a central, principal camera and two satellite cameras to reduce the effects of occlusions; this has been demonstrated by the EU MetaVision project, although not for 3D-TV, but for general production applications [43]. A similar approach has been investigated specifically in the context of 3D-TV production [16], [17]. This work is now being continued in the European FP7 project 2020-3D-Media (see Section IV-A). Several capture configurations have also been investigated later by the 3D4YOU project (see also Section IV and [24]).

The depth information or any number of (virtual) stereoscopic views of a scene can be derived from a 3D scene description using CGI methods. The most common way is using a surface description (e.g. triangular meshes) plus color or texture.

The automatic production of static 3D CGI scene descriptions from images has been pioneered by projects, like Mona Lisa [39], VANGARD [21] and others. The Mona Lisa project focused on the development of VR methods for (2D-) TV production and the integration of real and virtual scene elements. The EU MIRAGE project has demonstrated live generation of two-view stereo using this approach for synthetic content [7].

While Mona Lisa was capturing studio-based action in a 2D-representation (image+alpha) the ORIGAMI project [41] developed methods for a complete 3D object-based production using wide-baseline multi-camera set-ups, targeting special effects. However, the principles have been also applied to sport scenes in the UK TSB iview project [42] and the usefulness of the methods for 3D-TV productions have been recently demonstrated in the 3D4YOU and i3DLive projects (see also Section IV).

For the cinema it seems that the audience accepts wearing of glasses to enjoy two-view stereoscopic program s and this is also the preferred option of industry to introduce 3D-TV in the home. However, as mentioned previously, researchers in Europe championed for a long time the concept that viewers at home would prefer a glasses-free display system. This has led to a number of research projects and products of auto-stereoscopic displays and the concept still has a number of supporters.

Auto-stereoscopic display techniques were the subject of a number of European and national research projects. A number of projects started with a holistic view and developed end-to-end chains for production, transmission and display of 3D information. The EU VANGARD project was focused on the generation of static 3D scene models, but also developed an active barrier-based auto-stereoscopic display to view the generated content [13]. Similarly the UK Prometheus project developed methods to generate object-based 3D content of (studio-based) action and auto-stereoscopic displays based on integral imaging [40]. The EU PANORAMA and ATTEST projects developed auto-stereoscopic display techniques that were targeted to video-conferencing and 3D-TV.

Anticipating these developments, the MPEG group of ISO/IEC established at an early stage the 3DAV (3D Audio/Visual) ad-hoc group to investigate the needs for standardization in 3D [11], and it has since launched several standardization efforts on 3D data representations. These include multi-view coding (MVC) without additional depth information as well as video-plus-depth formats for single and multiple video streams. An overview of the different formats and their current status within MPEG is given in Section III-C.

III. Subjects of R&D Related to 3D-TV in Europe

This section outlines a number of research topics with European contributions.

[5]http://www.siggraph.org/.

A. New Production Tools

This section highlights examples of on-set and post-production technology investigated in Europe recently, which of course represent only a very sparse sample of past and ongoing research efforts. The first is about improving stereoscopic capture during shooting, the second discusses stereoscopic re-mapping as an alternative to depth-based view synthesis for changing the apparent depth range, and the third elaborates on 2D-3D conversion.

1) On-Set Stereo Assistance Systems: It is well known that improper creation of stereo content can easily result in a bad user experience. In fact, the depth impression from a 3D display creates an illusion in the human visual system and if not done properly, could produce consequences for human 3D perception such as eye strain and visual fatigue. Production of good stereo content is therefore a difficult art that requires a variety of technical, psychological, and creative skills and has to consider perception and display capabilities.

Therefore, to create good stereo, stereographers have to take into account a variety of conditions, guidelines and rules right from the beginning of the production chain. One main issue is to ensure that the whole scene usually remains within a so-called Comfortable Viewing Range (CVR) of the targeted 3D viewing conditions (e.g. ratio of screen width and viewing distance, also called presence factor). The 3D experience is generally comfortable if all scene elements stay in this limited depth space close to the screen. As the available depth volume is restricted compared to the real 3D world, the difficult job of a stereographer is "*to bring the whole real world inside this virtual space called the comfort zone*".

There are two main parameters by which this production rule can be controlled. One is the inter-axial distance (stereo baseline) which controls the overall range of depth, i.e., the depth volume of the reproduced scene. The other one is the convergence which controls the depth position of the scene in relation to the screen, i.e., which parts of the scene appear behind and which in front of the screen, respectively.

Further issues are the avoidance of undesired effects causing retinal rivalry. This refers to any kind of geometrical distortions (keystone, vertical misalignment, lens distortions, mismatches in focal length, etc.), to unbalanced photometry (color mismatches, differences in sharpness, brightness, contrast or gamma, etc.) and to perception conflicts (stereo framing, stereoscopic window violation, extreme out-screening, etc.).

Apart from a mismatching stereo baseline, these deficiencies can usually be corrected to some extent during post-production. Nevertheless, any careful planning and execution of stereo shooting tries to avoid them from the beginning. This includes an accurate rigging and calibration of the stereo cameras, good adjustment and matching of electronic and optical camera parameters and, above all, the adaptation of the stereo baseline to the depth structure of the scene content. However, this adjustment is time-consuming manual work and requires skilled staff to do it properly.

These efforts were accepted as long as 3D productions were only addressing a small niche market. However, due of the rapid increase of 3D productions during the last few years, there is now a rising demand for efficient 3D production tools assisting stereographers and the camera team on the set. The main goals of such assistance systems for stereo shooting are to ease rigging, to save time for adjustments, to change them quickly from take to take and to also allow less experienced camera staff to use proper stereo settings.

Fig. 2. STAN as demonstrated at NAB 2009.

One example of such an assistance system is the Stereoscopic Analyzer (STAN) [56]. Fig. 2 shows an application of the STAN to a mirror rig with two ARRI D21 cameras and a side-by-side rig with two MicroHD cameras from Fraunhofer IIS, as shown for the first time at NAB 2009.

An image-based scene analysis system estimates in real-time the relative pose of the two cameras in order to allow optimal camera alignment and lens settings directly on the set. In addition, it detects the position of near- and far objects in the scene to derive the optimal inter-axial distance (stereo baseline), and gives a framing alert in the case of stereoscopic window violation (when an object appears on the edge of one view and cannot be seen in the other). A 3D viewfinder visualizes the stereo analysis via an intuitive user interface. Analyzed stereo parameters can be stored as metadata for later postproduction or can be used for an automatic real-time correction of undesired vertical disparities and geometrical distortions through image rectification for live production.

2) Stereoscopic Re-Mapping: Post-production of stereoscopic 3D, real time or offline often requires estimation of depth or disparity data in order to perform view synthesis. Such view synthesis would then change stereo properties, i.e. change global and local depth impression, correct stereoscopic errors (framing violations, excessive disparity), etc. The state-of-the-art in 3D video editing relies on such depth-based approaches today [54]. However, methods that rely on depth or disparity data suffer from a number of drawbacks. Depth or disparity estimation is still an error prone task, often resulting in unreliable depth or disparity maps, which may lead to errors in the rendered output views. Interactive correction and editing of depth or disparity data is often necessary, which is expensive and cumbersome. Further, gaps often appear after rendering for different camera baselines which require in-painting, which is another mainly interactive task adding additional work and cost to the stereo editing process. Finally, depth or disparity based methods require knowledge about camera calibration to perform tasks like rectification, etc.

