Fraunhofer

TRENDS AND TECHNOLOGIES IN DIGITAL MEDIA

Edition 3

CONTENT

PREFACE

Digital media technology makes a lot of things possible that were unthinkable just a few years ago. 4K image definition and spatial sound even for portable devices and uninterrupted enjoyment of media – regardless of location, time, or end device – are prepared and are available for end users. 3D has long reached the living rooms of our homes. The trend continues towards ever-better definition, more dynamic color, and higher frame rates in order to increase picture quality even further.

As a trailblazer for future technologies in the area of digital media, the Fraunhofer Alliance Digital Cinema has stood for globally recognized standards, developments, and products for ten years. That's the reason our experts today are presenting the first systems and solutions of the next generation of media technologies. "Immersiveness" in the image and audio sector is a trend made possible by new means of production using multi-camera systems and special 3D audio technology. A single recording can now contain several views of an image or several sound objects. This opens up new creative opportunities in the editing stage that comes later. The first results from the Spatial-AV project, in which these technologies were developed and implemented, will be presented at IBC 2014. These include the options of navigating through 360° panoramas and selecting certain scenes interactively. Light field technologies allow film producers to avoid expensive retakes and easily integrate effects even after the fact. New 3D audio technology allows three-dimensional worlds of sound to be recorded and played back – or transmitted live straight from a concert hall. Standards such as MPEG-H, HEVC, and HE-AAC, in whose creation Fraunhofer institutes played a key role, open up a new world of sound and picture experiences for TV viewers.

As the spokesperson for the Fraunhofer Alliance Digital Cinema, I hope that you enjoy reading about and discovering new future technologies for the media sector.

Dr. Siegfried Foessel

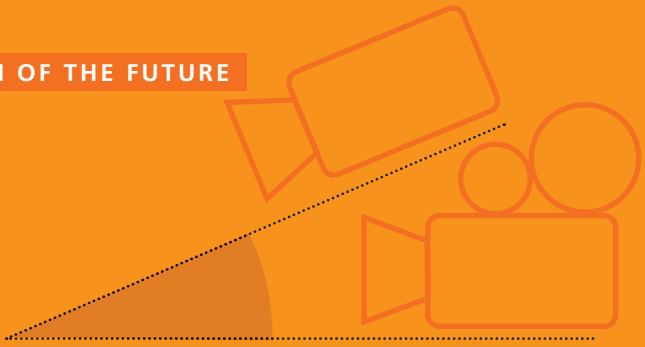

2014 FIFA WORLD CUP™ CHAMPIONS

SPATIAL MOVIE PRODUCTION – MORE CREATIVE AND BETTER VALUE

The days when the camera viewed a scene from a single position are long gone. Nowadays, special effects are in demand, and even more so if they're in 3D. Researchers in the Spatial AV project looked at how more creativity can be brought to three-dimensional cinema – and have developed technological solutions.

The camera revolves around the main character, who seems to be frozen in the middle of a jump – time seems to stand still for a moment and the camera shows the jumping figure from all sides. What two-dimensional movies can achieve can't be missing from three-dimensional ones. Of course, the 3D versions have the advantage of being able to pull the viewers in from their seats and whisk them away into an alternative fantasy world. But, the same principle applies: in order to make sure that movies bring in enough money, the makers have to constantly come up with new special effects. In the case of three-dimensional films, however, this drives up the production costs, which are already high to begin with. The best example of this is the movie Avatar, which, with a budget of $237 million, holds one of the top positions among the most expensive movies ever made.

"We want to create more opportunities for creativity – for both 2D and 3D productions," says Dr. Siegfried Foessel, head of department at the Fraunhofer Institute for Integrated Circuits IIS, "and that includes both picture and sound." The researchers at the Erlangen-based institute



have therefore joined with their colleagues at the Fraunhofer Institute for Digital Media Technology IDMT, the Heinrich-Hertz-Institut HHI, and the Fraunhofer Institute for Open Communication Systems FOKUS to look for the right solutions for the film industry. The project will run until the end of 2014 and the final presentation will take place on September at the IBC exhibition in Amsterdam Hall 8 B80.

3D productions made easy...

What makes 3D production so complex and therefore costly: instead of one, the cameraman must operate and focus two cameras. This is because the left and right eyes have a slightly different angle of vision – the two cameras imitate this effect. And, as if that wasn't stressful enough, the angle of inclination and the distance between the cameras must be adjusted constantly.

In the future, camera operators will no longer have to worry about things like that: it will be enough to focus one camera, and everything else will follow automatically. This is made possible by software developed by the researchers at IIS. "The second camera adopts the focus setting of the first one, and appropriate algorithms ensure that the cameras adjust to one another in an optimum manner," explains Foessel. There is already a prototype of the software. The cameras capture 25 frames per second and they recalibrate themselves automatically once per second. That's enough for most productions, but the IIS researchers are still currently working on further increasing the speed of calibration.

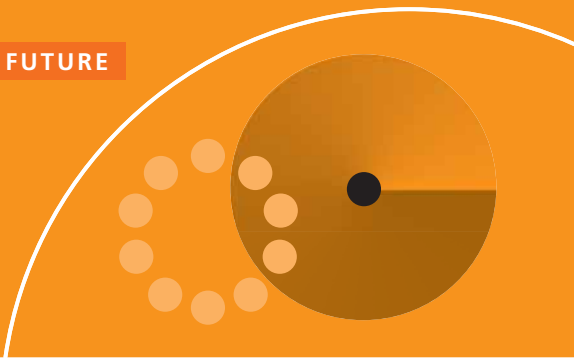
For more complex special effects, however, two cameras are no longer enough The IIS researchers have therefore set up a system comprising 16 cameras, which can be expanded even further as required. The trick to the system is in the software: the software uses the 16 camera images to generate depth maps that use gray values to specify how far the object in this pixel is from the viewer.

