# TRENDS AND TECHNOLOGIES IN DIGITAL MEDIA

# CONTENT

# TRENDS AND TECHNOLOGIES IN DIGITAL MEDIA

3D is here to stay, both in movie theaters and in our living rooms. However, there is still a need for technologies that make 3D production simpler and cheaper while helping provide viewers with the perfect media experience, in both visual and auditory terms. This is why the industry is seeking to refine new 3D-related methods and technologies.

This includes constantly optimizing recording and production processes for stereo 3D, panoramic and multi-view video or 3D audio as well as high frame rate (HFR), high dynamic range (HDR) or the light field technology for more freedom in post-production.When it comes to transmission of content there is also a need of new coding methods such as HE-AAC or HEVC. These methods efficiently encode high resolution pictures and multi-channel audio at limited bandwidth for adaptive streaming, even if the number of users and devices is increasing.

This booklet focuses the most important digital media trends and methods which researchers of the Fraunhofer Digital Cinema Alliance are working on presently.

Dr. Siegfried Foessel
(Spokesman of the Fraunhofer
Digital Cinema Alliance)

# 3D AND BEYOND: THE ULTIMATE MEDIA EXPERIENCE

### 3D with no room for improvement?

Producing a high-quality feature film, particularly in 3D, is a time-consuming and expensive proposition. The obvious example is »Avatar«, which cost 237 million dollars and took four years to make. Productions with less than ample budgets have no way of achieving the same high standard and so are practically doomed to fail at the box office, which is one reason why there are so few successful high-quality 3D films. To be able to create practical, marketable technology solutions for the future, scientists at the four member institutes of the Fraunhofer Digital Cinema Alliance, pooling their experience and expertise in video and audio technology, have launched the »SpatialAV« development project. During this project an intelligent, modular, multi-sensory recording and production system for immersive audio-visual media will be developed, which allows a feeling like being part of the scene for the spectator.

### SpatialAV – experience the future of recording and production technology

The goal of the SpatialAV project is automating workflows, which supports high quality and helps cut technical production costs. Using intelligent, networked cameras, microphone arrays and additional sensors providing, for example, positional, inertial or depth data, the system should be able to automatically adjust recording parameters, supply additional supporting data for post-production and so be sufficiently versatile to be used in a wide range of production and reproduction scenarios. The plan is to use recorded data to derive visualization and auralization data optimized for a specific type of reproduction: 2D, 3D, free-viewpoint or panoramic video. The recording and pro-

## SPATIAL AV – A JOINT RESEARCH PROJECT OF FRAUNHOFER DIGITAL CINEMA ALLIANCE

– goal:developing an intelligent, modular, multi-sensory recording and production system for immersive audiovisual media
– launch date: February 1, 2012
– project duration: 33 months

**"The SpatialAV project aims to develop innovative technologies to explore new creative ways of making films. The idea is to allow camera operators to go back to concentrating on mise-en-scene and telling the story. We must avoid cluttering them with camera setting, lighting, sound and other kinds of technical data that have been swamping film production sets ever since the advent of 3D."**

*Dr. Siegfried Foessel,*
*Project Manager SpatialAV*

**Fraunhofer-Institutes involved in the project::**

Fraunhofer IIS
self-calibrating cameras, use of new recording methods based on light field and HDR technologies; sensor fusion; post-production and distribution formats

Fraunhofer HHI
system design and equipment for production of 3D and panoramic content; algorithms for processing and cleaning up data for reproduction

Fraunhofer IDMT
3D audio, object-based audio coding, methods and tools for audio-visual coherence, portable 3D audio reproduction equipment

Fraunhofer FOKUS
3D dome projection and methods for adapting content to panoramic or spherical projection screens

duction systems to be developed are to be suitable both for classic 2D and 3D reproduction on planar (e.g. theater and television) screens and for three-dimensional screens, as used in panoramic and dome theaters.

## Transition from manual to automated stereo 3D production

Current 3D recording techniques mean that a great deal of manual adjustment and editing is necessary to guarantee high-quality video and audio reproduction. This makes 3D production comparatively to 2D expensive and time-consuming.

In the first step, Fraunhofer scientists will use operational meta and calibration data such as depth information or 3D room acoustics to automatically generate information that can help cope with changed production conditions or to create virtual point-of-views for autostereoscopic displays.

