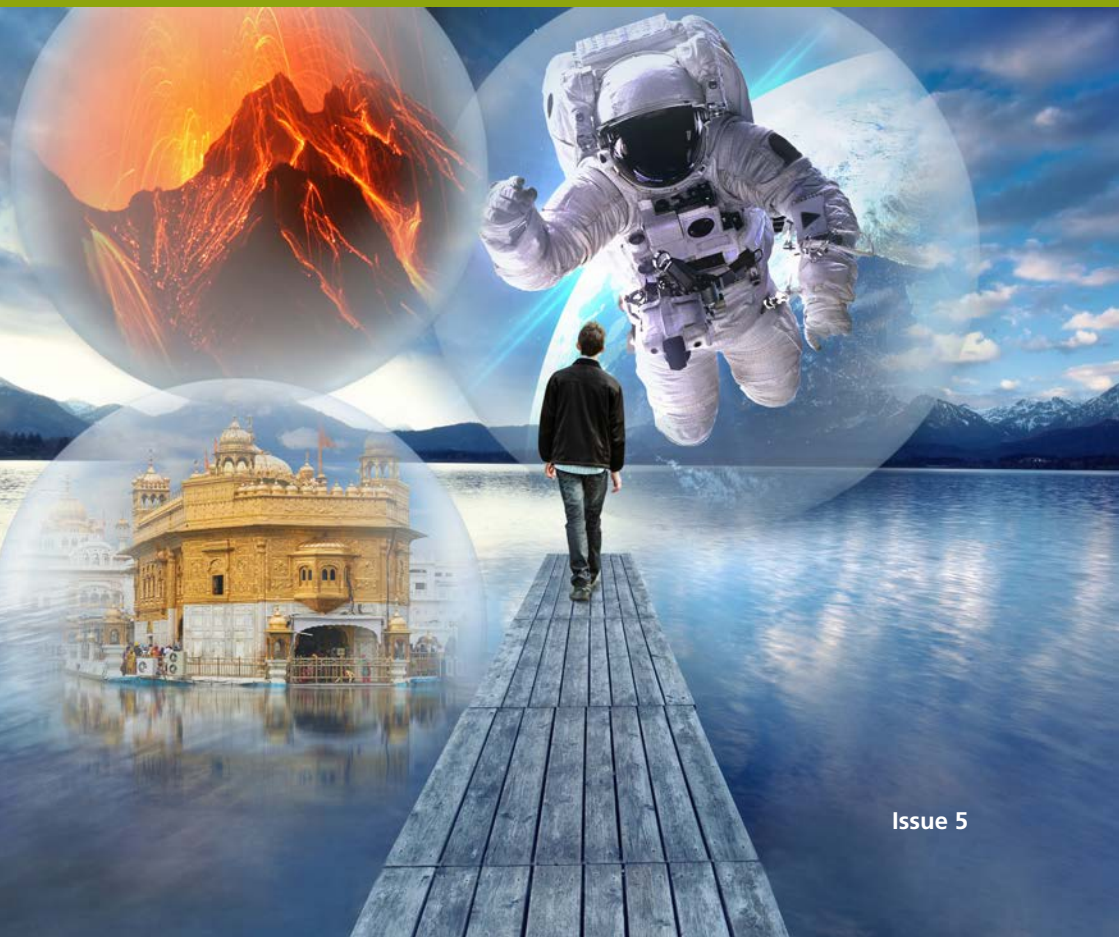
Fraunhofer

TRENDS AND TECHNOLOGIES IN DIGITAL MEDIA



Issue 5

CONTENT

PREFACE

Virtual Reality is one of the most appealing and promising new experiences when it comes to enjoying media content. The interactive selection of the view provides more than audio and picture – especially for all users who have grown up with Star Trek and computer games. Wearing VR glasses is no longer a barrier – what counts is the real-time view selection of video and sound content in best quality very close to reality. This immersive experience of real-life content makes VR a door-opener for new media business, in a way that is different from CGI-created content.

Advances in image sensor technology provide higher resolution, a wider color gamut, or higher sensitivity, which allows for better capturing of real scenes. New display technologies enable better reproduction and smaller dimensions. 360 degree capturing combined with 3D audio takes the user into the scene. Computation of high data volumes is possible even in smartphones. By combining VR content with binaural rendering from object-based audio, the realism will be intensified. What is more, the viewer can navigate through the scene, and interact with objects or content in the VR world.

With 360 degree the viewer can turn his or her head to the left and to the right, up and down. New production technologies, e. g. light-field, even extend these three dimensions of freedom to six and allow the viewer to move forward and backwards or step to one or the other side, too. But for VR, there is still some way to go. Bottlenecks such as parallax-compensated image stitching are demanding challenges to solve. And the questions of efficient business models for VR, which include the foundation of platforms and video channels, have still not been answered satisfactorily. We will give you an overview of new technologies that take you to your VR experience.

Dr. Siegfried Fößel

ENJOY VR

New technologies from Fraunhofer laboratories ensure that VR worlds can be enjoyed without limitation.

We need eyes and ears to perceive the world around us optimally. This also applies to virtual realities. We are only convinced by them when both the picture quality and sound are right.

Convincing all-round panorama

Researchers at Fraunhofer HHI have developed a technology to give us a realistic picture impression. At its core is the OmniCam-360: it records 360 degree panorama images, or "all-round" images, without parallax. This means that there are no kinks in the overall visuals that could make a violin bow or a hand "disappear". With other recordings, this happens frequently wherever the images from two cameras overlap. "Instead of placing the various cameras back to back in a star-shaped arrangement, the OmniCam-360 makes use of ten cameras that are arranged vertically. With the aid of mirrors they seem as if they are all looking at the scene from exactly the same point," says Christian Weissig, Project Manager at Fraunhofer HHI. "For the first time, automatic post-production is possible without manual correction."

Thanks to their OmniCam-360, the researchers were now able to support live-streaming applications for the first time – to transmit a concert by the Berlin Philharmonic. The users could enjoy the concert either on a smart TV or a computer, or even put on VR glasses to be in the "live" audience. Either way, the viewers had a much better view than the actual concert attendees, as they were

"standing" on the stage right behind the conductor.

The recordings made by the Omni-Cam-360 have a resolution that is ten times that of HD, or, to be more precise, 10,000 x 2000 pixels. As televisions and VR glasses are currently unable to display this resolution, the researchers are working on adapting the data transmitted to match the resolution of the end device in question. Furthermore, in the future they would only like to transmit the data that lies within the viewer's point of view – the "region of interest", as it is known.