Virtual cameras can do more

"We can use this depth map to generate any number of views from the 16 camera views – meaning that we have created a virtual camera, similar to films that are entirely computer-generated. That gives us a great deal of freedom; we can produce moving shots without having to move the real cameras at all, for example," explains Foessel. That would not be possible with only two cameras: too many views needed for the moving shot would be missing. The software prototype is ready and pilot projects are already planned.

Sound will be three-dimensional

The sound tracks of movies are conventionally recorded and mixed in 5:1 format, meaning that the various sounds are distributed across different channels. Experts refer to this as the channel-based method. If you've bagged a good seat in the middle of the movie theater, you will experience spatial sound as intended. But in the seats at the side, the effect is less than perfect: in this case, you will be sitting nearer to one loudspeaker and the sound it produces will therefore seem considerably louder than the others.



Wave field synthesis, however – a method realized for the first time by the researchers at the Fraunhofer Institute for Digital Media Technology IDMT – ensures that everyone in the audience hears three-dimensional sound, no matter where they are sitting. "Wave field synthesis is not channel-based, but object-based," explains Dr. Sandra Brix, head of department at IDMT. This means that individual noises, voices, or instruments are recorded as objects in their own right and can be placed as appropriate into the sound scene. This would allow a plane to acoustically "fly over" the audience, for example. To generate this spatial sound impression with individually audible sound objects, a large number of loudspeakers produce an acoustic wave front. This spreads throughout the entire space as a sound field, much like the ripples caused by a stone being dropped in water.

If, for example, the researchers wish to record a classical concert, they first capture the sounds produced by individual instruments and then put them back together to form a scene. "We developed a software tool with an intuitive user interface to check whether all these microphones work," says Brix. This allows recording engineers to move a virtual "listener" through the orchestra space during recording to check the sound impression created in different areas. Editing of recordings has also been simplified by the researchers: special software allows recording engineers to perform rough mixing of the three-dimensional sound scene using headphones – without the need for the loudspeaker set-up that will ultimately be required. They then only need to go to the studio for fine tuning.

Changes, however, are not only coming to movie theaters; how we experience movies and television on our own sofas is also going to be transformed. Soccer and concert fans, for example, will be able to freely select the camera perspective, turn full circle, and enjoy a panoramic view of the field and the stands during live broadcasts.

The viewer as a "cameraman"

"OmniCam360" makes it possible: when this camera is placed at the edge of the field near the center line, it is able to capture a 360° panoramic view of the whole stadium. "The camera only weighs 15 kg and is no larger than a normal television camera. This enables it to be carried by one person and fixed to a tripod," says Christian Weissig, project manager at the Fraunhofer Heinrich-Hertz-Institut HHI in Berlin, where the camera was developed. The Omnicam comprises a total of ten cameras, but there is no need for complex calibration. All you need to remember when operating the Omnicam is: unpack the camera, plug it in, and start filming. The camera has already been able to prove its capability in a range of test productions, including with the Vienna Philharmonic. The camera is now licensed and is being marketed.

If, however, you would like to produce film material for a dome-shaped screen, additional cameras are needed to point skywards – otherwise a panoramic view will be projected around the edge but there will be a gaping black hole on the ceiling. That's why the researchers at Fraunhofer FOKUS have developed a special process to merge image streams from individual cameras into a seamless picture in real time. This means that even dome-shaped movie theaters will be able to show live broadcasts in the future.

WORKING WITH LIGHT FIELD DATA

An unforgettable sunrise, a complicated stunt, or the unattainable emotion of the actors in the scene is unique and cannot be repeated. It's unforgivable so, if that's the moment when the sharpness wasn't set exactly or the right angle hadn't been chosen. Any number of things might happen at any second to prevent the right impression being captured on film.

One new technology that could be used here in the future is filming using multi-camera systems, known as light field technology. This means that a single recording contains several different views, which are then used in post-production to recover the creative opportunities that sometimes seem to go missing on set.

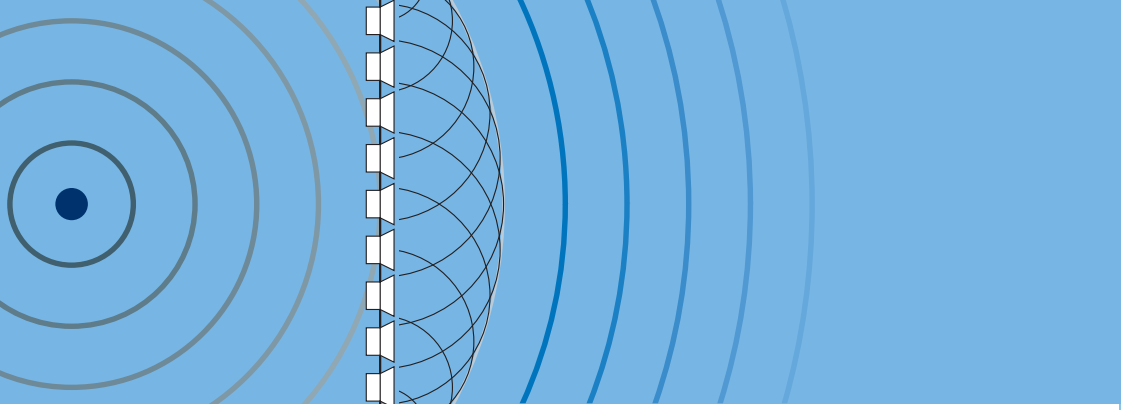
The processing of light field data, or data from multi-camera systems, is a new and promising trend in film production. This editing flexibility during post-production makes complex and expensive retakes or additional filming a thing of the past. Intelligent allocation of different views allows refocussing, changes in perspective, 3D effects, and changes in depth, as well as virtual camera movements in all spatial directions to be calculated. This allows as many additional views as necessary to be generated from the existing views.