It is no longer unusual for more than one camera to be used simultaneously for either 2D or 3D recording. Multi-camera systems are gaining importance in film and TV productions. Conventional stereo camera systems, omnidirectional cameras, multi-view video systems – each of these approaches was designed and optimized with one specific type of production in mind and consequently lacks flexibility. All are time-consuming and require a large number of staff to be present on set or in the studio. In addition, current systems are difficult or impossible to adapt to other production sets or changing conditions. The SpatialAV project aims to enable more flexible production scenarios by making it possible to adapt data for various purposes, so that similar or partly identical systems can be utilized in another context despite differences in recording conditions or camera and microphone placement. Besides stereo camera rigs, recording is increasingly being done using so-called multi-baseline stereo systems. A number of pilot projects are using complex systems involving, for instance, side-by-side

cameras or automatic sliders. As a camera records, additional data is automatically generated which can be used for depth, calibration and scene calculation. Based on one perspective, it is possible to calculate several views of a scene, which can be adjusted in post-production to fit the desired angle.

## Technologies for omnidirectional 3D

The impression of being physically in the middle of the action or looking up at a real starry sky is realized by using 2D scenes, and only for some special applications currently in 3D. However, demand is increasing for panoramic and dome theaters to provide a realistic 3D experience. Omnidirectional camera systems arranged in star or ball shapes are being used to film 3D scenes from various angles. Material shot in this way can be shown in special projection rooms and planetariums, as well as for creating moving panoramic backgrounds for virtual-set productions. For live sportscasts, viewers can be offered images of previous action that happened out of shot for the main cameras.

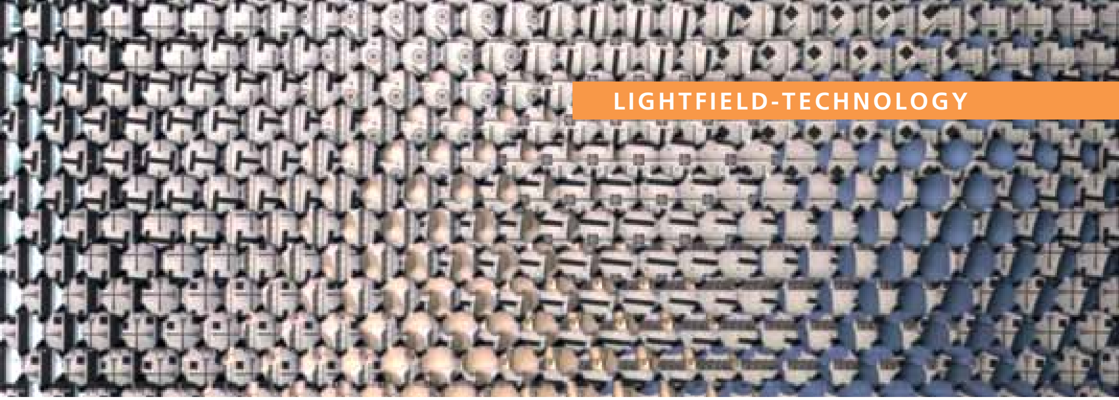
## The Future: Free-Viewpoint Video Systems and 3D Audio

In the film »Matrix«, the movements of protagonist Neo seem to be “freezed” until the scene continues from another viewing angle. In this case, these special effects that could be used only for one single point of time, should be used in motion with different speeds for real shots. This implies an intelligent multi-camera system shooting a central scene from several different angles. The director of photography then can move for example a virtual camera around the scene afterwards in the post production. Since such multi-camera systems are used for scene areas of varying size, each production traditionally uses a specially developed and calibrated system, which then works perfectly only in that context. This makes production extremely time-consuming, labor-intensive and hence expensive. SpatialAV aims to develop an intelligent, more resource-efficient method for recording data and preparing it so that it can be adapted for other scenarios. The result will be a modular recording and production set which can be easily configured and optimized for a wide range of applications.

**Quiet please – roll sound!**

Audio recording systems face the same challenges associated with video recording. It is common for large numbers of microphones to be individually placed close to each sound source, with the recorded sound manually and laboriously incorporated in post-production. Room acoustics are often recorded poorly or not at all, usually being added at the post-production stage. For this reason, a significant part of the audio development work undertaken in SpatialAV will relate to audio-visual coherence, i.e. the joint recording, blending and adjustment of the acoustic and visual perspectives.