Enveloping audio impressions

A convincing VR experience is not only based on what the eyes can see, as 50 percent of the experience depends on the authentic audio reproduction. Flaws in audio instantly ruin the illusion of an artificial world. However, if done properly, the 3D sound is authentic and allows users to enter the virtual world. The audio impressions lead users through the virtual environment where they become a part of what is happening around them. For example, a loud noise can simply change the direction of a user's gaze, allowing for entirely new types of storytelling.

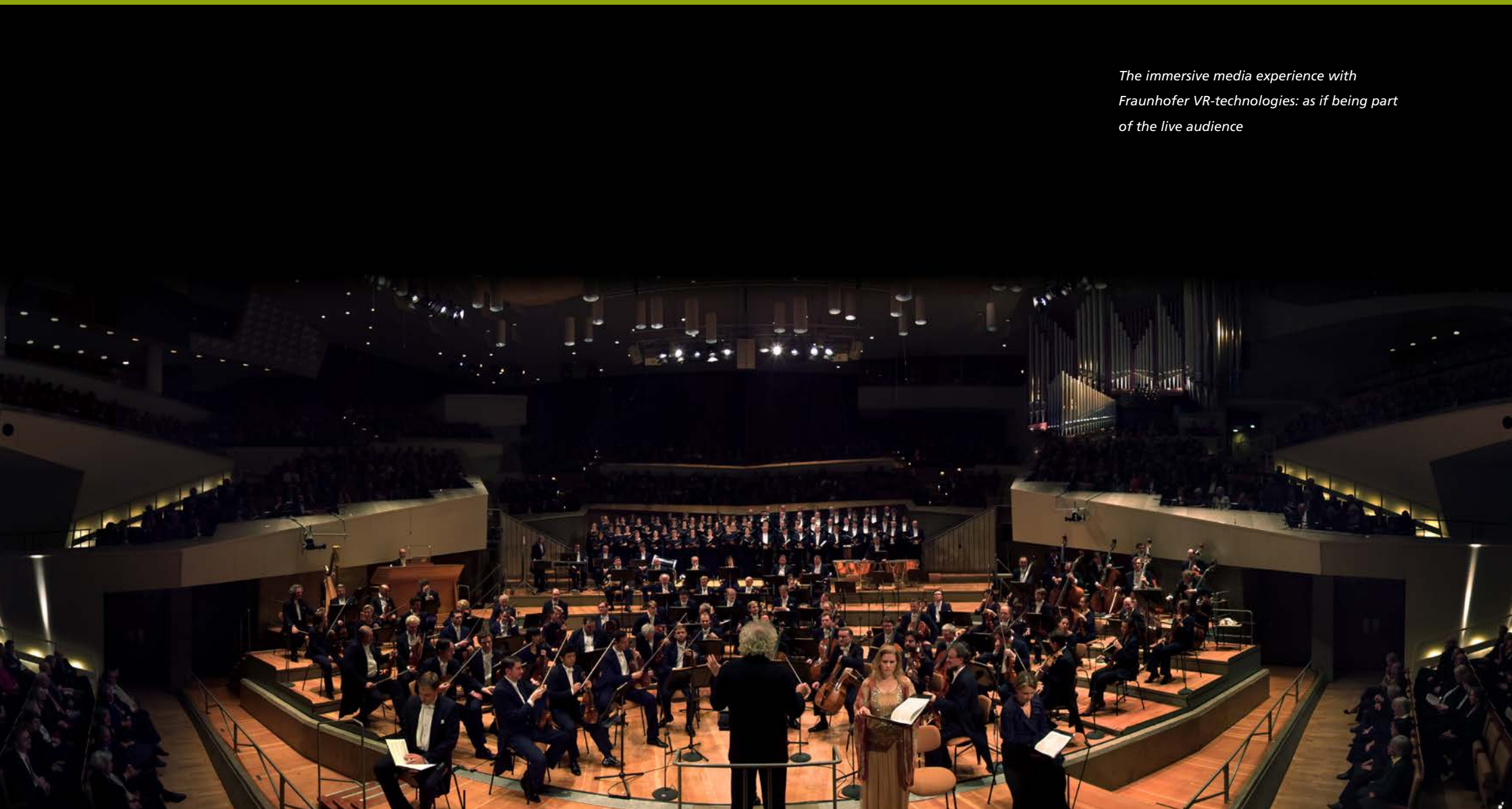
Creating convincing audio content for virtual worlds can be challenging; however, the developers at Fraunhofer IIS are working on innovative solutions to deliver this to consumers. One answer is Fraunhofer Cingo, today's leading solution for playback of enveloping audio for VR applications via headphones, which supports every audio format for surround and 3D sound. For a highly realistic listening experience, Cingo dynamically adapts the sound to the user's head movements in real-time. Fraunhofer Cingo is already being used in Samsung and LG's VR headsets.

In addition to playback, proper recording, production, and transport of audio signals are critical to a successful VR experience. For this reason, Fraunhofer IIS developed technologies for the complete VR audio chain of production. A specially optimized algorithm that can be applied to different types of cameras with either built-in or external microphones enables convincing 3D sound recordings. Additionally, post-production tools allow straightforward mixing and mastering of 3D sound while the audio codecs HE-AAC and MPEG-H, co-developed at Fraunhofer IIS, efficiently transmit surround and 3D sound, e.g. in a live-streaming app.

Thanks to Fraunhofer's audio and video technologies for VR, the next virtual concert will be a feast for the eyes and ears, where the users' experience will be as if they were attending a live event in a concert hall.

VIRTUAL REALITY – TECHNOLOGIES THAT TAKE YOU THERE!

The immersive media experience with Fraunhofer VR-technologies: as if being part of the live audience



REAL PEOPLE IN VIRTUAL 3D WORLDS

A 3D recording system replaces the synthetic avatar: At the moment, avatars – virtual people – stumble quite unnaturally through virtual worlds. A new technology will make the movements of virtual humans appear smoother and more natural.