Higher Flexibility for Post-production

Experts at Fraunhofer IIS are currently working on putting this light field processing into practice in a useful way. To this end, they are developing an auto-calibration method that can correct geometric distortions in the camera position (rectification) without test charts, calibration

patterns, or special markers. Depth maps with high pixel density are calculated based on the scene. These depth maps can be used to generate new high-quality views for 2D, 3D, or multi-view displays.

In practice, this means that a plug-in added to post-production software would allow a change in sharpness to be made without the need for a new recording, say during a voice-over scene or a dialog situation. The technology is currently in the first preliminary phase of implementation and testing, but offers great potential for the design of new, cost-efficient working processes in recording and production technology.



THE PERFECT SOUND EXPERIENCE

With SpatialSound Stage, SpatialSound Wave and SpatialSound Control, Fraunhofer IDMT offers new solutions for directional and three-dimensional sound emission for stages, live shows, planetariums, clubs, and theme parks.

Thanks to the SpatialSound Stage system, spectators can immerse themselves completely in what's happening on stage. Regardless the stage size the viewers can follow movements of the actors acoustically. Consistent sound quality and the correct spatial image are guaranteed for every seat. The operators of the world's largest lake stage in Bregenz, use this technology.

The SpatialSound Wave concept is designed for compact and efficient spatial sound reproduction in mobile or fixed installations. The freely configurable system works with any arbitrary loudspeaker arrangements - still guaranteeing a realistic acoustic impression. This can be achieved with sound objects which can be freely placed around the listening area, allowing a sound engineer to individually design an acoustical scene with plenty of effects.

The system impresses visitors of the Hamburg Planetarium during the planetarium's dramatic 3D shows.

The intuitive SpatialSound Control user interface complements the SpatialSound technologies. It allows simple production of audio material while using resources efficiently. The system offers sound engineers integrated apps for production and demonstration purposes and for live events. Thanks to the web-based multi-client solution, the software can be used by several people at the same time and can be controlled via various devices, such as PCs, tablets, or smartphones. In the area of live production, sound engineers can save a lot of time thanks to the straightforward configuration of the system.


LS+
DA
EQ

EASYDCP GOES IMF

With over 1,000 licensees around the world, the post-production software easyDCP has now secured its place on the market for the creation, playback, and encryption of digital cinema packages. The trend towards integration of the easyDCP functionality spectrum in post-production software offered by other companies is increasing; these include Blackmagic Design, Quantel, and Drastic – companies that want to be able to offer their customers a convenient addition to DCP generation via a plug-in. This is all made possible by a software API (application programming interface) from the Fraunhofer team, which offers direct integration into existing working environments via appropriate plug-ins.

After the first version of the new IMF format (Interoperable Master Format) has run through SMPTE standardization, this format will also be added to the scope of easyDCP's functions. IMF is a file-based solution to store high-quality image and audio data, as well as subtitles, in a uniform format – IMP (Interoperable Master Package). IMP is suitable for exchanging film material independently of manufacturers and devices, and is intended to save both time and money by using file-based work steps only. IMP is used as a master/source package for generating a wide range of various distribution formats. The IMF is intended particularly for data exchange between various post-production firms and for optimization of internal exchange.

easyDCP 3.0 will be presented at the IBC 2014. In addition to improved user guidance with enhanced drag & drop, this new version comprises support for the IMF format and significant advances in creating and checking subtitles.

“The software from Fraunhofer IIS has been embraced by the sector as one of the most far-reaching solutions in the production of digital film copies.”

Quantel, manufacturer of digital production equipment

easyDCP at a glance

- Software-based only and optimized to ensure that standard hardware can be used
- Simple and convenient to operate, allowing small and medium-sized production companies to generate DCPs that can be played around the world
- Contains tools for creating, playing, and encrypting DCPs
- Conforms to standards applicable to international specifications for digital cinema
- Platform-independent and reliable in the creation of DCPs
- Stand-alone or integrated as a plug-in in the post-production tools offered by many manufacturers
- Also used by film festivals such as the Berlinale
- Marketed around the world by distribution partner easyDCP GmbH

NEW CAMERA SYSTEM SIMPLIFIES 3D FILMING

The viewer climbs huge trees with the main character, creeps through fairytale forests, and dashes through deep ravines on the backs of dragons – and sees everything in three dimensions, just as if he or she were really there. However, this 3D technology that allows the viewer to be part of the action, has its price. The 3D film Avatar, for example, apparently had production costs of 300 to 500 million US dollars, making it one of the most expensive films of all time – even if these costs were not due entirely to the 3D technology, of course.

