# Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays

Jan Froehlich<sup>a</sup>, Stefan Grandinetti<sup>b</sup>, Bernd Eberhardt<sup>b</sup>, Simon Walter<sup>b</sup>, Andreas Schilling<sup>a</sup>, Harald Brendel<sup>c</sup>

<sup>a</sup>Wilhelm Schickard Institute, University of Tübingen, Sand 14, 72076 Tübingen, Germany
 <sup>b</sup>Stuttgart Media University (HdM), Nobelstraße 10, 70569 Stuttgart, Germany
 <sup>c</sup>Arnold & Richter Cine Technik GmbH & Co. Betriebs KG, Türkenstr. 89, 80799 Munich, Germany

## ABSTRACT

High quality video sequences are required for the evaluation of tone mapping operators and high dynamic range (HDR) displays. We provide scenic and documentary scenes with a dynamic range of up to 18 stops. The scenes are staged using professional film lighting, make-up and set design to enable the evaluation of image and material appearance. To address challenges for HDR-displays and temporal tone mapping operators, the sequences include highlights entering and leaving the image, brightness changing over time, high contrast skin tones, specular highlights and bright, saturated colors. HDR-capture is carried out using two cameras mounted on a mirror-rig. To achieve a cinematic depth of field, digital motion picture cameras with Super-35mm size sensors are used. We provide HDR-video sequences to serve as a common ground for the evaluation of temporal tone mapping operators and HDR-displays. They are available to the scientific community for further research.

Keywords: High Dynamic Range, HDR-Video, Wide Gamut, Tone Mapping.

## **1. MOTIVATION**

Current trends in mainstream motion picture imaging include higher spatial resolution, higher temporal resolution and the use of high dynamic range (HDR) imaging. Whereas high resolution and higher frame rate videos can be acquired today, using current generation motion picture cameras, there is no single HDR camera with a Super 35mm sized sensor available today. The marginal application of HDR-video stands in great contrast to the domain of still imaging, where HDR-image capture is well studied and commonly practiced. Although there has been research on HDR-video acquisition<sup>1,2,3</sup>, no cinematic HDR-video has yet been gathered.

For scientists developing next generation HDR monitors and temporal tone mapping operators, it is crucial to have cinematic HDR-content available, because image quality assessments can only be performed using high fidelity images. These images must be of sufficient spatial resolution, temporal resolution and dynamic range. Throughout the last decade, professional digital film cameras gained around 4 stops of dynamic range, from about 10 stops in 2001<sup>4</sup> to 14 stops today<sup>5</sup>. With high dynamic range displays on the horizon, we expect this trend to proceed. To simulate the dynamic range of prospective recording devices today, the combination of two exposures is needed. Figure 1 shows a comparison between the dynamic range of a professional motion picture camera, our dual camera setup and current display devices.



Figure 1. Dynamic range of current acquisition and distribution devices compared: Alexa Camera<sup>6</sup>, our rig (2\*Alexa), a TV conforming to ITU-R BT.1886<sup>7</sup>, a SMPTE RP 431-2<sup>8</sup> cinema projection and a Dolby PRM-4200<sup>9,10</sup> HDR-display.

Image quality is not only determined by the signal quality of the image acquisition system, but also by lighting, make-up and staging. As an example, a faithful skin tone reproduction of a non-powdered actor in typical room lighting will not appear life-like to most observers. Humans often appear to be unhealthy or look fatigue in reproductions, when filmed without cinematic lighting and makeup. This is because our visual expectation for high quality images is to see staged pictures. Especially when dealing with non-expert observers in user studies, staged images are important to avoid misinterpretation.

Our goal is to provide cinematic footage that covers the dynamic range of tomorrow's sensors, so that tone mapping algorithms and HDR-displays can be evaluated regarding their ability to handle these future videos today.

#### 2. METHODS

HDR still images and videos are often captured by taking multiple images with different exposures, one after the other<sup>2,11,12</sup>. When dealing with moving objects, artifacts can be introduced by not taking all exposures at the same time. See the darkened flames in Figure 2 as an example. Therefore it is essential to capture all exposures simultaneously.



Figure 2. Ghosting artifacts introduced by time sequential HDR-image capture using an Apple iPhone 4S mobile phone with 'HDR' feature enabled in the standard camera app.