The background is deceptively realistic: trees and houses look like real life, even the cars are not noticeably different from real vehicles. If you weren't wearing VR glasses, you might imagine that you were in real life instead of virtual reality. Only the "humans" are still moving somewhat unnaturally, and the textures – the weave in the characters' sweater material, their five o'clock shadow, or moles on their skin – also still have quite a way to go to appear natural. In order to copy how humans move, the creators of the avatars usually use motion capture: this tracking process takes a real person and places markers on their body; the movements of these markers are then recorded by a camera. Alternatively, the person can wear a special suit containing a large number of sensors that capture the movements of the arms, legs and other body parts, and forwards the data directly to the PC. This information is then used to animate the avatar in question.

Fluid, lifelike movements

Researchers at the Fraunhofer Heinrich-Hertz-Institute HHI have now developed a more expedient technology. "We record the real person from all spatial directions using 20 to 30 cameras and then calculate a robust three-dimensional model," explains Dr. Oliver Schreer, Group Leader at the Fraunhofer HHI. "We can then add this model directly to the virtual

world. We plan on accelerating capture and 3D modeling to allow us to integrate the person into the virtual world in real time. It is no longer necessary to animate the person any further." The result is that the person moves as fluidly and naturally in the VR world as in the real world; the person's motions are carried over one-to-one instead of having to animate them as previously. The texture also looks real to the viewer. The viewer can observe the avatar from any point of view of his or her choice – it is, for example, possible to move in 360 degree around the avatar. This makes one's virtual companion appear to be extremely true to life.

A 3D model makes it possible

The real trick is in the software that creates the 3D model. But let's start at the beginning: the cameras – all of which are in pairs – are distributed at optimum positions around the room. Each pair of cameras captures three-dimensional information about the real person. Similar to our two eyes, these cameras receive information about how far each body part is from the camera in each recording. The software calculates this depth information for each camera pair up to 50 times a second, which is also the rate at which each camera pair takes pictures. The software then fuses the data from individual camera pairs together. The result is a life-like three-dimensional copy of the person, together with their movements, which can be rendered for any viewing angle. Mapping faults and undercuts are now a thing of the past. For users, this means that they can see areas that were previously hidden when only a single camera pair was used – hidden by the hand that the person is holding in front of their torso, for example. This three-dimensional model can be integrated directly into the VR world – the real person "hatches" from the camera set and into the artificial world with all their natural movements.

TELEVISION WITH AN ALL-ROUND VIEW

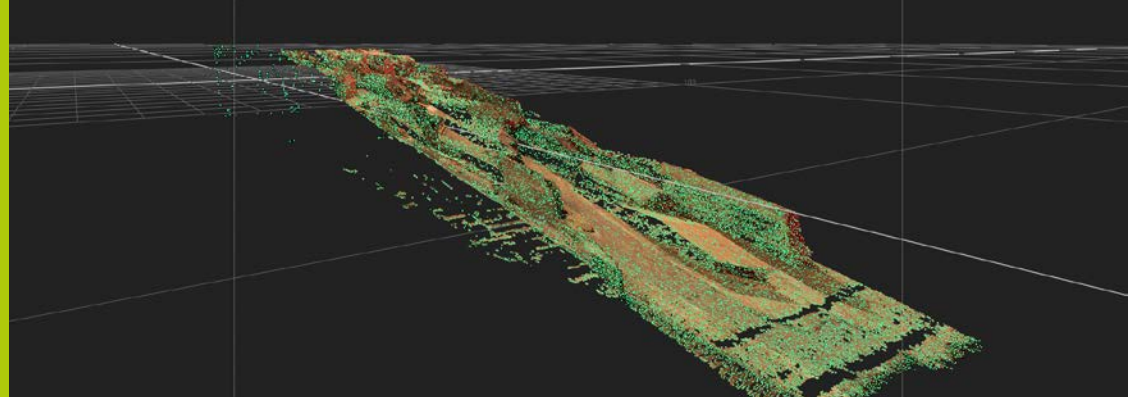
In the future, TV viewers will be their own "cameramen". The established standard HbbTV can be used for 360 degree TV transmission.

The helicopter blades whir as it soars over the mountain peaks, an eagle glides past the side window. How nice it would be to be able to turn your head and watch where the eagle goes. But, alas, you're watching the scene as part of a TV documentary from the comfort of your own couch and you have to settle for the perspective that the cameraman chose. But, with 360 degree TV, it's different: here, the end user can play director and decide which direction he or she would like to look. Dr. Stephan Steglich, Division Director at the Fraunhofer Institute for Open Communication Systems FOKUS, explains how it works.

Dr. Steglich, how should we imagine 360 degree television?
A special camera setup records the film in all spatial directions; the sound is also recorded spherically. For playback, we start by deciding on a specific viewing direction – the viewer's attention should be principally focused on the direction in which the action is happening. From here, the viewer can use his or her remote control to move their point of view from side to side, up, back, or down. It's no longer just one film; it's several films in one.

This requires large quantities of data to be transmitted. What technological requirements need to be fulfilled to bring 360 degree TV to viewers?
One possible solution we have identified is the recognized standard HbbTV, which is short for hybrid broadcast broadband TV. TV stations usually use this standard in conjunction with an Internet connection in order to show their viewers additional content. We use HbbTV in order to move functions that television sets previously couldn't perform onto the Internet – into a cloud, to be precise. So we make sure that the TV set only does what it's able to. This means that if the viewer chooses a specific viewing angle, the calculation is performed in real time in the cloud. The television doesn't even "notice" what it's doing.

What does the user need to have in order to be able to benefit from this technology in the future?
Basically all he or she needs is a modern television that is HbbTV ready. We deliberately left out additional devices such as VR glasses in order to be able to reach as wide a public as possible. A test version of our 360 degree television is already being broadcast. Our technology is even of interest for mobile end devices such as smartphones: the more complexity I can move from the device into the cloud, the longer the device's battery will last. The data volume is also reduced.



STANDARDIZATION ON LIGHT FIELD MODALITIES

One of the most important achievements of mankind is the transfer of knowledge to future generations. This is done by several methods: verbally, by drawings, in writing and printing, or lately by photographs and video. All these methods are exchangeable in some way; however it is a great deal more effort to describe a situation verbally than by a photograph. The phrase "A picture is worth a thousand words" became popular in the early 20th century and well describes the phenomenon that people can incorporate knowledge by visual sensors much faster than by hearing and interpreting words. In the last hundred years, this fact was implemented by capturing images with cameras, in the form of a still image as photograph, or in combination with sound as motion picture or video. Technology in the 20th century allowed fast imaging of the scene on a flat planar sensor, either on a photosensitive emulsion or on an electronic sensor. Several attempts were made to increase the dimensions of the reproduction of the scene, e. g. for stereoscopic 3D and also to get some depth information. But it is difficult to capture and reproduce the scene in the exact same way as an individual with only two images.