The goal is the development of production tools and infrastructure for high-quality, spatially accurate sound recording and reproduction. Additionally, a portable 3D audio reproduction system is to be developed to provide production staff on set with a first impression of the 3D audio recorded. Research and development activities are scheduled to last three years, starting in 2012.





# LIGHTFIELD – TECHNOLOGY FOR MOVING PICTURES

In today's digital world, film sets – especially in the context of high-quality 3D production – entail an extensive digital post-production team working in the immediate vicinity of the filming location, performing calculations und calibrations, reviewing footage, editing lighting and sound, creating preview clips and so on.

In the early days of cinema, by contrast, apart from the cast and directing team, there were only the camera, microphone and lighting on a film set. It is worth returning to that kind of work environment. Not only do highly complex work environments adversely affect results, the fact that they require a great deal of material and human resources as well as time means a constant escalation of costs. Consequently, the industry is looking for solutions that involve only a reasonable number of equipment on set while offering a large amount of freedom with regard to camera and scene parameters.

Lightfield technology makes it possible to produce a variety of shots of one scene and create a so called lightfield using a camera array (see picture above) and then to adjust angle and depth of field at the post-production stage. If slight modifications become necessary in the course of the shoot, they can be easily achieved in post-production.

# BEHIND THE SCENES: AUDIO CODECS IN BROADCASTING AND STREAMING

*Interview with Harald Fuchs, Senior Business Development Manager TV Broadcasting at Fraunhofer IIS*

**When watching video on the Internet or listening to web radio, one often comes across the HE-AAC audio codec. What is behind this unwieldy acronym?**

HE-AAC stands for "High-Efficiency Advanced Audio Coding". This is an extremely efficient MPEG audio codec, which makes it possible to achieve high-quality stereo and multi-channel audio at very low bit rates. Virtually every end user comes into contact with HE-AAC on a daily basis – from web TV to Internet radio to DAB+ and Digital Radio Mondiale and to satellite radio. It is also used in TV broadcasting. It's part of the DVB toolbox and used in most of the countries, including the UK and Sweden, that recently switched to second-generation terrestrial TV. South American television broadcasters also use the codec to ensure good sound quality. HE-AAC is the most common format for multimedia streaming on the Internet. For this reason, most radio and TV sets, set-top boxes, PCs, and mobile devices support it.

**Why is HE-AAC so popular?**

Bandwidth is a valuable commodity in both broadcasting and online delivery of multimedia content. HE-AAC is extraordinarily efficient, helping broadcasters and service providers to save bandwidth and in turn money. This comes with no compromise in quality: Many independent tests have shown that HE-AAC provides very high audio quality even at very low bit rates. Practically all modern equipment supports this codec, making content accessible on almost any device.

**Do you expect HE-AAC to become even more widespread?**

Yes, of course. Take the new DASH streaming standard, where the codecs that are used to stream multimedia content over http are currently being specified. Many global companies such as Microsoft, Qualcomm or Adobe are involved in developing and marketing DASH. HE-AAC is ideally suited for the use with DASH, as it not only combines very low bit rates with very high quality, it is also able to seamlessly switch between bit rates. This combination of efficiency and flexibility is unique.

**Radio and TV broadcasters often have difficulty finding the right balance between dialogue and background noise. I understand that in the future it will be up to the audience to adjust the respective volume levels – how?**

We are working on such a technology, called Dialogue Enhancement. It allows listeners and viewers to create their own audio mix by adjusting the speech volume separately from the rest of the audio signal. Considering that the balance between dialogue and background noise is a major source of audience complaints, broadcasters are clearly in need of a technology like this. Incidentally, Dialogue Enhancement is extremely easy to combine with HE-AAC, because the two technologies are designed to be perfectly compatible.