#### 2.1. Acquisition

To generate HDR-video with different exposures captured at the same time, a mirror-rig as shown in Figure 3a is used. A common glass pane with antireflective coating is employed as a beam splitter. This results in a ratio of around 1:16 between reflection and transmittance, shifting the camera exposures by 4 stops. To be able to use large sensor motion picture cameras, the mirror is mounted in front of the lenses, instead of splitting the light behind the lens, as proposed by Tocci et al.<sup>1</sup>. Thus, aperture, integration time and sensor gain can be kept at identical settings in both cameras. This results in the same depth of field and motion blur, but different signal to noise ratios, which are used to enhance the dynamic range. Both cameras are adjusted mechanically for geometric alignment and the integration times of the camera-sensors are synchronized to record exactly the same fraction of time.

When used with longer focal lengths lenses, the 1:16 beam splitter generated ghosting artifacts caused by double reflections in the mirror glass. Hence some shots are acquired using an alternative setup shown in Figure 3b. It employs a 50% semitransparent mirror instead of the 1:16 mirror. In this case, a neutral density filter has to be mounted in front of one camera to shift the exposure of this camera.



Figure 3. Schematic drawing of the mirror-rigs used for HDR-video acquisition.

We chose to use the Alexa camera, a CMOS sensor based motion picture camera made by Arri. The sensor is operated in a dual gain mode resulting in a dynamic range (full-well/read-out noise) of 14.8 stops (89 dB). We adjusted the exposure to capture extended detail in dark and bright areas compared to a typical exposure with a single camera. Frame rates of 24fps and 25fps enable both TV and cinema applications. The integration time is set to 1/50 second (172,8°/180° shutter) to achieve a cinematic look. Only the high-speed shots are recorded with a 356° shutter to gain one additional stop of exposure.

During in-camera image processing, the signals from the sensor are converted from analog to digital on two paths with different levels of analog amplifications. The shadows are reconstructed from the high gain path and the highlights from the low gain path<sup>13</sup>. The resulting 16bit 2880x1620 resolution RAW Bayer pattern image is then converted to 1920x1080 RGB-pixels and coded in 12 bit LogC wide gamut color space<sup>6</sup> to be recorded as QuickTime file using near visually lossless 330Mbit/s ProRes<sup>14</sup> intra-frame compression.

#### 2.2. Postproduction

In postproduction, the highlight-preserving image is aligned to the lowlight-preserving image to increase spatial fit. Depending on the accuracy of the camera alignment, either a homography transform, or warping through local disparity estimation<sup>15</sup> is applied to the highlight-preserving image. Subsequently the colors of this image are matched to the lowlight-preserving image through multiplication of the individual color channels. Then the two images are merged to one HDR-frame, by blending between the lowlight-preserving image and the highlight-preserving image, depending on the brightness of the individual pixels. Finally the border pixels are set to black, to mask pixels where no highlight pass is available due to the spatial displacement. See Figure 4 for an overview of the in-camera and postproduction image-processing pipeline.



Figure 4. Overview of the HDR-video processing pipeline.

All postproduction steps are carried out in a color space spanned by the wide gamut primaries listed in Table 1. The final OpenEXR files are also coded in respect to these primaries. See the supplementary material on the project website<sup>16</sup> for precompiled conversion matrices from Alexa wide gamut to sRGB and CIE 1931 XYZ.

	Red	Green	Blue	White
х	0.6840	0.2210	0.0861	0.3127
у	0.3130	0.8480	-0.1020	0.3290

Table 1. CIE 1931 chromaticity coordinates of the Alexa wide gamut color space primaries.

During postproduction, the rendered image sequences were screened using Filmlight's Baselight color grading system and a Dolby PRM-4200 HDR monitor. The images were displayed three times with exposure offsets -6, 0 and 6 stops to be able to see the full dynamic range that was captured.

#### **3. RESULTS**

The resulting scenes were staged and recorded at the Stuttgart Media University (HdM) between January and October 2013. We planned five categories of scenes, to focus on the HDR-challenges inherent to different types of film projects, e.g. documentary, advertising or A-movie. The used lenses and the amount of lighting and makeup are corresponding to the resources typically available in the respective productions. As an example, dolly grip was just used in sequences that are representative for movie or advertising shootings. In simulated documentary shots like "Bistro", only a reduced amount of makeup was applied.

In the following paragraphs, the visual content and the major technical specifications of all scenes are summarized. See the supplementary material on the project website<sup>16</sup> for additional technical details like the T-stops used and the choice of lenses.