A revolutionary approach to image production

With the advent of low-cost imagers, additional sensors for localization or depth acquisition, and new electronic display devices, true three-dimensional acquisition and reproduction for a scene are in sight. The approach is not only to capture and reproduce the two images seen by human eyes, but to capture the complete scene in multiple dimensions, to model the scene, and to render the images based on demands as in computer generated imagery (CGI). This can be the viewpoint, the viewing direction, or the focal point.

The theory behind this has been known since the end of the 19th or the beginning of the 20th century. It was developed by Faraday, Lippmann, and Gershun and is called light field theory or plenoptic. This states that, at every point in space, the light rays form a function based on position, direction, and intensity. The traditional cameras today capture the light rays in a specific position, but bundle light rays from one or more directions by the lens and record this on a photosensitive sensor. The aperture of the lens defines the number of bundled rays.

It is clear that we cannot capture the complete light-field as we cannot position a camera in every point in space without their influencing each other, but it is possible to sample the light-field in different ways. The simplest way is to use a circular camera array for capturing 360 degrees, allowing Virtual Reality (VR) reproductions in VR glasses. By using different arrangements of cameras and special sensors and optics, other applications are possible, such as VR with translational movement or 3D modeling or generation of point clouds with the option of merging real scenes with CGI scenes. Today, the industry is investigating which technologies allow which applications. This relates to Virtual Reality, Augmented Reality (AR), point clouds, light field cameras, visual effects VfX, depth enriched images, 3D models; in general, every process of mixing real scenes with computer-generated scenes.

Standardization of light/sound field technology

Within the standardization committees ISO/IEC JTC1 SC29 WG1 (called JPEG) and WG11 (called MPEG), several groups are working on these topics, discussing the possible modalities and correct representation forms. Attached is a list of the current working groups.

– **Joint Ad Hoc Group for Digital representations of Light/Sound Field for immersive media applications**
This group was a joint ad hoc group between JPEG and MPEG to determine the state of the art, potential modalities, and applications. The summary technical report can be found as JPEG Pleno Joint AhG Report on the JPEG website.

– **JPEG Pleno**
The Adhoc Group JPEG PLENO targets a standard framework for the representation and exchange of new imaging modalities such as light field, point-cloud, and holographic imaging. Several workshops for this topic were executed. At the last ICME 2016 conference, a grand challenge for light field image compression was offered.

– **JPEG Systems**
The subgroup JPEG systems investigates the backward compatible extension of JPEG for integrating multiple images and depth maps allowing depth enriched images.

– **MPEG Wave Field Audio**
As well as images, audio experiences are also position dependent and used for object localization by viewers. A group in MPEG investigates the use of wave field/sound field technology to support less restrictive viewing conditions, such as flexible point of view and object-based sound representations. As in the image field, microphone arrays or microphone for individual sound objects are used to capture the sound field.

– **MPEG Tools for VR**
VR is a fast-progressing technology with a strong market momentum. Several technologies are necessary to realize a VR workflow such as image stitching, metadata enrichment, development of optimized video coding technologies. To avoid fragmentation in this market, the standardization of formats, 3D projection methods, metadata, and signaling is necessary.

– **MPEG Light-Field Coding**
Light field enables many different special effects, like the "Matrix Bullet effect", refocusing, or depth-based editing. For next-generation cinematic movies in particular, 3D processing allows a much more immersive viewing experience. In combination with audio wave field processing new workflows are possible. The focus in light field coding is the representation and coding of light fields captured by lenslet light-field cameras or camera arrays.

– **MPEG Point Cloud Compression**
Images from cameras used in light field capturing are projections of light rays to the sensor. A more universal representation is a point cloud. In this cloud, the position, light rays and Bidirectional Reflectance Distribution Function, BRDF is defined for specific objects. Because of the limitation of data points during capturing, point clouds are sparsely distributed in most of the cases. The benefit of point clouds is the 3D representation, which allows new renderings from different view-points.

In the next few years, light field technology will become a highly interesting technology for improving the viewing experience. Today, this field is being very intensively developed. It is still unclear which representation formats are necessary and useful to allow optimized workflows.

The JPEG and MPEG committees will investigate and standardize new formats and methods for this new technology.

Overview by Dr. Siegfried Foessel, Fraunhofer Digital Media Alliance

OMNICAM-360 DEGREE FIRST LIVE TRANSMISSION WITH THE BERLIN PHILHARMONICS

Since the OmniCam-system was first developed at Fraunhofer Heinrich Hertz Institute in 2009, numerous video productions have been realized in cooperation with the Berlin Philharmonics – a reliable top-class partner. After various productions, the OmniCam-360 was installed to record the concert given on occasion of the 25th anniversary of the fall of the Berlin Wall in 2014. 2015, the documentary "Playing the Space" was released, in which the Philharmonics plays a leading role and which was also filmed with the panoramic camera system.

The compact OmniCam-360 from the Fraunhofer HHI is equipped with ten HD cameras that are attached to a mirror system. The individual produced images are corrected in real time and put together into a parallax-free UHD video panorama with a resolution of approx. 10,000 x 2,000 pixels. The OmniCam-360 measures about 50 x 50 centimeters and weighs around 15 kilos. Therefor it can be utilized for many different cases.
The collaboration with the Berlin Philharmonics was taken to a new level when the OmniCam-360 came to use in the first live transmission: The panoramic images were successfully streamed in UHD quality through the Philharmonics Digitals Concert Hall platform and shown online. By wiping and/or zooming, viewers can focus on the part of the picture that interests them the most. In broadcasts of concerts this could be the section of the orchestra with their favorite instruments, or the conductor, or they could just let their gaze wander.

MPEG-H: DELIVERING IMMERSIVE AUDIO TO THE WORLD'S FIRST TERRESTRIAL UHD-TV SYSTEM

Ultra High Definition Television (UHD TV) is the next best thing to reality. UHD TV transports crystal clear, high-resolution pictures that make you want to touch the screen to check that the race car will not drive through your living room. The 4K and 8K image resolution allows UHD TV to create the illusion of reality. In addition, it offers the extension to 3D video reproduction, High Dynamic Range Imaging (HDRI) to display image luminance in more detail and High Frame Rate (HFR) which provides a more fluent picture sequencing and image quality by exceeding the typical frame rates used today.