**How far along are you in developing Dialogue Enhancement?**

We first tested the technology in an experiment with the BBC during the 2011 Wimbledon Tennis Championships. And Swedish Radio is launching a trial this fall. This will involve an app on mobile devices that allows listeners to assign separate volume levels to background noise and dialogue. We're still at the development stage, though. What we hope the trials will do is yield answers to questions such as "What's the simplest way of integrating the technology into the broadcast workflow?", "What kind of adjustment options are broadcasters and audiences most interested in?" and "How can we make the technology really easy to use?". Our aim is to see new radio and TV remote controls provide access to Dialogue Enhancement functionality in two or three years' time.

### What other trends are there in audio coding?

Traditionally, speech codecs and music codecs represent two completely separate approaches. As the name indicates, speech codecs are optimized for voice transmission. At very low bit rates, say 8 kbit/s per channel, they work reasonably well for speech, but music sounds very poor. By contrast, music codecs are not optimized for voice transmission at very low bit rates, but play music well. The Extended HE-AAC codec bridges the gap between speech and music. It handles both speech and music equally well, even at low bit rates. This is going to be particularly interesting for future radio systems where talk and music are both important, but the amount of bandwidth available is very limited. At IBC we'll be demonstrating how Extended HE-AAC can be used in broadcasting.

### And what are the main current trends in the audio industry?

One important trend is the movement towards higher flexibility for the end-user, with technologies like Dialogue Enhancement at the forefront. Another is obviously the continued reduction of bit rates, in other words, higher transmission efficiency.A further trend is 3D audio which aims to provide a more immersive spatial sound, mirroring the trend towards 3D video and ever larger screens with higher resolutions. For 3D audio, unlike the current 5.1 multi-channel standard, the speakers are not all positioned at the same level; rather, additional speakers are placed at different heights. As well as achieving more natural sound reproduction, this opens up possibilities for acoustic design, for example, to 3D movie scenes.

*Interview by Janine van Ackeren*

# NEW STANDARD HEVC ENCODES FILMS MORE EFFICIENTLY

The opening ceremonies for the Olympic Games captivated countless viewers all over the world in front of their televisions, astounding them with a gigantic show. Relatively few people were able to have a live experience of the spectacle at the London stadium. Still, some of the fans watching the show felt as if they were there live, even though they were only sitting in front of a large cinema screen. That's because a few movie theaters showed the opening ceremonies in 8K-resolution, which corresponds to 33 megapixels (see graphics page 17).

The resolution on home televisions will soon be enhanced even further, conveying the feeling of being right in the middle of the action. Indeed, the successor to the full HD television set is already penetrating the market: the 4K display, also called 2160p format. These televisions have four times as many pixels as the TVs in our living rooms today. Still, the continuously growing number of pixels must also be fed with the matching content, so that the capabilities of the high resolution television can also be utilized. But to do so has always been tied to immense costs, until recently, and therefore was only considered for major events, like the Olympic Games.

### Twice as efficient

The previous standard for encoding data and sending it from the broadcaster to the home television set is known as H.264/MPEG-4 AVC. Theoretically, it has certainly been up to handling the mass of data; however practically, the broadcasting of higher resolution entails substantial costs: because an additional channel is needed for television broadcasts, and for Internet transmission, the server needs a wider bandwidth. A majority of the repu-

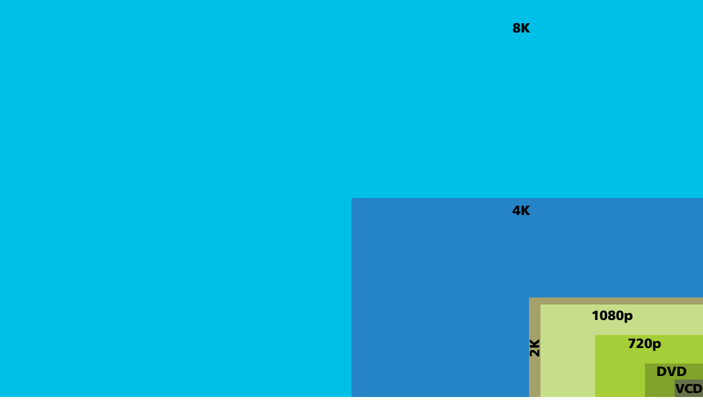
table electronics manufacturers have now joined forces to develop a new broadcasting standard together: HEVC, short for "High Efficiency Video Coding." The labs at the Fraunhofer Heinrich-Hertz-Institute HHI in Berlin, which played a critical role in the engineering of the H.264 predecessor standard, also made a substantial contribution to this new standard.