A camera for each eye

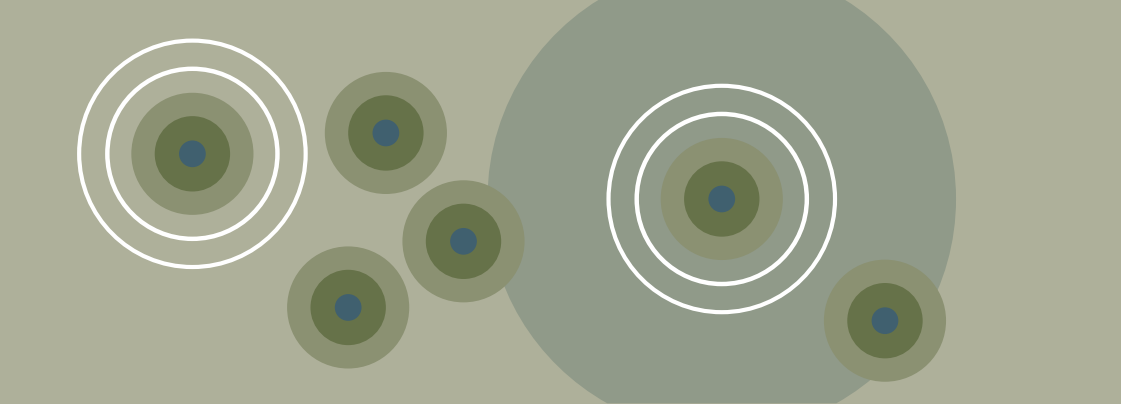
All the same, three-dimensional films have previously meant a great deal of time and effort for film makers, and therefore high costs. On the other hand, it's now hard to imagine the movie industry without 3D pictures. "That's why we are working on reducing the time and effort required," says Ralf Tanger, project manager at Fraunhofer Heinrich-Hertz-Institut HHI. Previously, all scenes have had to be filmed with two cameras in order to provide the left and right eyes with different viewing angles. The distance between the cameras depends not only on the size of the screen on which the movie will later be shown, but also on the depth range of the scene. This means that the film team needs to adapt this distance constantly.

In the future, cameramen will be able to save themselves this time and effort. "We have developed a trifocal camera system that allows the camera spacing to be adapted during editing rather than on the set," explains Tanger. "The system comprises a main camera and two satellite cameras mounted at a fixed distance

from the main camera. This means that work on set is almost like during a 2D production." Software used during editing takes the images from this camera system and calculates depth maps – images that use various gray tones to show how far the object at a certain pixel is from the viewer. The software uses these depth maps to generate the two camera perspectives needed for the optimum 3D image, whether it's for the screen in a movie theater or for TV.

Feasibility test at Disney

But does the system really meet the expectations in practice? "We can definitely answer that question with a yes," Tanger is pleased to say. The movie giant Disney has already tested the system – successfully. The result is the short film Make Believe, which Disney presented for the first time at NAB in April 2014. This film will also be presented at IBC from Disney on the big screen on Saturday, September 13th, from 4:30 - 5:30 p.m.



ENTERING NEW SPHERES WITH ATMOSPHEA

Since April 2014, visitors to Hamburg Planetarium have been able to witness for themselves how impressive the interplay is between 3D images and sounds. The ATMOSPHEA sound system is responsible for the new 3D audio experience. The basis of the ATMOSPHEA system is the *SpatialSound Wave* technology developed by Fraunhofer IDMT. The new sound system came about in cooperation between Fraunhofer IDMT and Shure Europe GmbH. For this innovative audio solution, Shure Europe GmbH licensed the SpatialSound Wave technology from IDMT under the name ATMOSPHEA.

ATMOSPHEA makes Hamburg Planetarium the first in the world to offer this unique combination of image content and sound sources that can be positioned freely around the room. 60 loudspeakers and four subwoofers were installed behind the 21 m-wide dome of the planetarium. Individual control of each speaker allows up to 32 audio sources to be located and activated freely within the dome. This allows the ideal sound to be produced for every seat. The audio content and acoustic environments are heard by every visitor with the right perspective; the best sound is not limited to preferred seats.

But it's not just at Hamburg Planetarium that visitors can look forward to the new sound quality. Jena Planetarium and Kiel Planetarium, have also invested in the 3D sound technology of Fraunhofer IDMT.

"The combination of image content and the new sound system offers completely new kinds of design options and experiences for immersive media content. We are looking forward to using them. This will be enormously attractive to artists and producers from the area of audiovisual media." *Thomas Kraupe, director of Hamburg Planetarium*

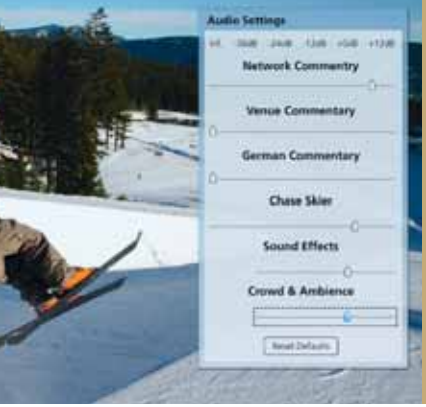
TELEVISION AUDIO FOR THE 21ST CENTURY VIEWER

Today's consumers enjoy the convenient, interactive, and personalized experiences offered by new media services on a variety of devices, from home theater TVs to tablets and mobile phones. Fraunhofer's new TV audio system, based on the MPEG-H audio standard, will allow TV broadcasters to offer a similar audio experience on all the devices, while allowing viewers to tailor that experience to their liking.