But video reproduction is only one part of the equation for a convincing TV experience. Sound is the other central ingredient to authenticity. Gladly, technological developments on the audio side are equally innovative: The next-generation audio standard MPEG-H 3D Audio takes three-dimensional sound from the real world into the living room. MPEG-H provides a full audio immersion experience for TV viewers, which creates the feeling of "being there" rather than just watching a TV program.

The MPEG-H Audio standard, mainly developed by Fraunhofer IIS, was selected as A/342 Candidate Standard for the new television standard ATSC 3.0 in May 2016. Shortly after, the South Korean standards organization Telecommunications Technology Association (TTA) announced that the audio technology would be part of the Korean standard for terrestrial UHD TV. If governmental officials approve the standard in September, MPEG-H 3D Audio will provide an entirely new sound experience in the world's first 4K terrestrial TV system.

If approved, South Korean consumers would be the first to experience interactive and immersive sound, the two main features of MPEG-H 3D Audio. The technology enables users to adjust the sound mix to their preferences. For example, consumers would have the ability to choose between different commentators in a sporting event or improve the intelligibility of dialogues by decreasing ambience volume. Furthermore, by adding 3D audio components for additional height information the codec delivers true audio immersion.

With the 2018 Winter Olympic Games in Pyeongchang, South Korea on the horizon, the new TV terrestrial system is expected to launch in February 2017, starting with the Seoul metropolitan area. According to Korean government representatives, it will then be expanded to cities near the Olympic Game venues. By 2021, the service will be available nationwide.

The MPEG-H 3D Audio codec is ready to use in broadcast systems, as demonstrated by Fraunhofer IIS in a series of events. Fraunhofer's latest presentation of a complete ATSC 3.0 broadcast chain with MPEG-H Audio was at KOBA Show in Seoul, South Korea in May 2016.

MPEG-H 3D Audio is already being integrated into professional broadcasting equipment. Most recently, the broadcast equipment manufacturers DS Broadcast and Kai Media announced the availability of MPEG-H 3D Audio in their latest 4K encoder products.
More product announcements, including TV sets from major manufacturers for the Korean market, are expected this fall.

XHE-AAC: THE CODEC OF CHOICE FOR STREAMING AND DIGITAL RADIO

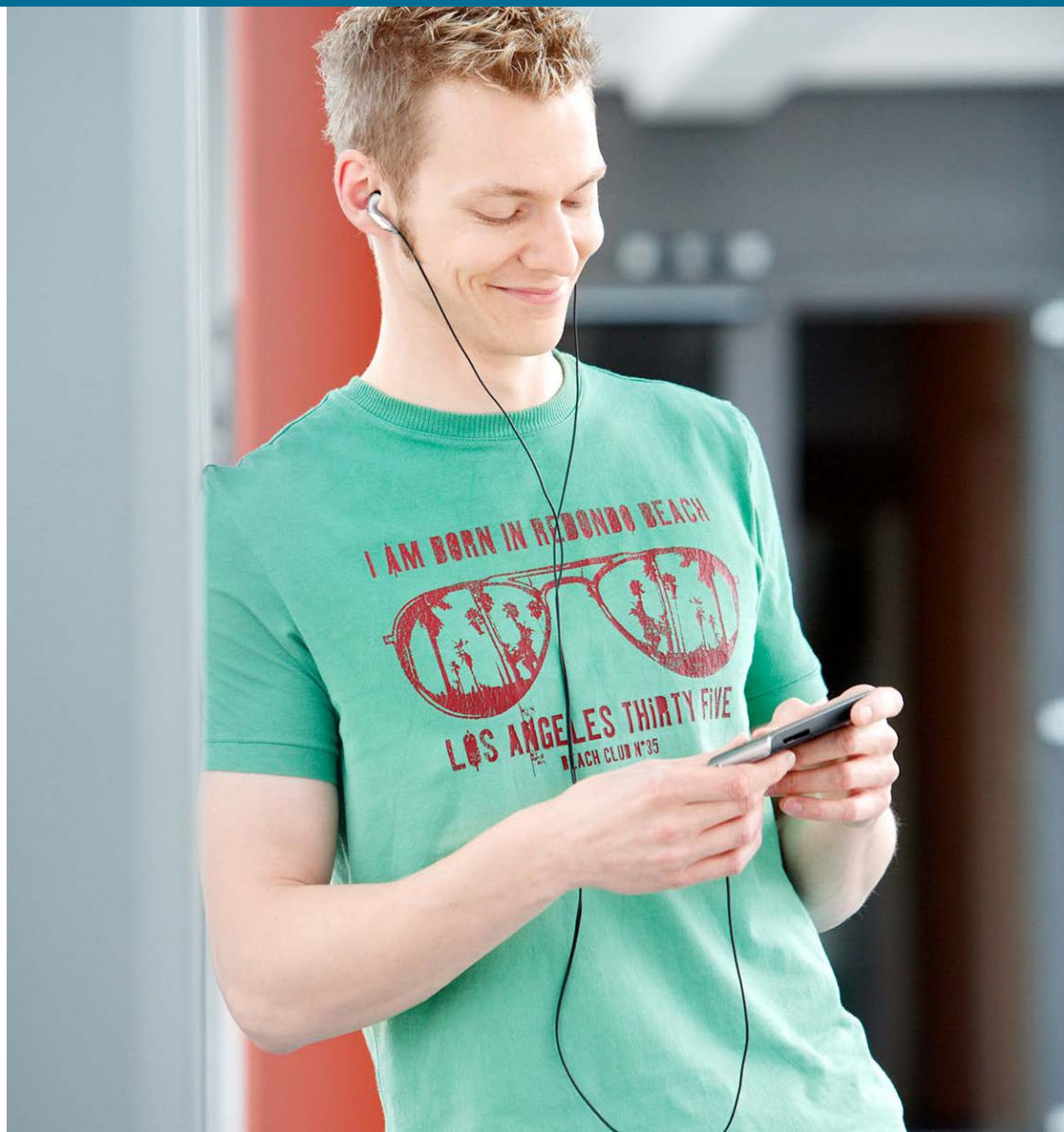
As featured by Radio World International on 13 June 2016

Today, mobile music streaming services are facing a series of challenges. Service providers are suffering due to the inability to serve the full range of potential customers because of real-world bandwidth limitations combined with a high playout/CDN cost. In addition, today's mobile user experience is far from perfect. Streaming services are eating into consumers' monthly data allowances and service reliability is far from what we have grown to appreciate from classic broadcast coverage.