Recently, alternative methods for 3D editing have been proposed which do not rely on dense depth or disparity estimation but use warping in the image domain instead [25]. Neither calibration nor in-painting is necessary within practical limits. The proposed stereoscopic warping computes a sparse set of disparities for a number of feature points in the images that can be

Fig. 3. (Left) Original stereo image; (right) remapped stereo image. The cow is shifted back inside the screen while keeping volume and background remains unchanged.

tracked with high accuracy and reliability. Additionally, a novel kind of 3D saliency information is extracted from the images, which correspond to regions of high user attention in the stereoscopic views. Sparse disparities and saliency are used to compute a warping function related to the stereo pair, which distorts both images in a way that a desired/postulated change of depth perception is implemented. The warping respects salient elements keeping them as undistorted as possible and distributes stronger deformations in less salient image regions.

The desired/postulated changes of depth perception are formulated via a mathematical framework of disparity mapping operators, i.e. mapping functions $\Phi(d)$ that convert a given input disparity distribution into a desired/postulated output disparity distribution. These mapping functions can be non-linear and locally adaptive. Definition is derived e.g. from aspects of stereo perception, technological issues like display capabilities and viewing conditions, and artistic constraints. An example is illustrated in Fig. 3 (from [26]). The left image shows an original stereo pair in anaglyph representation. The right image is the result of a stereoscopic mapping operation. The head of the cow was shifted back ward, reducing the depth position without changing depth volume, while the background remained unchanged. The operation was non-linear in disparity range and locally adaptive. Note that this is not just a simple shift parallax manipulation.

Fig. 3 is only one example of a manipulation of depth perception. In principle the technology enables full control over the disparity of a stereo pair after it was captured. The application range is huge. Besides post-production (correction of errors such as stereoscopic window violations, changing depth as desired) the algorithm could be implemented in camera systems or close behind in the signal flow to automatically correct stereoscopic errors. This would ensure good stereo quality for instance in live broadcast scenarios. Display adaptation is another application area. Depth can be scaled to adjust the depth impression for different viewing conditions. A depth impression controller on the remote control could be used like we use brightness or contrast controllers of a TV set today. Finally, it has been shown that the algorithms can be used for 2D to 3D conversion, i.e. generation of stereoscopic video from a monoscopic source.

3) Offline 2D-3D Conversion: Offline semi-automatic 2D to 3D conversion has two applications: The most obvious one is in transforming existing 2D content libraries (movies, documentaries, games, . . .) into a new 3D version of it (recent examples being Star Wars by Georges Lucas, and Oceans by Jacques Perrin). The second business opportunity is about creating completely new 3D content that is totally or partially shot in 2D mode for reasons related to the difficulty of using 3D stereo camera rigs in extreme shooting conditions (such as helicopter, aerial or sub-marine stunts) or for reasons linked to the shooting process preferences of the film director (Alice in Wonderland by Tim Burton).

In this paper, we are not describing R&D work that has been or is being done in terms of fully automatic real time 2D to 3D conversion that is being introduced in some (but not all) 3D-TV sets. We consider that, whatever the level of efforts spent in R&D on these concepts, the achieved 3D experience will never be satisfactory enough and could even jeopardize the 3D content deployment to homes since the consumer world would not accept eye strains and headaches associated with poor 3D experiences and this would discourage people from investing in 3D equipment (TV sets, Blu ray players, set-top boxes, camcorders, . . .).

On the contrary, offline 2D to 3D conversion aims at the production of high quality content at the post-production stage. It started as a fully manual process. Thanks to R&D projects, which have been, and are still, being conducted in Europe and around the rest of the world, more and more semi-automatic conversion tools are being developed to save time and money. The main technical challenges are the generation of a depth map from a single view, synthesizing a second view out of it using semi-automatic in-painting techniques and propagating the results to neighboring frames through specific tracking techniques. Nevertheless, the manual operations of artists are still intensive and critical in the process of achieving high quality results.

In Europe, R&D efforts started just before the year 2000 leading to some products that are already available on the market.

In 2007 the BlueBox service suite of advanced 3D content creation tools that converts 2D video content to stereoscopic 3D was released [28]. The major 3D content creation service on the BlueBox allows the content creators to generate high quality semi-automatic conversion of existing 2D material. Nevertheless, this software suite was mainly dedicated to the production of 3D content for auto-stereoscopic 3D displays using 2D-plus-Depth format.

In France, proprietary techniques have been developed to convert existing 2D movies into high quality stereoscopic 3D content for cinema applications or special venues (such as theme parks). An experience of more than 15 years has been built in this area and provided solutions are competitive compared to non European companies working in this domain. Documentaries have been produced (e.g. extreme sports). It was shot in 2D and converted after ward in post-production into 3D stereoscopic format.

The ANR[6] French research project "3D Comfort & Acceptance" is aiming at better understanding of 3D vision in terms of human factors and comfort and creating a label of quality and recommendation that will help 3D broadcasters, 3D producers and 3D movie directors to provide high-quality 3D content to the audience on cinema, TV and PCs.

[6]ANR : Agence Nationale de la Recherche.

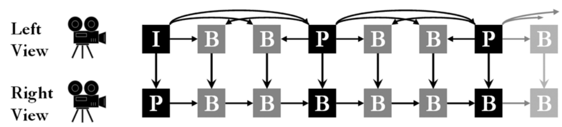

Fig. 4. Stereo coding, combined temporal/interview prediction.

As the movie business, and in particular Hollywood, is the prime user for 2D to 3D conversion, another international entertainment services company, with R&D centers based in Europe, has launched a 3D research program including the topic of 2D to 3D conversion. The goal here is to develop algorithms to accelerate the conversion process by providing more efficient semi-automatic tools that will assist the artist with his work.

Another set of tools for stereoscopic 3D production is being developed in the UK-based TSB[7] i3Dlive project. The project follows two main streams: the first of these is the development of semi-interactive offline tools for the production of 3D movies, which aim to generate stereoscopic 3D from a one principal camera, augmented with one or more witness cameras in a narrow-baseline configuration. The other work stream looks into fully automated real-time 3D modeling from wide-baseline multi-camera configurations for broadcast applications, including stereoscopic 3D and free-viewpoint video. More details on this project are presented in Section IV-C2.

B. Coding

In this section we elaborate on coding of 3D video and associated data. International standards, mainly MPEG and ITU, are most relevant. Specifically contributions from Europe are highlighted here, which have always been part of global efforts with strong participation from around the world. More detailed overviews of coding for 3D can be found, for example in [55].

1) Available Standard 3D Video Formats: This section gives an overview of available 3D video standards divided into video only and video plus depth formats.

Video only formats: Research on video coding has a long tradition in Europe. For instance the MOMUSYS project played a leading role in the development of the first parts of the MPEG-4 standard, which was a global effort with contributors from around the world. MPEG-4 as well as its predecessor MPEG-2 already provided the means to encode stereoscopic video such as the MPEG-2 Multi-view Profile. The basic principle is to increase coding efficiency by combined temporal/interview prediction as illustrated in Fig. 4. For more than 2 views this is easily extended to Multi-view Video Coding (MVC). A corresponding MPEG-ITU standard was released in 2008, which is an extension of H.264/AVC [26]. It can also be applied to 2 views. This is specified in the so called Stereo High Profile. MVC is currently the most efficient way to perform stereo and multi-view video coding. The first commercial adoption of MVC was in the Blu-ray 3D specification. MVC was developed with the strong participation of researchers from the 3DTV Network of Excellence, who contributed basic technical solutions as well as leadership and management of standards development.