Fraunhofer's system utilizes MPEG-H's object coding to provide interactivity, which allows viewers to adjust the elements of the sound mix to their personal preferences, boosting hard to understand dialogue or creating a "home team" mix of sports broadcasts. Although it is based on different technology, this system builds upon Fraunhofer's pioneering trials with the BBC on Dialogue Enhancement.

This system also uses MPEG-H Audio's ability to transmit immersive sound with additional front and rear height speaker channels, improving upon today's surround sound broadcasts by providing a truly realistic and immersive audio experience on par with the latest cinema sound systems.

Based on the evolution of the AAC codec family as well as new technologies, this new system's audio codec offers unsurpassed coding efficiency allowing the additional channels or objects needed for interactive or immersive sound to be transmitted at bit rates used today for 5.1 surround broadcasts.



3D Sound in the living room

Fraunhofer's new TV audio system builds upon fifteen years of TV audio experience in providing half the world's TV surround sound. Broadcasters such as the BBC, NHK, and TV Globo rely on the AAC codec family to provide surround sound to their viewers. Fraunhofer has designed this new system to be compatible with the existing systems and practices used today for AC-3 or HE-AAC broadcasting and to provide an easy, staged implementation of new features, allowing broadcasters to deploy them when they choose.

While home theater enthusiasts are likely to install nine or eleven separate speakers to hear the ultimate 3D sound quality from the new system, today's convenience-oriented consumers will need a more accessible way to hear immersive sound. Fraunhofer has developed a conceptual prototype of a soundbar that delivers 3D audio without the wires, multiple external speakers or complex equipment settings needed for traditional speaker-based 3D sound that could be incorporated into future TVs or soundbars.

RAZOR-SHARP TV PICTURE: A FIRST FOR THE BUNDESLIGA

FC Bayern Munich beat SV Werder Bremen 5:2 in the Allianz-Arena on April 26, 2014 – so a pretty run-of-the-mill Bundesliga game, you might think. But you would be wrong: At least as far as the transmission of the game goes, it was a world first. The broadcaster Sky transmitted the game – live – in 4K definition, which means there were four times as many pixels as previously found in HD. For this, Sky used the new HEVC video compression standard that was developed by renowned electronic manufacturers together with researchers from the Fraunhofer Heinrich-Hertz-Institut HHI and made public about a year ago. The advantage of HEVC, which stands for "High Efficiency Video Coding" is that it only needs half the bitrate of its predecessor H.264 to provide the same picture quality – allowing it to transmit twice as much data on the same channel. That means that HEVC is practically made for the ultra-high definition of 4K television, which is where H.264 begins to hit its boundaries.

First ever live coding of 4K images with HEVC

The coding of the 4K camera images from the stadium was carried out by the HHI researchers together with their colleagues at Rohde & Schwarz. This means that they developed the software and hardware that convert the camera images into a coded HEVC bitstream, which is then transmitted to televisions via satellite. And what was so special about that? "For the first time, we were able to encode recordings live, meaning in real time," Benjamin Bross, who manages the HEVC project at the HHI, was pleased to explain. This type of real-time encoding contains a lot of challenges. While the

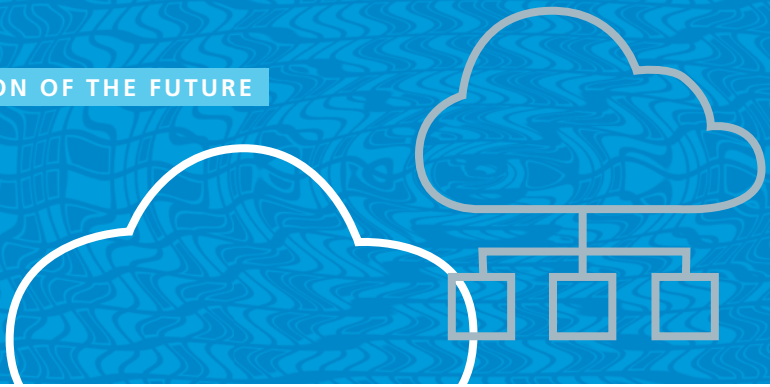
previous standard, H.264, divides the image to be transmitted into blocks of 16 x 16 pixels, HEVC subdivides using variable block sizes. The encoder must therefore decide which block sizes make the most sense on an image-by-image basis. The researchers will demonstrate the new encoder at the IBC trade fair – live, of course!

Down to the smallest detail…

So far, so good. But what does it benefit the viewer if he or she can view TV images in 4K and via HEVC coding in the future? The greatest advantage must be the razor-sharp picture in which you can make out every small detail. On a football pitch, for example, one camera can capture the entire pitch. This gives the viewer at home a strategic overview, while still allowing him or her to see every detail.

"HEVC offers a lot of possibilities. It's now time to get creative and take advantage of these possibilities," says Bross.

4K television sets are already on the market, but for the time being they are far from cheap. What's more, the benefits are still limited: there are still only Internet streaming services and no television channel in 4K. By the end of next year, estimates Bross, that might all by quite different – at least in the case of pioneers such as Sky. And, in a few years, 4K televisions might be the norm in our living rooms.