The advantage of HEVC:The standard requires half the bandwidth for high quality video transmission. But how is it being done? "Parts of H.264 were subsumed and optimized," explains Dr. Thomas Schierl, group manager of multimedia communications at HHI. "One example is the block size: whereas H.264 subdivides the transmission image into blocks of 16 by 16 pixels, HEVC instead carves the image into blocks of varying sizes with up to 64 by 64 pixels. These larger blocks can be encoded considerably more efficiently." If an object is seen in the image that moves to the side, then this movement occurs smoothly. The standards establish the movement data for each block – data that is ordinarily transmitted once per block. Because the blocks in HEVC are substantially larger than in H.264, correspondingly less movement data are needed.

## Parallel Processing

Compared to H.264, since the computational effort for the higher coding efficiency increases sharply to encode or decode the images, HEVC in the standard design allows computer units to work parallel with each other. Either the image is separated out into several parts, known as tiles, whereupon each processor works on one of them, or in the wave front method, where the processors each handle one block of lines in the image. These methods allow encoder manufacturers to get implementations and products to market rapidly.

The development is scheduled for completion in January 2013. Thereafter, new televisions, smartphones and PC units will presumably contain decoders that convert




data – encoded with HEVC – into high-resolution television images. The HEVC standard for 3D movies should follow in one to two years.

## HEVC for video telephony and video streaming as well

The new standard will deliver benefits to video telephony as well. It too, was hitherto largely based on H.264. With HEVC, the image quality can be increased substantially at the same data rate. Likewise, the transmission can be adapted for web video-streaming. MPEG-DASH, a transport format for multimedia streaming, currently enables viewers to watch judder-free videos via the Internet. Today it allows the transport of H.264-encoded contents as well as other standards. The researchers are planning to extend DASH by April 2013 in such a manner that it can also transmit HEVC-encoded videos.

**HEVC in short:**

– Television resolution is constantly improving – new 4k displays have four times as many pixels as the TVs today

– More pixel need more data – so this must go hand-in-hand with transmitting the data more efficiently

– Reputable electronics manufacturers have now joined forces with Fraunhofer researchers to develop a new broadcasting standard for video compression: HEVC

– The standard HEVC is twice as efficient as H.264, therefore an high quality film needs only half of the bandwidth.

– The development is scheduled for completion in January 2013

# HIGH DYNAMIC RANGE MAKES MOVIES LOOK MORE BRILLIANT

The charming view of picturesque houses along the Italian coastline is meant to look like a holiday video. The brilliance of the view on the movie is lacking because the beautiful white houses in the sun are overexposed and the small lanes are too dark. When adjusting the exposure time, you fix the minimum or maximum brightness. In the areas where the set is brighter or darker than the chosen brightness, the image information gets lost and pixels are either black or white.

For an optimized shooting High Dynamic Range (HDR) should be used, means a larger interval between the brightest and darkest pixel. Many cameras record several time-sequential images with different exposure times to create this effect. This varies the exposure and the different bright images are merged afterwards, creating a blurring effect. The blurring is created in moving pictures because the objects occur in different images at different positions.

The scientists of the Fraunhofer-Institute for Integrated Circuits IIS in Erlangen solved this problem by using only one image exposure time and a specific sensor filter. A large range of contrast and a high level of sharpness for the moving objects can be reached with this method. "We put a special filter with a non-regular sampling in front of the camera sensor which shades some of the pixels," explains Joachim Keinert from IIS. "Special algorithms were used to calculate an optimized exposed image."

The algorithm as well as the computer simulation can be demonstrated. Further tests with a prototype camera will follow soon.

# THE FUTURE FOR CINEMA: IMMERSIVE CINEMA AND STEREOSCOPY

The combination of immersive cinema and 3D stereoscopy (3DS) represent the future of cinema: it upgrades the virtual reality experience by offering multidirectional viewing with a corresponding depth feeling. It rises accordingly new technical challenges both in the creation and diffusion of its content. With its experience in setting up and controlling large displays with very high resolution and immersive installations such as digital domes, the Fraunhofer FOKUS competence center VISCOM can foresee the needs to help the creators.