[7]TSB: Technology Strategy Board.

A simple way to use existing video codecs for stereo video transmission is to apply temporal or spatial interleaving. This is also referred to as frame-compatible coding. With spatial interleaving, resolution is reduced so that the data from left and right views can be packed into a single frame. There are various ways of arranging the data, e.g., a side-by-side format in which the right view is squeezed into the right side of the frame and the left view into the left side of the frame, or a top-bottom format in which left and right views are squeezed into the top and bottom of a frame, respectively. The data for each view may also be filtered using a quincunx sampling (or checkerboard format) and interleaved; the samples may also be packed into one of the other arrangements. With time multiplexing, the left and right views are alternating in time, with a reduced temporal resolution for each view. Amendments to H.264/AVC have been developed that signal the new frame packing arrangements as so called SEI messages. This signaling could be used at the receiver to de-interleave the video and render stereo to the display. Of course, legacy devices without knowledge of the interleaving format or the new signaling will not be able to perform the de-interleaving and hence such video encoded in this format is not usable for those devices. The simplicity and compatibility to existing infrastructure makes stereo interleaving formats very attractive for fast market introduction.

Video plus depth format: The ATTEST project laid the foundations for a number of technologies in 3D-TV including usage of the so-called video plus depth format (V+D), as illustrated in Fig. 1. A video signal and a per pixel depth map is transmitted to the user. From the video and depth information, a stereo pair can be rendered by 3D warping at the decoder. Per pixel depth data can be regarded as a monochromatic, luminance-only video signal. The depth range is restricted to a range between two extremes Z_{near} and Z_{far} indicating the minimum and maximum distance of the corresponding 3D point from the camera respectively. Typically this depth range is quantized with 8 bits in a logarithmic scale, i.e., the closest point is associated with the value 255 and the most distant point is associated with the value 0. With that, the depth map is specified as a gray scale image. These Grey scale images can be fed into the luminance channel of a video signal and the chrominance can be set to a constant value. The resulting standard video signal can then be processed by any state-of-the-art video codec.

In some cases such depth data can be efficiently compressed at 10–20% of the bit rate which is necessary to encode the color video [27], while still providing good quality rendered views. However, for more complex depth data the necessary bit rate can reach the color bit rate. Recently, alternative approaches for depth coding based on so-called Platelets were proposed, which may perform better than state-of-the-art video codecs such as H.264/AVC [29].

The ability to generate the stereo pair from V+D at the decoder is an extended functionality compared to conventional stereo video. It means that the stereo impression can be adjusted and customized after transmission. Also more than 2 views can be generated at the decoder enabling support of multi-view displays and head motion parallax viewing within practical limits.

The concept of V+D is particularly interesting due to the backward compatibility and extended functionality. Moreover it is possible to use available video codecs. It is only necessary to

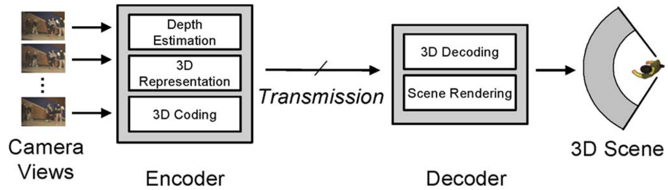

Fig. 5. Advanced 3D video processing chain.

specify high-level syntax that allows a decoder to interpret two incoming video streams correctly as color and depth. Additionally, information about depth range (Z_{near} and Z_{far}) needs to be transmitted. Therefore MPEG specified a corresponding container format "ISO/IEC 23002-3 Representation of Auxiliary Video and Supplemental Information", also known as MPEG-C Part 3, for video plus depth data [30] in early 2007. This standard already enables 3D video based on video plus depth.

However, the advantages of V+D over conventional stereo video are offset by increased complexity for both sender side and receiver side. View synthesis has to be performed after decoding to generate the 2nd view of the stereo pair. Before encoding, the depth data have to be generated. This is usually done by depth/disparity estimation from a captured stereo pair. Such algorithms can be highly complex and are still error prone.

2) Advanced 3D Video Formats: The 3D video formats presented so far are ready to use, however, they do not satisfy all requirements in an efficient way. This includes wide range multi-view auto-stereoscopic displays and free viewpoint video, where the user can chose their own viewpoint. Display adaptation is possible but very limited with V+D. Such advanced 3D video applications require a 3D video format that allows rendering a continuum of output views or a very large number of different output views at the decoder. MVC does not support a continuum and is inefficient if the number of views to be transmitted is large. V+D support only a very limited continuum around the available original view since view synthesis artifacts increase dramatically with the distance from the virtual viewpoint.

The basic concept of advanced 3D video formats is illustrated in Fig. 5. The decoder receives a coded representation (bitstream) of data in the advanced 3D video format. The data is decoded and then used for rendering of arbitrary views within a certain operating range. This allows all requirements to be satisfied. At the encoder a real world 3D scene is typically captured by multiple cameras, and a 3D representation is extracted from the multiple camera signals. Most formats under study use multiple depth maps for 3D scene representation [31], [32], but other approaches based on 3D meshes, point clouds, quads and other geometric primitives are under study as well. Sender side processing thus includes depth estimation or another kind of 3D reconstruction. Any appropriate technology and processing can be applied (including synthetic creation, animation) as long as the output is in the correct 3D video format. With that the required decoupling of content creation and display is achieved.

The advanced 3D video processing chain involves a number of highly complex and error prone processing steps, such as depth estimation or some other 3D reconstruction at the sender, and rendering of virtual views at the receiver. Examples of such advanced 3D video formats include multi-view video plus depth (MVD) [33], [34], layered depth video (LDV) [35], or depth enhanced stereo (DES) [36].They have in common that they use per pixel depth maps, video and other data.

Based on evolving market needs, MPEG is now considering a new phase of standardization. The new standard should support advanced applications as described in the previous section via an advanced 3D video format in efficiently coded representation [37]. A number of European projects have been and/or continue to be active in this context including for instance 3D4YOU, 3DPHONE, 2020 3D MEDIA, MOBILE3DTV, MUSCADE.

3) Forums and Activities Related to Standards: 3D@Home is an international consortium aiming at accelerating the adoption of quality 3D into the living room. It comprises five Steering Teams covering the whole range from creation and transmission up to 3D displays and human factors. Almost one third of the members are based in Europe; the rest are located in the US and in Asia. 3D@Home started its work with an inaugural meeting during NAB 2008. Meanwhile a large set of white papers, first evaluation content, a 3D display technology classification, a glossary and a set of expert recommendations for good 3D content have been produced.

In fall 2008 SMPTE (Society for Motion Picture and Television Engineers) established a Task Force on 3D to the Home to evaluate the use cases and requirements for a 3D Home Master standard. In the beginning the idea was mainly to build a standard around Stereoscopic 3D technology. Driven mainly from European participants the scope was widened quickly to ward a more flexible approach including depth maps as an optional component. The Task Force finished its work in April 2009 with a final report. In summer 2009 the Work Group on 3D Home Master (located in the 10E Essence Technical Committee) was established. In 2011 a first version of the standard can be expected. Motivated by discussions in the WG, an additional AHG "3D Frame Compatible" was initiated mainly to list and categorize the different possibilities to store Stereoscopic 3D in a frame compatible way. Further, SMPTE recently started activities for specification of disparity data associated with stereo video in a group called "AHG on 3D Disparity Maps".