BETTER STREAMING

The FAMIUM development platform allows for new applications in adaptive video streaming: The framework can be used to create media content in DASH format and to transmit it in high quality over the Internet. One special highlight: the content can be distributed to several devices. You can also add advertising or integrate Digital Rights Management (DRM) in order to encrypt certain content for, say, a paying audience only.

Be it a TV show or a live concert recording – it's becoming more and more common for multimedia content to be offered as a stream and played in real time in a browser. The problem, however, is that the picture starts to jerk if the bitrate deteriorates. This can be prevented if the video quality adapts automatically to the available transmission rate. This kind of adaptive streaming is made possible by the ISO standard MPEG DASH (Dynamic Adaptive Streaming over HTTP). DASH allows videos to be modified with regard to definition, codec, or bitrate to ensure that they are suitable for the user's device and network connection, and that they play without jerking.

Player Inside

The FAMIUM framework allows various multimedia content to be created in DASH format. The software was developed by scientists at the Fraunhofer Institute for Open Communication Systems FOKUS in Berlin. "An additional part of FAMIUM is its player," explains Stefan Kaiser, scientist at FOKUS. "This allows movies and live streams to be played directly in the browser – no additional plug-in is required."

The multiscreen framework that allows content to be distributed to various devices is another component of the plat-

form. "We implemented a scenario to allow us to watch a Formula 1 race on two screens. While the TV is showing the normal race, a second screen – such as a tablet – provides perspectives from the cockpit," explains Stefan Kaiser.

A further feature allows us to include advertising in the multimedia content. The advertising can be either static or dynamic, which means that it is integrated spontaneously during playback. Pre-determined times can be entered to define when the advertising is shown. The researchers will demonstrate how this works at the International Broadcasting Convention IBC in Amsterdam in the fall.

They will also show how FAMIUM can be used to encrypt content for, say, a paying audience. Various Digital Rights Management (DRM) systems can be applied to the media content. DRM integration is particularly important when it comes to the player. FAMIUM also allows protected media to be played in browsers without the necessity of a plug-in. To this end, the researchers use the two W3C specifications Media Source Extensions (MSE) (Working Draft) and Encrypted Media Extensions (EME) (Candidate Recommendation).

A/V ANALYSIS: QUALITY CONTROL, IDENTIFICATION, AND FORENSICS

Songs, speech, films, or videos – broadcasters, video producers, and digital archives work with huge quantities of audiovisual data. Without special tools, using, searching, and managing this content is often difficult. The QC components from Fraunhofer IDMT help to investigate digital picture and sound recordings as well as generate and analyze and various types of metadata to be harvested automatically. For example, interference and quality problems can be recognized, global movements in videos such as panning, tilting, zooming or camera shake can be detected (video motion detection), and speech or music can be differentiated.

An additional field of application is segment-based identification of video material. Many TV broadcasters rely on film material from agencies for their reports, and change it with fading or by altering the sequence of images and sounds. This makes it difficult to look for contributions in archives or on the Internet. Support can be provided by processes that can identify video segments that come from the same recording or source. To this end, fingerprints are extracted from the videos and compared with one another in a multi-stage automatic process. This can be used, for example, to find out which reports used the same material and which TV channels reported on a certain topic. This makes it easier to track information regarding rights and origins, avoids unnecessary retention of duplicates, and allows the integrity of segments to be monitored.

Processes such as this are now to be complemented with newly developed audio forensic tools that detect the use of different microphones in recordings, can locate cuts in audio material, and can de-

tect previous coding steps and the parameters thereof. These processes can improve segment-based video identification by taking information from the audio track into account. Furthermore, the technical quality check can be expanded by, say, automatically detecting codec and bit rate of earlier coding steps and thus avoiding the use of lesser-quality material or unintentional duplicate coding. These processes can still be used to annotate steps and processing parameters automatically during production.

Due to a lack of time, editors often include incomplete or erroneous annotations, even though they are indispensable for later searching and for managing large quantities of data, such as the devices used for the recording. Most of all, however, the audio forensic tools described here can also help to check whether audio clips that are thought to be authentic weren't really edited after the recording in order to subvert their meaning. This may be of great importance when using user-generated content for journalistic research, for example.

"Our A/V Analyzing Toolbox is a comprehensive collection of analytical components that can be used to optimize the content management and production processes in the areas of broadcasting, digital archives, and the distribution of digital goods."

Dr. Uwe Kühhirt, head of department for metadata

WE ALL HEAR DIFFERENTLY

Everyone has their own sound preferences and their own volume sensitivity. Furthermore, our hearing begins to deteriorate from the age of 50. In Europe, around 70 million people live with a hearing impairment. Of these, fewer than 25 percent use a hearing aid[1]. This results in difficulties understanding speech. Given this background, one of the focuses of the research carried out by the Oldenburg-based Project Group Hearing, Speech and Audio Technology of Fraunhofer IDMT is on integrating hearing aid technology processes into telecommunication and consumer electronics.

The scientists develop signal processing methods and operating interfaces that allow users to adapt the audio signal from TV sets, radios, MP3 players, the Internet, or phones to their individual hearing needs. A hearing test from an acoustician is not necessary, unlike when a hearing aid needs to be adapted. Users can select from several presets to find the sound quality and volume that is comfortable for them. Adaptive signal processing than automatically optimizes every audio signal to this hearing profile. Speech signals that were previously too quiet, too loud, booming, subdued, or shrill are played as an even sound pattern with improved comprehensibility. Initial studies have shown that people with low and medium hearing loss can understand speech signals better with this system. Sound adjustments, however, are also made by those with normal hearing – with large differences between individual preferences.