However, there is an answer. Recently, the MPEG working group in ISO standardized and published the latest member of the highly successful AAC family of audio codecs, "xHE-AAC" (Extended High-Efficiency Advanced Audio Coding).

xHE-AAC combines and significantly improves the functionality of two formerly separated worlds: On the one hand the general-purpose audio codecs and in particular today's de-facto standard MPEG HE-AAC, on the other hand codecs for speech-only content at very low bit rates represented by AMR-WB+.

xHE-AAC has the ability to provide good quality audio down to 8 kbps for mono and 16 kbps for stereo services regardless of the type of content (e.g. talk radio, a music program or a jingle). Even in the mid bit rate range, xHE-AAC shows signif-



icant quality improvements over established codecs, while for higher bit rates it converges to the same transparent quality known from the other members of the AAC family.

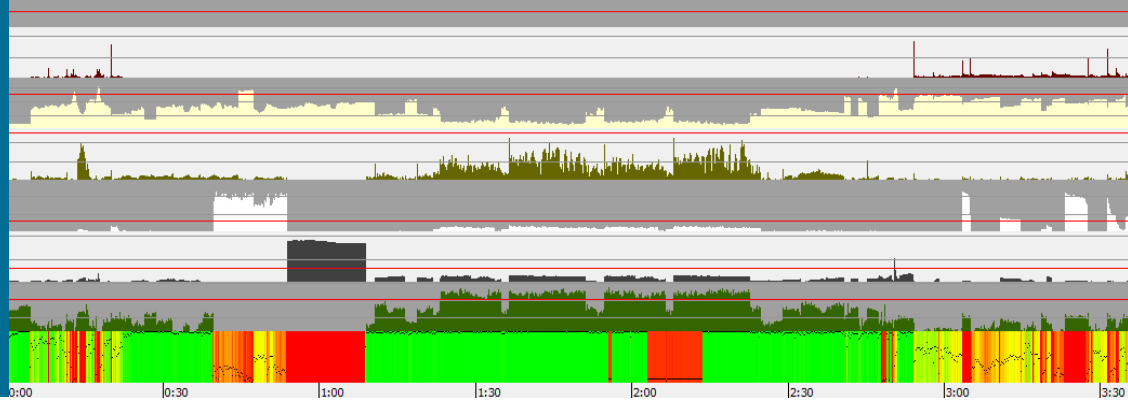
Due to its universality, the global Digital Radio Mondiale (DRM) digital radio standard has adopted MPEG xHE-AAC as its primary audio codec. India is on the verge of becoming the world's largest digital radio deployment. In February 2016, attendees at Broadcast Engineering Society of India Convention in New Delhi witnessed the first DRM transmission based on xHE-AAC by India's public broadcaster All India Radio. DRM ready receiver chipsets and receivers support xHE-AAC out of the box, as well as all providers of professional DRM encoders and multiplexers.

For audio streaming over IP to mobile devices, xHE-AAC promises a revolutionary impact on both user experience and business opportunities. Streaming service providers for the first time can now start targeting potential customers limited by affordable 2G contracts who so far could not be served using existing audio codecs, but in the case of India represent roughly 90 percent of mobile users. Furthermore, service providers can benefit from a dramatic cut in monthly costs of data distribution. Sample calculations for mid-size services show a saving potential of 75 percent of the CDN cost once mobile apps start requesting the low bit rate xHE-AAC streams instead of the mp3 and HE-AAC versions.

Consumers with 4G/LTE contracts around the world benefit from the increase in service reliability when leaving the well-covered city centers and falling back to 2G-coverage, or joining mass events with plenty of mobile data usage.

In April, Via Licensing announced the start of the official xHE-AAC licensing program, which includes the functionality of the existing AAC program. This makes up for the fact that xHE-AAC capable audio decoders are downward compatible to all existing AAC based content. The licensing program also makes it more convenient for manufacturers to mix xHE-AAC and AAC-only devices and apps.

Professional streaming encoders are already available by Telos Alliance and StreamS.



POST-PRODUCTION TOOLS

IMF – a uniform exchange format for post-production

easyDCP, the post-production software from Fraunhofer IIS, has offered processing of IMPs (Interoperable Master Packages) in addition to DCP generation for quite a while now. The necessary Interoperable Master Format, IMF, proves its interoperability successfully over and over at various Plugfest events and serves as a universal exchange format within the movie production. Since standardization by the SMPTE (Society of Motion Picture and Television Engineers), IMF has been establishing itself more and more in the media industry.

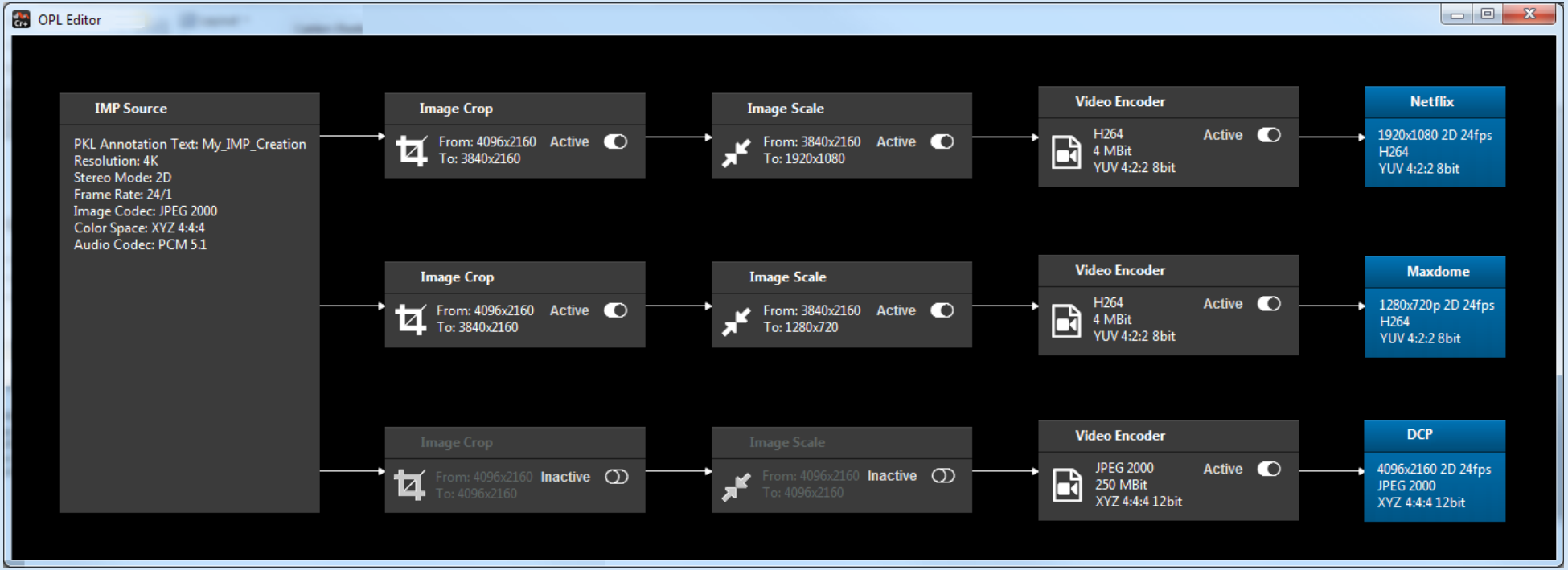
The advantages of the IMF are obvious: a uniform and internationally standardized post-production format guarantees seamless exchange of content in the highest picture quality; it is no longer necessary to support a lot of different exchange formats, therefore the risk of quality losses due to unwanted decoding, image-conversion and encoding steps is minimized. At the same time, the IMP serves as a master for the creation of distribution formats. In practice, it is not uncommon that up to 100 versions of different distribution formats need to be created from a single master and checked simultaneously. IMF allows for the automatic creation of videos with different technical parameters such as video compression format, spatial resolution of the target device, audio speaker set-up, or different national versions (subtitles, soundtrack in the local language etc.). Using the IMF-concept called Output Profile List OPL, this work can be automated, since the conversion of the IMP into a certain distribution format (e. g. iTunes) is described here.