C. End Devices

There have been a number of European projects looking into the developments of new 3D-TV end devices, in particular auto-stereoscopic displays. There are a number of products available to the market from European companies (e.g. Tridelity, Alioscopy, NewSight), mainly for specific applications.

Current research is focusing on getting end devices into the entertainment segment, i.e. 3D-TV for the home or on mobile devices.

1) 3D VIVANT: The 3D VIVANT[8] project is intending to provide an immersive 3D visual and spatial sound experience without the need for spectacles or headphones. The project is designing a 3D 'Holoscopic' single aperture camera which will provide real-time capture of 3D scenes. The images will ultimately be displayed on a dedicated 3D display using the principles of holographic geometry for high quality viewing without spectacles.

[8]http://www.3dvivant.eu/.

Other parts of the project investigate 3D object extraction from captured images, depth processing, and delivery using a scalable 3D Holoscopic robust video coding method for transmission. In addition to broadcasting, the project is also developing new applications of 3D Holoscopy for the internet to provide online hyper linking of 3D objects. The project plans to develop a new spatial sound system for a more natural representation of sound fields alongside captured images and playback through loudspeaker arrays in a domestic environment.

Real 3D—Digital Holography for the Capture, Processing, and Display of 3D and 4D Real-World Objects': The Real 3D project[9] aims to overcome the restrictions found in conventional stereoscopic and auto-stereoscopic display technologies, like mismatches between apparent object depth and the observer's focus distance, by using holography [50]. A 3D holographic acquisition system based on digital camera technology will be developed, with at least 8 million pixels each of side length no greater than 9 μm, arranged non-uniformly in a circular configuration. The acquisition system will be capable of recording holographic video of the 3D scene.

On the display side a 3D holographic display system based on liquid crystal on silicon (LCOS) technology will be developed. The reconstruction system will be capable of displaying holographic video of the 3D scene.

2) Helium 3D: HELIUM3D[10] aims to develop a new auto-stereoscopic display that overcomes limitations of currently-available technology. The project approach is based on direct-view RGB laser projection via a low loss transparent display screen to the eyes of viewers. The display aims to support multiple viewers in a quality that is superior to currently-available displays in brightness and color gamut. The display will employ viewer and gesture tracking and interaction to maximize the viewing experience.

D. Human Issues/Perception

Since 2009, more and more 3D movies have been released, and the success of Avatar by James Cameron has confirmed the interest of the audience for this way of telling a story. In movie theaters, 3D digital projection technology is now mature and also the 3D basic shooting grammar is respected most of the time (but not always) by content creators. Spectators can then enjoy a 2-hour-long feature film with no or limited eye strains. Also, wearing glasses in this context seems to be accepted by the majority of people.

Nevertheless, there is a huge step between occasionally watching a 3D movie in a cinema and watching 3D-TV almost every day at home. Even if the 3D content is technically well-produced (which is again not always the case), the viewing conditions are totally different (duration, frequency, viewing distance, screen size, ambient light, comfort of the glasses, etc.). In addition to movies, 3D live events such as sports and concerts will be distributed to the home; these events cannot be as well produced as movies as no off-line content improvement can be performed during post-production. Finally, not all TV programs will be in 3D in the future, forcing consumers to put on and take off their glasses frequently which could lead to a kind of weariness. The understanding of these human factors is key to successfully deploying 3D-TV in the home. This has been well understood by the different parties and a number of R&D projects have been and are still being conducted to measure the Quality of Experience (QoE) associated with 3D-TV programs.

In Europe, several R&D activities have been launched in this area:

Futurim@ge was a French funded research project in the area of the "pole de compétitivité Images & Réseaux". The main goal of this project was to study future image formats like 1080p, High Dynamic Range (HDR) and 3D. One of their work packages involved the evaluation of performances of various 3D technical solutions from the user stand point. They presented some QoE results on five particular topics:

- Evaluation of the influence of depth rendering on the QoE for auto-stereoscopic displays
- Description of a crosstalk-reduction algorithm
- Presentation of a protocol for subjective tests to evaluate visual fatigue
- Metric proposal for the evaluation of 3D picture quality
- Impact evaluation of the reduction of 3D image resolution on user perception

In 2009, a human perception related platform called 3D-Fovea was founded in France. It is made up of a partnership between the Ophthalmology departments of CHU[11] Brest and CHU Angers, GIS[12] Marsouin, content providers and equipment manufacturers for the digital cinema industry like E3S, which provide 3D active glasses. The objective of 3D Fovea is to test equipment and study how spectators perceive 3D effects. This consortium will propose new tools to measure eye strain and to assess the impact of all elements of the 3D-TV chain, from content production to content rendering. 3D Fovea will develop recommendations in terms of both technologies to support 3D and content creation guidelines.

MUSCADE[13] is a 3 years long IST-FP7 European research project that started in January 2010 (See Section IV-A. for more details). One aspect of MUSCADE is "Quality of Experience" intending to define quality metrics for 3D video and 3D audio visual content. This considered in a holistic way including aspects of 3D audiovisual content (quality of audio and video) created, delivered and presented using the MUSCADE technologies developed in other work packages.

IV. European 3D-TV Research Projects

This section gives an overview over a selection of recent European projects not discussed in previous sections. It starts with recent EU funded projects and a selection of national projects.

A. EU IST FP7

Table I lists a selection of recent research projects funded under the IST (Information Society Technologies) area of Framework 7 of the European Commission [46] that contribute to 3D-TV. The table gives an indication of the areas in which

[9]http://www.digitalholography.eu/.

[10]http://www.helium3d.eu/.

[11]CHU : Centre Hospitalier Universitaire : University Hospital Center.

[12]GIS: Groupement d'Intérêt Scientifique : Scientific research consortium based in Brittany, France working on information and communication technologies.

[13]MUSCADE: Multimedia Scalable 3D for Europe.

TABLE I
RECENT EU-FUNDED PROJECTS

Project name	3D-Production	Trans-mission	End-devices	Mobile 3D
MUSCADE	X	X		
3D4YOU	X	X		
2020 3D media	X	X		
i3Dpost	X			
SkyMedia	X	X		
DIOMEDES		X		
3DPHONE			X	X
MOBILE3DTV		X		X
3D VIVANT	X	X	X	
Real3D			X	
Helium 3D			X	

the projects are active. An overview and web-links to the projects can be found in [48].

1) 3D4YOU: Content Generation and Delivery for 3D Television: *3D4YOU*[14] is working on advanced depth-based 3D representation formats and coding including dedicated capture technologies and conversion approaches [56]. The results are integrated into a broadcast chain to demonstrate the viability of the approaches investigated. 3D4YOU is developing technologies to generate advanced depth-based formats as described in Section III-B-2 from multi-view setups with a number of cameras, as well as hybrid setups comprising video and time-of-flight cameras. Special emphasis lies on disparity estimation, occlusion layer generation as well as fusion of time-of-flight and video. The achievable visual quality of the different formats is evaluated in subjective tests.