Together with end users, the scientists optimize the applied algorithms and also carry out research in defined application scenarios to see which operational interfaces are suitable for audio signal adjustment – from classic control knobs to two-dimensional graphics on touch pads and voice control.

[1] *See: Heger, D.& Holube, I., 2010, Zeitschrift für Audiologie, 49(2); Kochkin, S., 2009, Hearing Review,16(11).*

10 YEARS OF RESEARCH AND DEVELOPMENT FOR THE FUTURE OF CINEMA

The greatest challenge facing cinema at the beginning of this century was the introduction and use of digital media. In the year the Fraunhofer Digital Cinema Alliance was founded, 2004, a lot of areas of media had not yet been completely converted to digital – cinema, as the last bastion of analog film technology, was at risk of being overtaken, and indeed left behind, by Internet-based and other digital formats that were pushing their way onto the market. Computer animations and special effects like those used in movies by pioneers such as George Lucas were expensive and time-consuming exceptions. At the very beginning, cameras that could satisfy the high standards of big-screen technology digitally were not available, and neither was the technology for additional editing that would allow seamless digital work up to the final product without having to sacrifice quality when compared to 35 mm film.

This was a challenge accepted by the four Fraunhofer institutes IIS, HHI, IDMT, and FOKUS (previously FIRST) when they came together to form the Fraunhofer Digital

2002
Project: Technologies for Digital Cinema

2004
Foundation of the Fraunhofer Alliance Digital Cinema

2005
DCI specifications

2006
CineVision 2006

2007
Worldscreen: Layered Scheme Compression for Digital Cinema and Cross Media Conversion

Publication of the DCI specifications

Cinema Alliance. With a central contact point, the institutes offer new technologies and developments for the film industry, productions, and filmmakers.

The institutes made a name for themselves in the industry through developments such as the ARRI-D20/21 – one of the first digital film style cameras in the world. This was followed by the creation of the technical specifications for digital cinema for Hollywood studios, the DCI (Digital Cinema Initiatives) test plan, which is now applicable around the world.

The 3D issue was supported by assistance systems for the cameraman, the STAN, as well as real-time conversion processes for 2D and 3D.

For post-production, the scientists developed new software for generating, playing, and encrypting digital cinema package DCPs, which is now used by more than 1000 firms around the world, including many well-known names from the movie industry. The institutes offer new systems for sound, such as those for 3D sound and wave field, and are opening up new worlds of sound and experience to viewers in movie theaters, planetariums, and at home.

2008
System specifications for Digital Cinema in Germany

2011
Foundation of the 3D Innovation Center

2012
Partner of the Berlinale for input checking of DCPs

2014
Spatial AV: Development of a smart, modular, multi-sensor recording and production system for immersive audiovisual media

FOCUS AND RESEARCH TOPICS MOVIE AND MEDIA WORKFLOW

ACQUISITION

Image and Sound Technologies
- Smart Cameras
- Stereoscopic 3D Systems
- Multi-camera Systems
- Camera Arrays
- High Dynamic Range HDR

- Microphone Arrays
- Spatial Sound Systems

Coding
- JPEG 2000
- Low-complexity
- Object-based Audio

POST-PRODUCTION

Mastering
- DCP Digital Cinema Package
- IMF Interoperable Master Format
- Archive Formats

Scalable Video Coding
- JPEG 2000
- Production Streaming Concepts

Processing & Effects
- Multi-camera Data Processing
- Light-field Processing
- 2D/3D Conversion
- Depth Maps Audio/Video

Room Simulation
Multi-Platform Transcoding
Content Analysis
Digital Archives

META DATA MANAGEMENT
QUALITY CONTROL
SECURITY

DISTRIBUTION

Image/Video Coding
- JPEG 2000
- H.264/MPEG-AVC
- H.265/MPEG-HEVC
- MPEG-DASH

Audio Coding
- MP3
- AAC, HE-AAC, xHE-AAC
- MPEG-H 3D Audio

EXHIBITION

Display & Projection
- 3D Projection
- Dome Projection
- 3D Panorama
- 360 degree, Navigation Functionality inside Scenes

Sound
- Spatial Audio Systems
- Immersive, Personalised and Interactive Sound Reproduction

Devices
- Flat Loudspeakers
- 3D to Multi-view Conversion Autostereoscopic Displays

THE FUTURE IS STANDARDIZATION

Creating uniform worldwide standards is a basic requirement for guaranteeing high-quality recording and transmission of pictures and sound. The Fraunhofer institutes in the alliance are therefore active on all important standardization committees, together with representatives from industry, in order to ensure that this requirement is met for new technologies and future generations of end devices.

JPEG 2000 – the recognized picture format for cinema and broadcast

The JPEG 2000 formats previously familiar from digital cinema are being complemented with new profiles for mastering broadcast and Internet-based films. At the same time, profiles are being developed in JPEG2000 that are suitable as a long-term format for film archiving. This will give JPEG2000 a complete set of profiles for professional editing of films, from contribution and mastering to the distribution of films to theaters and for long-term archiving.

Another growing trend is the increase in the dynamic range of the pictures. In order to be able to store and transport these pictures adequately, experts in the Joint Photographic Experts Group and the Moving Picture Experts Group are working on coding higher bit depths – from 12 to 16 bit integer representation for device-specific picture data to scene-referenced half-float representations of picture data.