## 3.1. Wide Gamut and Moving Lights

An annual fair is filmed on location to provide color-saturated highlights and fast moving colorful objects. The saturated lights are dominant light sources that illuminate the actors with different colorful shades. The overall brightness and color of the scenes change very fast, both outside on the fair, as well as inside the beer-hall. To provide even more light changes, the scenes are edited together as a sequence with multiple cuts.

Carousel Fireworks	2536 frames 25fps	Night / Exterior Establishing Longshot	Crowded street on an annual fair with illuminated fun rides in the background
		Fullshot	Moving carousel with colored lights
AND MARTIN		Mediumshot	Girl watching carousel
		High Angle Longshot	Fireworks
	"Carousel Fireworks" is a sequence of shots acquired under available light at an annual fair during the night. The distinction of this scenery is to present colorful self illuminated objects and dark surroundings at the same time. Changing colored light sources illuminate the scenery including cloth and skin tones of the actors. The moving carriages of the carousel create blurred light sources that are both filmed in standard speed and slow-motion. The firework provides bright colored highlights glittering against a uniform black sky with the moon in frame. Later, the contrast gets reduced through smoke introduced by the firework.		
Beerfest Lightshow	2035 frames	Night / Interior	Crowded hall with fast moving lights.
	25 fps	Multiple Longshots	
	"Beerfest Lightshow" is filmed on location in a moody, smoky beer-hall while a lightshow is performed. This lightshow includes various kinds of fast switched and moving lights that send out bright and colorful light beams. Additionally, a mirror bowl reflects neutral colored light beams. Laser beams flash up and strobe- lights as well as blinder-lights brighten up the scenery temporarily. The cameras were exposed to capture a compromise between lowlight detail and highlights. As a result, some highlights are clipped.		

Table 2. Overview of scenes providing wide gamut stimuli and moving lights.

## 3.2. Low Key Scenes

As in the preceding table, the following scenes are also filmed in low-key with flickering light sources on location. But in this case, the light sources are mostly black body radiators covering a wide range of color temperatures.

The mood of the sceneries is dominated by the natural single light sources in combination with dim low-key sceneries.

Fireplace	952 frames 24 fps	Dawn / Exterior Fullshot Night / Exterior Medium Fullshot	Tilt down from defocused branches to group of persons at campfire. Persons standing beside and behind flames at a fire site stoking up the fire	
	The "Fireplace" scene offers a flickering light source with fast moving flames in front of a dim surrounding at dawn. The warm light of the campfire illuminates the persons that are surrounded by snowy scenery. The fire provides a strong color contrast to the bluish ambient light at dawn. Moving torchlights and flying sparks against the dark background at night provide high contrasts combined with fast movements.			
Smith Welding	1102 frames	Mixed Light /	A smith creates a light arc and flying sparks	
	25 fpsInterior, Fullshotby welding iron.In the dark mixed light of a blacksmith's shop, the low-key scenery is brightened up by an intense, fast moving point light source. The bluish welding arc is reflected on various textures of the scenery and forms a color contrast between the yellowish spraying sparks and the warm fire in the background.			
Smith Hammering	467 frames 25 fps	Mixed Light / Interior, Fullshot	A smith carries a forging blank from a fire to his anvil. Sparks fly when the iron is pounded.	
	Hammering incandescent iron at red heat creates bright spraying sparks that dominate the scenery in front of a dark background. In this setup, the fire and the forging blank were heated up to 800°C (1470°F).			

Table 3. Overview of low-key scenes with firelight.

## 3.3. Sunlight Scenes

The "Sunlight Scenes" represent typical conditions of documentary filming. They are captured in the field under natural light conditions and regular sunlight. Hence their dynamic range exceeds the latitude of our rig. We opted to partly leave the framed sun orb and specular highlights in the clipping range, in order to save details in the shades.

A travelling camera accentuates the appearance of textures. This can be seen on natural objects in the landscape scenery, ("Fishing") and synthetic materials in an environment of architecture ("Cars").

In both longshots of the following table, the brightness of the scene changes substantially over time, to provide a challenge for temporal tone mapping operators.