2) 2020 3D Media—Spatial Sound and Vision: *2020 3D Media*[15] is an Integrated Project within the IST-FP7. The project is developing a complete capture to display chain including 3D audio/video and immersive projection systems.

For capturing video along with a per-pixel depth map, the project is using a hybrid setup with an active structured light approach and trifocal disparity estimation using two additional satellite cameras. To capture multiple streams in real-time a solid-state-memory based capture solution is being developed, allowing the uncompressed recording of multiple videos using 10G and 1G Ethernet interfaces. In addition 2D to 3D image conversion algorithms are being developed exploiting dedicated segmentation algorithms. Post-production tools allow mixing of all implemented video formats. The work on coding focuses on the optimal bandwidth allocation for video and depth maps. Since formats vary significantly depending on the available technology in cinemas, a networked distribution system is being developed, allowing easy access to the right content. A flexible and modular multi-projection system covers a wide range of applications from 3D up to fully immersive displays. To support true 3D audio a complete processing chain including sound de-convolution and spatial rendering according to the target room properties is being developed.

[14]http://www.3d4you.eu/.

[15]http://www.20203dmedia.eu/.

3) i3DPost: Intelligent 3D Content Extraction and Manipulation for Film and Games: The IST-FP7 *i3DPost*[16] project is developing tools to extract 3D information from video. This data is then used in film and games production. The project is not specifically targeting stereoscopic 3D production, but the components being developed can support stereoscopic productions. Research is investigating the use of 3D information from camera tracking to improve motion estimation, matte extraction (or keying) and other common post-production operations. In addition, research is developing novel methods to capture structured models of the set and actors from dense 3D video reconstructed from multiple on-set "witness" cameras. This technology allows automatic reuse and manipulation of performers with changes in viewpoint and lighting.

Integration of multiple view 3D video analysis with on-set production allows the creation of video-quality actor and set models. Post-production tools have been extended to allow robust separation and manipulation of scene elements. On-set capture of actor performance is instrumented in a structured form for processing and rendering in a conventional graphics production pipeline whilst maintaining the visual quality of faces, body, and clothing movement. The representation of performance enables the modification in post-production of camera view, movement, and lighting to render novel content.

4) MUSCADE: Multimedia Scalable 3D for Europe: MUSCADE[17] intends to define, develop, validate and evaluate the technological innovations in 3D-TV capturing, data representation, compression, transmission and rendering required for a technically efficient and commercially successful Multi-View 3D-TV broadcast system. This is the next step after stereoscopic 3D broadcasting.

MUSCADE is developing an intelligent audio-visual capture and production assistance system designed for accurate capture of four or more video cameras along with multi-channel audio. The 3D-TV representation format that has been developed is generic in terms of supporting all current display technologies from 2D, S3D and multi-view up to light-field while also being scalable depending on the current bandwidth of the transmission channel. Several transmission technologies are being tested. New methodologies to link 3D video with 3D audio are being investigated and developed. The project is investigating and implementing techniques for scalable and robust transmission of 3D-TV content. In addition an interactive platform is being developed exploiting new ways of interaction in a true 3D environment. Finally, the project is defining and developing a quality metric to measure the Quality of Experience (QoE) for 3D audio-visual media.

5) Skymedia: Skymedia[18] aims to develop an end-to-end production pipeline for multimedia content taking shots from multiple points of view. An important concept of the project is the use of unmanned aerial vehicles (UAV) for the capture of HD and stereoscopic video of live events. The project also aims to develop wireless technology to gather and distribute the content

[16]http://www.i3dpost.eu/.

[17]http://www.muscade.eu/.

[18]http://ict-skymedia.eu/skymedia/.

either to mobile devices equipped with stereoscopic displays or to immersive mid-air screens, based on technology developed by FOGSCREEN Inc, who are one of the project partners.

6) Diomedes: The IST Diomedes[19] project focuses on coding and transmission of 3D immersive entertainment program s to the end user [49]. This includes multi-view video and 3D audio information.

3D Quality of Experience and 3D visual attention models will be used to code the audio-visual content into scalable layers. The distribution will combine DVB-T broadcast and P2P streaming models.

7) Mobile 3D-TV—Mobile 3DTV Content Delivery Optimization Over DVB-H System: Mobile 3D-TV[20] aims to develop TV services for mobile users. Mobile 3D-TV works on optimal formats for 3D video content for mobile usage taking into account the specific requirements in terms of compressibility and rendering efficiency of a small mobile device. Different 3D video representations are being evaluated including two-channel stereo video, video plus depth and mixed resolution stereo. A large set of the test material used in the project is available in a database[21] for the public. Several coding standards have been investigated for their usability in the mobile environment with a focus on video-plus-depth, mixed resolution stereo coding (MRSC) and MVC which has been chosen for the end-to-end system. Since DVB-H is the targeted distribution method, broadcasting equipment has been developed to set up a complete transmission chain. In addition simulation tools have been developed for application, linkage and physical layers.

8) 3DPHONE—All 3D Imaging Phone: The 3DPHONE[22] project aims to develop technologies and core applications by developing an all-3D imaging mobile phone. The aim of the project is to realize all fundamental functions of the phone i.e. media display, user interface (UI), and personal information management (PIM) applications in 3D, and usable without any stereo glasses. In particular the projects aims to develop a prototype 3D mobile phone equipped with an auto-stereoscopic display and multi-camera capture. Furthermore, new user interfaces and applications that make use of the specific 3D capabilities are being investigated.

The project is analyzing several video formats, including three depth-based formats, with respect to their suitability for bi-directional mobile video communication. For MVD[23] an extensive study has been carried out on depth coding, rendering algorithms and evaluation methods. For the prototype, classical stereo video (CSV) has been chosen and the H.264/MVC based coding solution has been optimized for the OMAP[24] platform.

B. Knowledge Networks

In 2004–2008 the European commission funded the 3DTV Network of Excellence [47] as part of Framework 6. The network was a platform for interdisciplinary research groups on all issues related to 3D-TV. The network has not been renewed, but has informally been replaced by the "3D, IMMERSIVE, INTERACTIVE MEDIA CLUSTER" [48]. The cluster is open to members of EU-funded projects in the field and aims to foster synergy effects between projects.

The European Broadcasting Union (EBU) has formed a 3D-TV Study Group [44] to gather requirements and information to advise the EBU members on 3D-TV related issues, especially in the production domain. The study group will look into topics of stereoscopic 3D production including cameras, codecs, sub-titling and production workflow.

C. National Program s

1) France: In June 2009, Mozart's famous Opera Don Giovanni was shot both in HD and 3D live in Rennes, France by a French consortium made up of more than 15 partners, and broadcast to cinema venues, in Rennes' town hall and in other locations in France. This project was led by the Opera House of Rennes under the umbrella of the "Images & Réseaux" local cluster. It was a world first since it was the first time that a cultural live spectacle was broadcast in real time, in 3D and using spatial sound based on HOA technology (high order Ambisonic). This event was a showcase of recent R&D achievements in the domain and triggered the setting up of a new and more ambitious collaborative project called 3DLive[25].

The project aims at creating a best-in-class competence in France for 3D stereo content production, distribution and rendering. This R&D project, which started in July 2009, is developing missing or non-optimized key technological bricks of a 3D-TV chain and will demonstrate its achievements through regular 3D shooting events (culture, sport, live TV shows); lessons learned from one event will allow improvements to be made for the next event.