Data protection is also playing a more and more important role, particularly in Europe. For this reason, efforts are under way to protect private data even within JPEG pictures; this may refer to areas of the picture or to metadata. Great store is set by reverse compatibility, i.e. existing reproduction systems can only see unpro-

tected data, while new tools may also be able to access protected data.

Additional activities that are only just beginning to be developed refer to the coding of free viewpoint scenes and the conflation of real and virtual scenes or data in the augmented reality sector. Additional information about objects detected in the vicinity play an increasingly important role.

HEVC – video compression for next-generation broadcasting

Based on developments by Fraunhofer HHI, the new standard for video compression, HEVC – High-Efficiency Video Coding, known as H.265 or MPEG-H, was completed and made public a year ago. Compared with its predecessor, H.264 / MPEG-4 AVC, it saves 50% bit rate while retaining the same subjective quality. On the one hand, this allows networks to broadcast UHD video commercially via satellite, and on the other hand it allows the transmission of an increased number of HD channels with limited bandwidth such as terrestrial transmission.

One year after this first version of the standard, the first extension for professional applications was completed; these are known as range extensions. These add 4:2:2 and 4:4:4 color subsampling and a monochrome format to HEVC, and also allow coding with higher bit depths up to 16 bit. At the same time, work is being done on scalable video coding and 3D video coding. The scalable (SVC) and 3D multiview (MCV) extensions have been available since July. The 3D extension with coding of depth information for auto-stereoscopic screens is slated for completion in February 2015.


AUX
MON 1
MON 2
FX
PAN
MUTE
15/16
17/18
PFL SOLO

Standardization in the audio sector

In the Audio & Multimedia Division, the standardization activities concentrate on mobile telephony and broadcasting.

To enhance the quality of mobile phone calls and to create a unique communication experience, Fraunhofer IIS is working on a new standard for the efficient coding of audio signals, which will be finished by the end of 2014. Thanks to Enhanced Voices Services (EVS), telephone calls sound more natural, are less stressful and meet the highest audio quality demands. The newly-developed audio codec offers excellent audio quality at very low coding delays.

In addition to mobile telephony, Fraunhofer IIS is also working on technologies for the future of television. In 2013, Fraunhofer's proposal was chosen for a new 3D audio technology by the MPEG standardization organization to provide a basis for the future open ISO standard MPEG-H audio for the delivery of high-quality 3D audio content. The launch of this new standard is scheduled for the spring of 2015, and will enable the efficient transmission and reproduction of 3D audio signals over all speaker systems and devices. This type of technology is just one of the ways that Fraunhofer IIS is bringing television into the 21st century.

TRENDS IN POST-PRODUCTION

Without post-production, no movie makes it to the big screen, TV, or your smartphone. Digital post-production isn't just for cutting the raw film material, it's also for the addition of light, music, sound, and special effects. But that is complex. Post-production usually takes about as long as filming itself does. In particular, the various input and working formats result in time-consuming and costly processes. Help is now expected to come from a uniform, standardized format called the Interoperable Master Format IMF for short. The Society of Motion Picture and Television Engineers (SMPTE) developed this standard together with the major Hollywood studios. It is intended to simplify data exchange between different post-production areas and internally during the intense phase of editing. The IMPs (Interoperable Master Packages) produced can then be used as a universal master and source format for various data formats for distribution – whether it be for film and cinema formats, broadcasting, or streaming.

Post-production in the cloud

Particularly in post-production, which requires a lot of computing power, considerations are starting to be given to cloud-based solutions. Particularly for animated films, where complicated models for objects and scenes often have to be calculated using enormous computing power, cloud computing can offer a cost-effective and adaptive alternative to individual rendering farms. At the moment, there are many doubts regarding security, and these are to be considered separately when dealing with special media clouds. One particular challenge is posed by live productions and outside shots. Because the effort required for laying cables may

not be too great, there is frequently insufficient bandwidth to be able to transmit the huge quantities of high-definition data. Heavy compression is not desired at this stage of production. The solution is to be found in the use of scalable data formats. Fraunhofer scientists are among those carrying out this research.

Scalable data formats

The approach of the scientists: to begin with, only the data from the main camera is transmitted in high definition; satellite or auxiliary cameras offer a preview version of the video recorded. If, however, something interesting is spotted in a preview, the high-definition data stream can be requested from exactly this camera on demand in real time. The advantage of this is that you can seamless switch to any currently running camera and there is limited bandwidth, no image is lost. The recording or transmission is only delayed and readjusts to the currently important images for the recording. In consequence, overall production is able to change resolutions on request at any time and therefore use the available resources efficiently. Together with film archives, Fraunhofer experts are working on efficient content management. The management system is intended to use a database to allow content stored at different locations to be accessible and ready to play. Systems of this nature are also of interest to digital archives, for example.

PUBLISHER'S DETAILS
Fraunhofer-Gesellschaft
Communications
www.fraunhofer.de

Concept: Angela Raguse,
Fraunhofer-Allianz Digital Cinema
Editorial Content: Angela Raguse, Birgit Niesing
Production: Christa Schraivogel

Photo acknowlegments
Cover, page 9: shutterstock
page 18: Planetarium Hamburg
page 23, 38: dpa
page 26, 34, 36: istockphoto
page 31: photocase

Graphics + Layout
Vierthaler & Braun