Fishing Longshot	834 frames 25 fpsSunrise / Exterior LongshotA fisherman stands in front of a lake. The camera is travelling towards him."Fishing Longshot" represents a typical sunrise in nature. It shows scenery with a lake, trees and a fisher in front of a sunny sky. The rising sun is just coming up and shines on surfaces like wood and moving water. Shafts of sunlight illuminate the morning haze in soft gradients and are mirrored on the water while specular highlights contrast to dark foreground objects.			
Fishing Closeshot	371 frames       Sunrise / Exterior       A fishhook is thrown into a lake and pulled         25 fps       Closeshot       out.         In "Fishing Closeshot", sunlight reflections are glistening on the water surface         stimulated by a fishhook. The exposure is set to highlight protection excluding the reflection of the sun.			
Cars Longshot	820 frames 25 fpsDay / Exterior LongshotCars and flags on a stone paced plaza in front of a building made of steel and glass."Cars Longshot" is a backlit scene with the sun in frame, which generates specular reflections on architecture and cars. The light that comes from the back emphasizes textures of stone and synthetic materials like steel, glass and the flags. The camera pans from the bright visible sun to the dim entrance of a building. Blacks get richer and contrast increases throughout the shot, as stray light is reduced at the end of the camera pan. The sun orb was clipped intentionally to retain details in the blacks.			
Cars Fullshot	442 frames 25 fpsDay / Exterior FullshotBlack car stands on plaza and is captured by a moving camera."Cars Fullshot" shows directional sunlight on a black car. The material appearance of the car finish, metal structures and glass windows are emphasized by a camera-travelling. The exposure is set to highlight-protection excluding the hotspots of specular reflections. Thus dark shades like under the car are preserved.			
Cars Closeshot	414 frames 25 fpsDay / Exterior CloseshotDetails of a standing black car are captured by a moving camera."Cars Closeshot" shows specular reflections of directional sunlight moving over the front of a car. The surface feel of the car finish and the glass windows are emphasized through a close framing. The exposure was set to highlight-protection excluding the hotspots of specular reflections of the sun.			

Table 4. Overview of sunlight scenes.

# 3.4. High Contrast Skin Tones

The following sequences focus on the reproduction of skin tones in different lighting situations, ranging from documentary ("Bistro") to a movie scene ("Poker"). They are set up under controlled lighting conditions in a studio. Skin tones are illuminated with very different intensities. These highlighted faces, hands, hair and textiles, are partially lit up and exposed to the maximum luminance available in the latitude of the recording device, but they are as well presented in marginal illumination, when actors move away from the light sources.

The mean luminance of the "Poker Travelling Slowmotion" scene changes massively over time by temporarily covering the brightest areas in the image. In contrast to that, the change of luminance in the "Showgirl 2" scene is accomplished by varying brightness and type of the illumination.

Bistro	969 frames 24 fps	Day / Interior Medium Fullshot Fullshot	A man sits at a table with the sun shining on him through a window. A waiter steps from shade into sunlight followed by a woman coming from the dark part of the room. She points at him with a gun.	
	through a wind sunlight at the scenery that re documentary f	a available-light situation, where the sun shines e light. This scenery combines local bright bistro-chamber. Thus it creates a high contrast ghting situation typically encountered in are staged to show skin tones, hair and textures partly in sunlight and partly in shade.		
Poker Fullshot	600 frames	Night / Interior	Gamblers sitting at poker table.	
	24 fpsFullshotIn the "Poker Fullshot" scene a poker club is set up to demonstrate extreme high- and low-lights in the same frame. The fine structured white tablecloth is lit up by a single source hydrargyrum medium-arc iodide lamp (HMI) and represents a high contrast to the dark room with many details in the shades.			
Poker Travelling Slowmotion	1947 frames	Night / Interior	Gamblers smoking and playing cards.	
	24 fpsMediumshot"Poker Travelling Slowmotion" is based on the same setup as the scene "Poker Fullshot", but recorded in slow-motion with 120 fps. The actors are framed closer and smoking cigarettes. An over-the-shoulder camera-movement covers the bright table and reveals it again. Apart from the table that is again illuminated by a single-source HMI-light, candles can be seen in the background scenery.			
Showgirl 1	776 frames	Night / Interior	Actress sitting in front of a makeup-mirror.	
	25 fpsCloseshotIlluminated by warm tungsten lighting."Showgirl 1" shows an actress sitting in front of a makeup mirror. Tungsten light bulbs illuminate her skin tones and create specular highlights on her costume, makeup, jewelry and other reflecting props. This glamorous mood is often intended in film productions to emphasize the beauty of an actress.			
Showgirl 2	341 frames 25 fps	Night / Interior Closeshot	Actress standing up from makeup-table while light changes from tungsten to daylight.	
	The "Showgirl 2" scene executes a light-change from tungsten light to bright stage-light from an HMI-lamp. Thus the skin tone of the actress is shown in two extreme lighting situations throughout one take. The dull feather boa serves as a diffuse white reference, whereas the glistening of the costume and jewelry is brighter than diffuse white.			