2) United Kingdom: The main UK funding body for fundamental research in this area is the Engineering and Physical Science Research Council (EPSRC), which funds academic research. Several government agencies like the Technology Strategy Board (TSB) fund more applied collaborative R&D projects.

One project active in a field related to 3D-TV is the TSB i3DLive project[26]. The project is developing new tools for the extraction of 3D information from live action [45].

The project follows two directions: The project leader "The Foundry" is developing tools that allow the extraction of 3D information from live action using one principal camera with the help of "witness" cameras. A typical application would be on the set of a movie production. In this case the principal camera would be a high-grade film camera and the witness cameras could be video cameras. This application scenario would allow the production of movie-quality stereoscopic 3D program s without special stereo rigs.

The second application scenario of i3DLive considers wide-baseline camera configurations and aims to develop 3D processing in real-time (or near real-time) at video frame rate. The wide-baseline arrangement is well-suited to studio set-ups with locked-off cameras. In this configuration fast 3D reconstruction

[19]http://www.diomedes-project.eu/.

[20]http://sp.cs.tut.fi/mobile3dtv/.

[21]Mobile 3D-TV database: http://www.mobile3dtv.eu/video-plus-depth/.

[22]http://www.3dphone.org/.

[23]MVD: Multiple Video plus Depth.

[24]OMAP: Open Multimedia Application Platform.

[25]http://www.3dlive-project.com/.

[26]http://www.bbc.co.uk/rd/projects/2009/10/i3dlive.shtml.

methods can be applied based on the computation of the visual hull. The resulting 3D data can be used to either generate special effects, like free-viewpoint video, or stereoscopic content.

In September 2010 the i3DLive project captured two live events as part of a Super High Vision (SHV) test transmission between BBC in London and NHK in Tokyo. The SHV camera, developed by NHK provides a resolution of 7680×4320 pixels. The i3DLive project aims to demonstrate the use of the high-resolution camera input for high-quality special effects and stereoscopic content.

3) Germany: In spring 2008 eight German companies and research institutes started the R&D project PRIME (**Pr**oduction and Projection Technology for **I**mmersive **Me**dia)[27] with the mission to develop trend-setting technology and business-models for the implementation of 3D media into cinema, TV and gaming [52]. The project is funded by German Federal Ministry for Economy and Technology (BMWi). As shown in Figure 8, the research activities of PRIME cover innovative technologies for conventional 3D cinema, first generation of 3D-TV introduction based on glasses, next generation 3D-TV using auto-stereoscopy, production and visualization of panoramic 3D video for immersive applications and implications of 3D video to ward 3D gaming. To test the PRIME technology under real production and working conditions, a couple of 3D productions have been released in the framework of the project. Examples are 3D short movies like "24 Hour Race" which received the ITVA gold award [57], [58], "Topper gibt nicht auf!" shown first time at the INSIGHT OUT symposium in 2010 [59], and "Berlin Philharmoniker", an orchestral rehearsal with Sir Simon Rattle, probably the first 3D movie of a symphony orchestra [53].

Then, in September 2010, PRIME carried out two 3D live productions. The first one was a 3D live-on-tape production of a concert with the UK pop band "Marina & The Diamonds" at the SWR3 New Pop Festival in Baden-Baden, Germany. The event was recorded with six stereo rigs, including crane, dolly and steady-cam in the audience as well as three point-of-view cameras at the stage. For the live-on-tape production the six stereo signals were mixed in real-time at a 3D OB van and stored at a 3D HD recorder. The live-on-tape production will be broadcast in early 2011 by ARTE.

In late September 2010, the MMZ (Mitteldeutsches Medien Zentrum) in Halle, Germany, carried out a 3D live transmission of a pop concert with the German Hip-Hop band "Die fantastischen Vier (Fanta4)" from the Steintor theater in Halle into 94 cinemas in five European countries. The Prime partners were responsible for live 3D capturing with five stereo rigs including crane, dolly and on-stage steady-cams, for stereography, real-time stereo corrections as well as for live 3D directing and editing.

V. Conclusions and Future Directions

3D-TV has been a very active field of research in Europe since at least the 1990s. The European research community is covering all aspects of 3D-TV from capture, post-production, and coding, to transmission and end-user terminals. The framework set up by the European Commission provides funding that allows research on long term work strands and the establishment of chains of projects of strategic significance. This European scheme is complemented by national and regional research funds that seek to support more local economic interests.

One of the long-term goals set out in the late 1990s was to overcome a number of known draw-backs of the traditional two-view stereoscopic representation. That led to a number of projects investigating new depth-based representations, production methods and new types of displays.

Although some auto-stereoscopic displays have been introduced to the market, most are targeting specialized applications and are still not yet seen to be mature for the use in the consumer segment. Most restrictions are affecting the ability to watch 3D-TV program s in a quality that is equivalent to glasses-based two-view displays in a living-room environment with any number of viewers. Another aspect that might need more attention from future research is the fact that auto-stereoscopic displays provide a different viewing experience than two-view displays: Current state-of-the-art displays provide a limited depth range (usually of the order of 1m), which is centered on the (physical) display. Two-view displays can reproduce a much wider range of depth, and are in particular able to present objects at infinity (e.g. on the horizon). This has implications for the production or conversion of the content. For example, content that was produced for stereoscopic displays, such as cinema content has a certain production grammar. That means the content producers use certain concepts to "tell their story". An automatic conversion of stereoscopic content for a medium such as an auto-stereoscopic display which has very different characteristics might therefore not be easily achievable at a level that is acceptable to the creative content producers and consequently the end user. These problems also apply to conversion of existing 2D content to 3D.

Although depth-based representations have not been used as a delivery format in the recent implementation of first regular 3D-TV services, they remain very powerful tools for production and post-production. They provide far more flexibility to adjust parameters, e.g. the range of depth, than the conventional stereoscopic format. Furthermore, a depth-based 3D representation helps or is required in post-production to position additional graphical elements, like special effects or captions at the right place in 3D.

This paper has focused on the research perspective of 3D-TV. The broadcast industry has very recently implemented the first regular 3D-TV services based on two-view stereo. Many technological choices are been made pragmatically. It is likely that more research will be needed after this introductory phase, in the light of wider user feedback.

[27] http://www.prime3d.de/.

References

[1] R. B. Johnson and G. A. Jacobsen, "Advances in lenticular lens arrays for visual display," presented at the Current Developments in Lens Design and Optical Engineering VI, Proceedings of SPIE, San Diego, CA, Aug. 2005, Paper 5874-06.

[2] C. Wheatstone, "Contributions to the physiology of vision.—Part the first: On some remarkable, and hitherto unobserved, phenomena of vinocular vision," *Philosophical Trans. Royal Soc. London*, vol. 128, pp. 371–394, Received and Read June 21, 1838.

[3] D. Brewster, *The Stereoscope: It's History, Theory and Construction*. London: Hastings, 1856.

[4] P. Barr, "Flying false flags," *Scotsman*, Jun. 2002.