## Challenges at production

The production workflow of digital domes and other immersive cinema mostly follows the production workflow of regular cinema. Especially when 3DS is added the content has to be entirely created by computer generated imagery, in order to offer a true 3DS experience in every direction. A virtual scene is developed and each frame is created by ray tracing (a technique in which an image is generated by tracing the path of light through pixels in an image plane) and image rendering. This is an optimization of rasterization which is the process of computing the mapping from scene geometry to pixels in an image plane. This has a cost both in hardware and time of processing because of the high image resolution and the level of detail in the scenes. Regarding regular cinema, 3DS is like looking through a window. To consider a single viewpoint is usually enough for the rendering. Therefore the strongest 3D effect is perceived by viewers seated in the center of the cinema room. A reasonable head movement is tolerated as long as the eyes stay approximately parallel to the horizon.

## Immersive cinema

In contrast, bringing 3DS into immersive cinema is a challenge. We can still assume that a main viewer is located in the center of the venue, surrounded by the screen, but now with a real freedom of head movement, the audience expects to perceive correct 3D in every direction. But since the eye positions of the spectators move and bank, the generation of correct omnidirectional 3D stereo images is not possible with the standard pinhole camera paradigm. To compute 3DS for immersive surround displays, innovative techniques like lens-shaders had to be developed. These shaders are used in ray-tracing, to simulate the head and eye movement in regards to the display surface. The new solution which appeared recently was to be able to specify these movements by means of simple maps. These new tools allow the creators to work more efficiently on the story-telling by spending less time on the technical aspect of the production. This is a real improvement in the production workflow of 3DS film for immersive cinema.

## Support in film production

Fraunhofer FOKUS is collaborating with content creators to optimize the content production, especially in 3D rendering for multidirectional viewing. Film director and owner of Softmachine GmbH Peter Popp in his recent new production ("Life of Trees") received technical support (consulting for 3DS rendering) from Fraunhofer FOKUS for the rendering of 3DS film for dome (computer power resources for rendering). Both partners take benefits from such collaboration: understanding and solving new technological challenges in stereoscopic dome rendering.

# 3D INNOVATION CENTER

3D does not only mean Hollywood Blockbusters. 3D also promote Germany in its traditionally strong economic sectors automotive and medicine. But how is 3D going to develop? What we needed was a place where a wide and varied range of actors with all their know-how can congregate, a place where technologies and innovations can be tested and presented and where we can chew over business models and marketing strategies together. Brain storming doesn't work alone.

The 3D innovation center offers a communication platform for providers and users plus a marketing instrument for advertising, sales and PR – in the shape of joint presentations at trade fairs and conferences & through joint mailings and overarching press and PR work.

Due to proactive collaboration in working groups the center also becomes a development platform and testbed for 3D technologies, applications and infrastructures. Experts can exchange their know-how and professional qualifications through workshops and seminars held in cooperation with universities and the industry.

**Contact**

3D Innovation Center
Kathleen Schröter
Tel. +49 30 31002-424
kathleen.schroeter@hhi.fraunhofer.de
www.3dinnovationcenter.de/en

# FRAUNHOFER DIGITAL CINEMA ALLIANCE

The Fraunhofer Digital Cinema Alliance is focused on solutions for an end-to-end digital media and digital cinema workflow in order to drive enhanced moving picture and audio experiences.

The Fraunhofer institutes of the alliance provide a network of intelligence with a pool of experts that provide future-oriented technologies and international standards to allow customers to stay ahead of the market.

Members are the following Fraunhofer Institutes for

- Digital Media Technologie IDMT
- Integrated Circuits IIS
- Telecommunications, Heinrich-Hertz-Institut HHI
- Open Communication Systems FOKUS

**Contact**

Fraunhofer Digital Cinema Alliance
Angela Raguse M.A.
Phone: +49 9131 776-5105
alliance-dc@iis.fraunhofer.de
www.dcinema.fraunhofer.de/en

**Editorial Notes**
Fraunhofer-Gesellschaft
Presse und Öffentlichkeitsarbeit
www.fraunhofer.de

**Editors**
Mandy Kühn, Angela Raguse

**Editorial Assistant**
Janine van Ackeren

**Photography**
Matthias Heyde, Matt Lancashire,
panthermedia, Piffl Medien,
istockphoto

**Illustration**
Vierthaler&Braun