Integrated quality check

The great number of largely automatically created videos means that a comprehensive and, indeed, automatic quality check of the transcoding process is desirable in order to save time and money. Usually the material is inspected for errors and artifacts that may arise during lossy transcoding or conversion into other color spaces or bit depths. The Fraunhofer researchers' approach is a direct integration of various quality check modules into the transcoding processes. To this end, the scientists will present a version of the software that can be used both to create the IMP and, in particular, to check the generation of distribution formats from an IMP automatically.

At IBC 2016, the Fraunhofer Institutes IIS and IDMT will showcase software components that detect problematic sections of video and summarize them in a test report. In the first demonstration version, the test results will be shown as a simple "traffic light function" with red/amber/green marking. In future, the user will also receive a detailed quality report for those parts of the distribution formats that need further attention. The quality report is presented in a visual time-line directly integrated in the easyDCP software suite in order to establish an intuitive way to check the problematic scenes if there are any. The first fault modules for quality checking with predefined parameters have already been integrated into the overall easyDCP system.

For 2017, the IMF suite with transcoding and expanded checking functionality is planned to be rolled out to the first pilot users.



Output Profile List OPL allows for the automatic creation of different distribution formats

MICO – MULTIMODAL MEDIA ANNOTATION AND SEARCH

In order to provide improved search and recommendation capabilities for ever-increasing amounts of content, cost-effective technologies for automatic extraction of metadata from raw media objects are needed. One of the core challenges in this domain is the need to shift from using individual standalone extractors to multimodal and context-aware extractor workflows, which can provide substantially improved search functionalities and results.

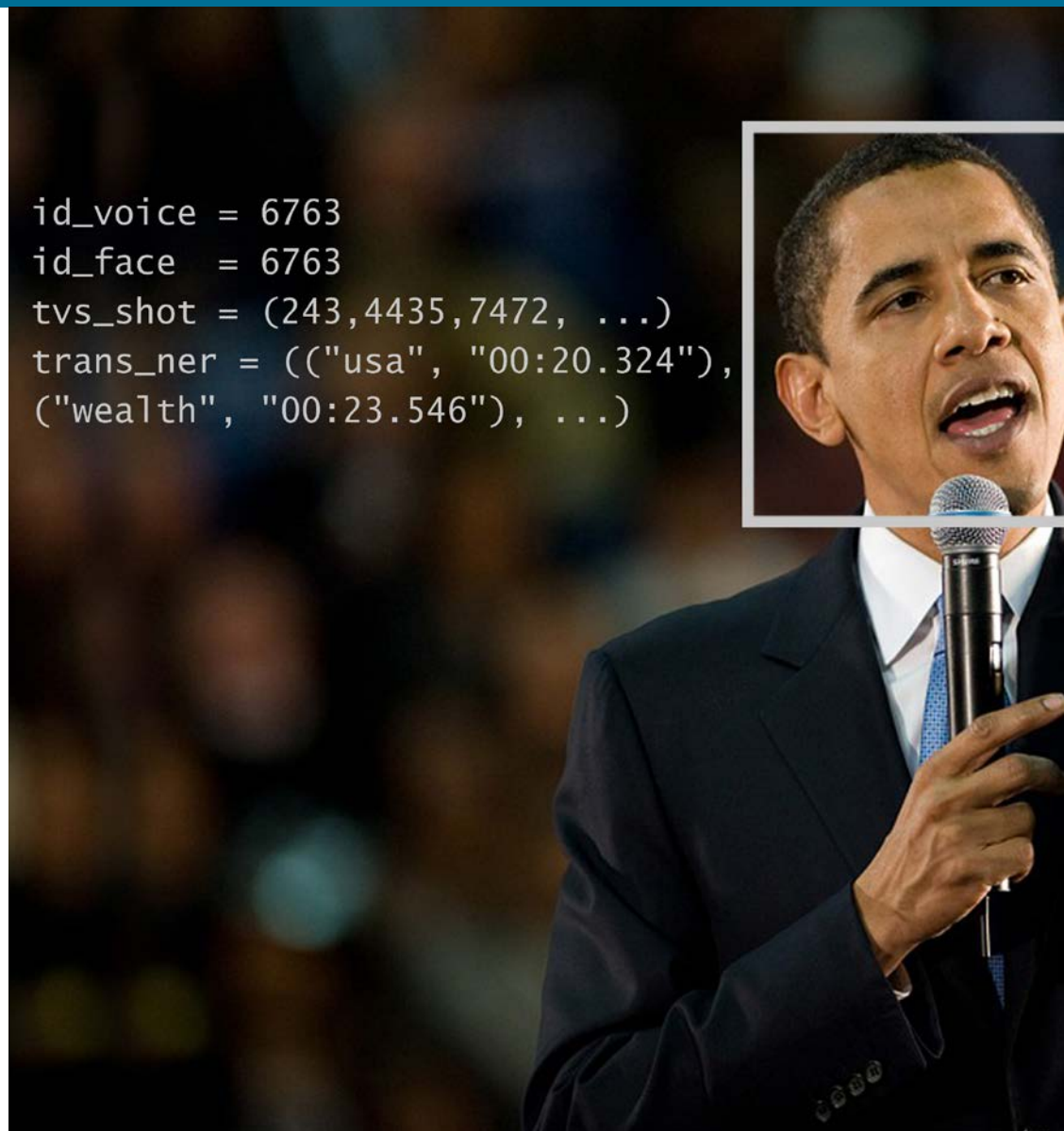
One example: even to realize seemingly simple search functionalities to find video segments where a (specific) person talks about a specific topic, a news video provider needs these extractors for annotation:

- Speaker recognition
- Speech recognition
- Named entity recognition
- Temporal video segmentation
- Face detection and recognition

By multimodal fusion and reuse of results from extractors, annotation quality can be improved at the same time, e. g. by combining speaker and face recognition, or by combining speaker and speech recognition. In order to realize such potential, the following challenges need to be addressed:

- Cost-effective integration of heteroge neous extractors
- Flexible extractor workflow orchestration
- An extensible, common metadata model
- Availability of an expressive, media-aware query language
- Recommender systems using both annotations and collaborative filtering

Fraunhofer IDMT has been active in the domains of A/V extractor implementation,



integration and orchestration, and recommendation for many years. We believe that multimodal, complex annotation and recommendation are keys to future media systems, but it requires an integrated technology platform addressing all aforementioned aspects. Such a platform has now been developed by the EU R&D project MICO, with all core functionalities being provided as business-friendly OSS.

Fraunhofer IDMT has provided significant contributions to the project and, together with other MICO partners, intends to further build on and extend the platform to realize the potential of multimodal annotation, search and recommendation.

easyDCP – ONLY A FEW CLICKS TO YOUR DCP

In order to create a DCP (Digital Cinema Package) easily and compliant to all the requirements, or to test such packages subjected to delivery, Fraunhofer IIS developed the easyDCP software suite. More information about easyDCP, which is available as standalone tool-set – as well as plug-ins for various post-production solutions incl. BMDs Resolve – can be found at *www.iis.fraunhofer.de/easydcp*

FRAUNHOFER CINGO®

Fraunhofer Cingo enables an immersive sound experience on mobile and virtual reality devices. For device manufacturers and service providers, Cingo is available as optimized software implementation for all major PC and mobile platforms, including iOS and Android. Equipment manufacturers using Cingo are Samsung (Samsung Gear VR), LG (LG 360 VR) and Google (Nexus family).

LIGHTWEIGHT IMAGE CODING FOR VIDEO OVER IP AND MEDIA CONTRIBUTION

The Lici® codec ensures image-by-image, visually lossless transmission of high-resolution video with compression ratios of 1:2 to 1:6. It features extremely low latency, high throughput and requires little logic to implement. Lici® is used for Video over IP applications and for contribution in professional movie production. IP Cores are available - licensing conditions can be sent on request. *www.iis.fraunhofer.de/lici*

MPEG-H

MPEG-H Audio provides interactive, immersive sound for TV and VR applications. It is available as software implementation to chip manufacturers, broadcasters and consumer electronics manufacturers. The first professional broadcast encoders that support MPEG-H Audio software are DS Broadcast´s BGE9000 4K Ultra HD Encoder and the new 4K UHD Live Broadcast Encoder KME-U4K from Kai Media.

FRAUNHOFER DIGITAL MEDIA ALLIANCE

As an one-stop competence center for digital media we provide for our customers scientific know-how and the development of solutions that can be integrated in workflows and optimize process steps.

The members of the Digital Media Network are actively working in renowned organizations and bodies like International Standardization Organization ISO, ISDCF (Inter-Society Digital Cinema Forum), SMPTE (Society for Motion Picture and Television Engineers), FKTG (German Society for Broadcast and Motion Picture), and in the EDCF (European Digital Cinema Forum).

Fraunhofer Institutes in the Digital Media Alliance jointly offer innovative solutions and products for the transition to the digital movie and media world of tomorrow. The Institutes in the Alliance are available as renowned contacts and partners for all of the digital topics connected to digital media, digital movies, and standardization, as well as new cinematography, audio, and projection technologies, post-production, distribution, and archiving. The goal of the Fraunhofer Digital Media Alliance is to quickly and easily help find the right contacts, partners, and suitable technology.

The Fraunhofer Institute members are
- Digital Media Technologie IDMT, Ilmenau
- Integrated Circuits IIS, Erlangen
- Telecommunications, Heinrich-Hertz-Institut HHI, Berlin
- Open Communication Systems FOKUS, Berlin

Contact
Fraunhofer Digital Media Alliance
Angela Raguse M.A.
Phone +49 9131 776-5105
alliance-dc@iis.fraunhofer.de
www.digitalmedia.fraunhofer.de

Publication Information

Fraunhofer Digital Media Alliance
c/o Fraunhofer Institute for
Integrated Circuits IIS

Am Wolfsmantel 33
91058 Erlangen, Germany

Concept and Editor
Angela Raguse
Fraunhofer-Allianz Digital Media

Layout and production
Kathrin Brohasga
Kerstin Krinke
Paul Pulkert

Photo acknowledgements
Cover picture: Fraunhofer IIS/fotolia.com
Page 3: Fraunofer IIS/Karoline Glasow
Page 4: Fraunhofer IIS/David Hartfiel
Page8/9: Fraunhofer HHI
Page 11: Fraunhofer HHI
Page 15: Fraunhofer Fokus
Page 20: Fraunhofer HHI/Berlin Philharmonic Orchestra/Monika Rittershaus
Page 23: Fraunhofer IIS/Frank Boxler/Valentin Schilling
Page 24: Fraunhofer IIS/Matthias Heyde
Page 27: iStock.com/sweetym
Page 29: Fraunhofer IIS/Fraunhofer IDMT
Page 30/31: Fraunhofer IIS
Page 33: CC0 Public Domain
Page 37: Fraunhofer IIS/Kurt Fuchs
Page 37: DS Broadcast