[5] R. Tiltman, "How "stereoscopic" television is shown," *Radio News*, Nov. 1928.
[6] M. Ziegler, "Digital stereoscopic imaging and applications. A way towards new dimensions," in *RACE II Project DISTIMA, IEE Colloq. Stereoscopic Television*, London, UK, 1992.
[7] C. Girdwood and P. Chiwy, "MIRAGE: An ACTS Project in Virtual Production and Stereoscopy," in *IBC Conf. Pub., No. 428*, Sep. 1996, pp. 155–160.
[8] "ACTS PANORAMA project," [Online]. Available: http://cordis.europa.eu/infowin/acts/rus/projects/ac092.htm
[9] M. Ziegler, L. Falkenhagen, R. ter Horst, and D. Kalivasd, "Evolution of stereoscopic and three-dimensional video," *Signal Process. Image Commun.*, vol. 14, no. 1–2, pp. 173–194, Nov. 6, 1998.
[10] H. Imaizumi and A. Luthra, "Stereoscopic Video Compression Standard "MPEG-2 Multi-view Profile"," in *Three-Dimensional Television, Video, and Display Technologies*, B. Javidi and F. Okano, Eds. New York, NY: Springer, 2002, pp. 169–181.
[11] A. Smolic and D. McCutchen, "3DAV exploration of video-based rendering technology in MPEG," *IEEE Trans. Circuits Syst. Video Technol.: Spec. Issue Immersive Telecommun.*, vol. 14, no. 3, pp. 348–356, Mar. 2004.
[12] A. Redert, M. Op de Beeck, C. Fehn, W. IJsselsteijn, M. Pollefeys, L. Van Gool, E. Ofek, I. Sexton, and P. Surman, "ATTEST—Advanced three-dimensional television systems technologies," in *Proc. 1st Int. Symp. 3D Data Process., Vis., Trans.*, Padova, Italy, Jun. 2002, pp. 313–319.
[13] C. Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV," in *Proc. SPIE Stereoscopic Displays Virtual Reality Syst. XI*, San Jose, CA, Jan. 2004, pp. 93–104.
[14] W. Matusik and H. Pfister, "3D-TV: A scalable system for real-time acquisition, transmission and autostereoscopic display of dynamic scenes," in *Proc. ACM SIGGRAPH*, Los Angeles, CA, Aug. 2004, pp. 814–824.
[15] 3D Image Processing, "3D real-time HDTV: An overview," White Paper, 2005 [Online]. Available: http://www.3d-ip.com
[16] R. Tanger, N. Atzpadin, M. Müller, C. Fehn, P. Kauff, and C. Herpel, "Trinocular depth acquisition," *SMPTE Motion Imaging J.*, vol. 116, no. 5/6, pp. 206–212, 2007.
[17] R. Tanger, N. Atzpadin, M. Müller, C. Fehn, P. Kauff, and C. Herpel, "Depth acquisition for post-production using trinocular camera systems and trifocal constraint," in *Proc. Int. Broadcast Conf.*, Amsterdam, The Netherlands, Sep. 2006, pp. 329–336.
[18] A. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems," *Proc. SPIE*, vol. 1915, pp. 36–48, Feb. 1993.
[19] B. Mendiburu, *3D Movie Making—Stereoscopic Digital Cinema from Script to Screen*. Burlington, MA: Focal, 2008.
[20] F. Zilly, M. Müller, and P. Kauff, "The stereoscopic analyzer—An image-based assistance tool for stereo shooting and 3D production," in *Proc. ICIP Spec. Sess. Image Process. 3D Cinema Prod.*, Hong Kong, Sep. 2010, pp. 26–29.
[21] R. Mohr, R. Buschmann, L. Falkenhagen, L. Van Gool, and R. Koch, "Cumuli, panorama, and Vanguard project overview," *3D Structure Multiple Images Large-Scale Environ., Lecture Notes Comput. Sc.*, vol. 1506/1998, pp. 1–13, 1998, 10.1007/3-540-49437-5_1.
[22] S. Jolly, M. Armstrong, and R. Salmon, "Three-dimensional television: A broadcaster's perspective," *Proc. SPIE*, vol. 7237, 2009.
[23] "Special issue on 3D-TV horizon: Contents, systems, and visual perception," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pt. 2.
[24] B. Barczak *et al.*, "Display-independent 3D-TV production and delivery using the layered depth video format," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pt. 2, submitted for publication.
[25] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Non-linear disparity mapping for stereoscopic 3D," *ACM Trans. Graphics (SIGGRAPH)*, Jul. 2010.
[26] P. Merkle, A. Smolic, K. Mueller, and T. Wiegand, "Efficient prediction structures for multi-view video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, Nov. 2007, Invited Paper.
[27] C. Fehn, P. Kauff, M. Op de Beeck, F. Ernst, W. Ijsselsteijn, M. Pollefeys, L. Vangool, E. Ofek, and I. Sexton, "An evolutionary and optimised approach on 3D-TV," in *Proc. IBC, Int. Broadcast Conv.*, Amsterdam, Netherlands, Sep. 2002.
[28] Philips 3D Solutions, "BlueBox," [Online]. Available: http://www.business-sites.philips.com/3dsolutions/3dcontentcreation-products/bluebox
[29] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Mueller, P. H. N. de With, and T. Wiegand, "The effects of multi-view depth video compression on multi-view rendering," *Signal Process.: Image Commun. (2008)* [Online]. Available: 10.1016/ j.image.2008.10.010
[30] "Text of ISO/IEC FDIS 23002-3 Representation of auxiliary video and supplemental information," Marrakech, Morocco, ISO/IEC JTC1/SC29/WG11, Jan. 2007, Doc. N8768.
[31] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," in *ACM SIGGRAPH and ACM Trans. Graphics*, Los Angeles, CA, USA, Aug. 2004.
[32] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Process.: Image Commun.. Spec. Issue 3DTV*, Feb. 2007.
[33] A. Smolic, K. Müller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "Intermediate view interpolation based on multi-view video plus depth for advanced 3D video systems," in *Proc. ICIP 2008, IEEE Int. Conf. Image Process.*, San Diego, CA, Oct. 2008.
[34] K. Mueller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "View synthesis for advanced 3D video systems," *EURASIP J. Image Video Process.*, vol. 2008, 10.1155/2008/438148.
[35] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "Reliability-based generation and view synthesis in layered depth video," in *Proc. IEEE Int. Workshop Multimedia Signal Process. (MMSP)*, Cairns, Australia, Oct. 2008.
[36] A. Smolic, K. Müller, P. Merkle, P. Kauff, and T. Wiegand, "An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution," in *Proc. Picture Coding Symp. (PCS)*, Chicago, IL, May 2009.
[37] *Vision on 3D Video*, ISO/IEC JTC1/SC29/WG11 N10357, Video and Requirements Group, Lausanne, CH, Feb. 2009.
[38] The Economic Impact of the UK Film Industry, "Oxford economics," Jun. 2010. [Online]. Available: http://www.ukfilmcouncil.org.uk/media/pdf/i/r/The_Economic_Impact_of_the_UK_Film_Industry_-_June_2010.pdf
[39] L. Blondé *et al.*, "A Virtual Studio for Live Broadcasting: the Mona Lisa Project," *IEEE Multimedia*, vol. 3, no. 2, pp. 18–29, Summer 1996.
[40] M. Price, J. Chandaria, O. Grau, G. A. Thomas, D. Chatting, J. Thorne, G. Milnthorpe, P. Woodward, L. Bull, E.-J. Ong, A. Hilton, J. Mitchelson, and J. Starck, "Real-time production and delivery of 3D media," in *Proc. Int. Broadcast. Conv.*, Amsterdam, Netherlands, Sep. 2002.
[41] O. Grau, R. Koch, F. Lavagetto, A. Sarti, S. Tubaro, and J. Woetzel, "The ORIGAMI Project: Advanced Tools for Creating and Mixing Real and Virtual Content in Film and TV Production," in *IEE Proc. Vis., Image Signal Process.*, 2005, vol. 152, pp. 454–469.
[42] O. Grau, G. A. Thomas, A. Hilton, J. Kilner, and J. Starck, "A robust free-viewpoint video system for sport scenes," in *Proc. 3DTV Conf.*, Kos, Greece, 2007.
[43] O. Grau, S. Minelly, and G. A. Thomas, "Applications of depth metadata," in *Conf. Proc. Int. Broadcast. Conv.*, Sept. 2001, vol. 2, pp. 62–70.
[44] "EBU 3D-TV study group," [Online]. Available: http://tech.ebu.ch/3dtv
[45] H. Kim, M. Sarim, T. Takai, J.-Y. Guillemaut, and A. Hilton, "Dynamic 3D scene reconstruction in outdoor environments," in *Proc. IEEE Symp. 3D Data Process. Vis.*, 2010.
[46] European Commission, "ICT in FP7," 2006 [Online]. Available: ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/fp7-ict-4poverview-v4.pdf
[47] "3DTV network of excellence," 2008 [Online]. Available: http://www.3dtv-research.org
[48] "3D. immersive, interactive media cluster web page," [Online]. Available: http://www.3dmedia-cluster.eu
[49] S. T. Worrall, A. M. Kondoz, D. Driesnack, M. Tekalp, P. Kovacs, Y. Lapid, H. Gokmen, and P. Aicroth, "DIOMEDES: Content aware delivery of 3D media using P2P and DVB-T2," *NEM Summit*, Sep. 2010.
[50] T. J. Naughton, C. Falldorf, L. Onural, P. Ferraro, C. Depeursinge, S. Krueger, Y. Emery, B. M. Hennelly, and M. Kujawinska, "Capture, processing, and display of real-world 3D objects using digital holography," in *9th Euro-American Workshop Inf. Opt. (WIO)*, Helsinki, Jul. 2010, pp. 1–3.
[51] "Super hi-vision (ultrahigh-definition wide-screen system with 4000 scanning lines)," [Online]. Available: http://www.nhk.or.jp/digital/en/super_hi/index.html

[52] "The PRIME project," [Online]. Available: http://www.prime3d.de/en.html
[53] "The Berliner Philharmoniker in 3D," [Online]. Available: http://www.digitalconcerthall.com/en/info/3d
[54] A. Smolic, P. Kauff, S. Knorr, A. Hornung, M. Kunter, M. Mueller, and M. Lang, "3D video post-production and processing," *Proc. IEEE*, vol. 99, no. 4, Apr. 2011.
[55] K. Mueller, P. Merkle, and T. Wiegand, "3D video representation using depth maps," *Proc. IEEE*, Apr. 2011, Invited Paper.
[56] Fraunhofer HHI, "Stereoscopic analyser (STAN)." [Online]. Available: http://www.hhi.fraunhofer.de/en/abteilungen-am-hhi/bildsignalverarbeitung/anwendungen/stan/
[57] Mebucom, "ITVA gold award for KUK film 3D trailer," (in German) [Online]. Available: http://www.mebucom.de/archiv/produktion/ITVA-GOLD-AWARD-f%C3%BCr-KUK-Film-3D-Trailer-340
[58] Nvidia, "Video with Nvidea 3D vision," [Online]. Available: http://www.nvidia.com/object/3d-vision-3d-movies.html
[59] HFF, "InSightOut," [Online]. Available: http://www.insightout-training.net/

Oliver Grau (M'10) received the Diploma (Master) and the Ph.D. from the University of Hanover, Germany.

From 1991-2000 he worked as a Research Scientist at the University of Hanover and was involved in several national and international projects, in the field of industrial image processing and 3D scene reconstruction for computer graphics applications. In 2000 he joined the BBC Research & Development Department in the UK. He was working on a number of national and international projects on 3D scene reconstruction and visualization. His research interests are in new innovative tools for visual media production using image processing, computer vision and computer graphic techniques and he published a number of research papers and patents on this topic.

Dr. Grau was and is active as reviewer for scientific journals, research bodies like EPSRC, EC-FP7 and as a program committee member of several international conferences. Further he was the initiator and chair of CVMP, the European Conference on Visual Media Production in London.

Thierry Borel received the electronics engineering degree from Ecole Supérieure d'Electronique de l'Ouest (ESEO) in Angers, France in 1986.

He started his career in Thomson in 1987 as a development engineer in Germany for 4.5 years working on first Thomson's LCD TV prototypes. In 1991, he moved back to France and worked on the first LCD RPTV and Plasma TV sets. He participated in collaboration with the David Sarnoff Research Institute to the elaboration of the first A-Si Integrated Drivers on projection LCD light valves. From 1996 to 1997, he participated in the development of the first automatic stereo camera of Thomson. From 1998 to 2005, he was appointed Display lab manager and conducted various projects related to flat display technologies. In 2005, he was appointed Signal Processing research lab manager and led various teams working on Color Correction for cinema, Color Management and Anti-Camcorder concept studies. He has coordinated 2 European research project (FP4-LCOS4LCOS, FP6-OSIRIS) in the display and 3DTV areas. From March 2007 to September 2008, he moved to Thomson/Technicolor Research lab in Beijing as Technology Development Manager. From July 2008, he is the leader of a research project inside Technicolor aiming at proposing end to end solutions for 3D TV and 3D Cinema workflows with and without glasses. He is now based in Rennes, France.

Mr Borel is member of SMPTE and DVB standardization organizations, a member of SID. He is also the chairman of a Steering Team in the 3D@Home consortium.

Peter Kauff, photograph and biography not available at the time of publication.

Aljoscha Smolic received the Dr.-Ing. degree in electrical and information engineering from Aachen University of Technology (RWTH), Germany, in 2001.

He joined Disney Research Zurich in 2009, where he leads the "Advanced Video Technology" group. Prior to that, he was Scientific Project Manager at Fraunhofer HHI in Berlin. He conducted research in various fields of video processing, video coding, computer vision and computer graphics and published more than 100 referred papers in these fields. Since 2003 he had teaching appointments at Technical University of Berlin, ETH Zurich, Universitat Politècnica de Catalunya (UPC), Universidad Politècnica de Madrid (UPM) and Universitat de les Illes Balears(UIB).

Dr. Smolic received the "Rudolf-Urtlel-Award" of the German Society for Technology in TV and Cinema (FKTG) for his dissertation in 2002. He is Area Editor for Signal Processing: Image Communication and served as Guest Editor for the PROCEEDINGS OF THE IEEE, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SIGNAL PROCESSING MAGAZINE and other scientific journals. He is Committee Member of several conferences, including ICIP, ICME, and EUSIPCO and served in several Chair positions of conferences. He chaired the MPEG ad hoc group on 3DAV pioneering standards for 3D video. In this context he also served as one of the Editors of the Multi-view Video Coding (MVC) standard.

Ralf Tanger (S'95–A'96–M'03), photograph and biography not available at the time of publication.