Table 5. Overview of scenes providing high contrast skin tones.

# 3.5. Still Life

The still life contains high contrast and standard skin tones, all in one reference image. This can be useful when comparing monitors or checking image processing pipelines.



481 frames<br/>25 fpsNight / Interior<br/>MediumshotA couple with dark and pale skin tones is<br/>standing behind a color-checker and a<br/>transmittance gray scale.A couple with dark and pale skin tones and clothes is lighted by means of a high<br/>contrast backlight. This provides bright highlights on skin, hair and clothes as<br/>well as dark areas. The square above the gray scale on top of the Ulbricht sphere<br/>represents a black/stray-light reference with virtually no radiance emitted.

Table 6. Still life HDR test image scene.

## **4. LIMITATIONS**

Recording HDR video by means of a mirror-rig makes it possible to capture a dynamic range, that single sensor cameras will probably only be capable to capture in the future. But using a mirror-rig comes along with significant limitations. The handling of the fully rigged recording device is very limited as it weights about six times more than a single digital film camera. Additionally it has to be kept wired to a recording unit. Therefore the rig cannot be placed as flexible as a single camera.

Due to the large mirror, the rig suffers from stray light and lens flares. Even if these could be reduced using a mirror with a higher-grade coating, stay light and lens flares will always be more severe due to the much bigger matte box.

Further artifacts introduced through the mirror are double contours in the highlight-protecting pass of long shots that can be seen in "Poker" or "Showgirl 2". These double contours result from a second reflection of the transmitted light rays when leaving the glass of the mirror in the direction of the lower camera. They could be avoided using thinner beam splitters.

Besides the double contours, the mirror also introduces polarizing artifacts. This means, that polarized light, like from the water surface in the "fishing" shots, or the car finish in the "car" shots, is split with another ratio than the non-polarized light from the surrounding. This results in diverging brightness and colors between the different areas in the image depending if they are reconstructed from the lowlight-preserving camera or the highlight-preserving camera.

In daylight shots, specular highlights and the sun orb are often clipped because lowering the exposure to capture them would have resulted in loosing too much detail in the blacks. While this doesn't present a problem for the intended use of the video data set, true radiance maps cannot be recovered using a setup consisting of two motion picture cameras that are shifted by 4 stops in exposure.

Finally the sensor of the Alexa camera is a rolling shutter sensor. Rolling shutter artifacts can be observed at very fast motion, e.g. frame 97031 from the "Fireworks" scene.

#### **5. CONCLUSION**

We present a cinematic wide gamut HDR-video test set, designed for the evaluation of temporal tone mapping operators and HDR-displays. The scenes are staged according to common film production techniques and captured using two stateof-the-art motion picture cameras mounted on a mirror-rig. The settings have been chosen to present challenges to HDRpresentation such as bright, saturated colors and brightness changes of different speed and magnitude. Human faces including high contrast skin-tones, hair and eyes complete the cinematic settings. As a consequence, this test-set can be employed to compare the rendering of material appearance and colors on different HDR-displays or to evaluate temporal video tone mapping operators. Due to the cinematic look and staging, user studies that rely on image quality assessment by non-specialist viewers can be conducted without causing irritations due to non-staged reproductions.

Preliminary screenings of the sequences on a Dolby PRM-4200 HDR-display show a more realistic appearance of material and enhanced perception of space, compared to viewing the images on a standard BT.1886 display. Therefore, these image sequences may also serve as a basis for further investigations on HDR-video perception such as the evaluation of the impact and the acceptance of HDR-footage to a normal audience.

All clips can be downloaded at the project website<sup>16</sup>.

#### **AKNOWLEDGEMENTS**

We would like to thank the companies Arri, C-Motion, Dolby, Filmlight, Image Engineering and Zeiss for lending us their equipment. In addition to that, we are very grateful for the enthusiastic support from the students in the HDRproject group at Stuttgart Media University (HdM) who helped to shoot the sequences and perform the postproduction. Among these we would like to especially point out Jascha Vick and Jonas Trottnow for their superior engagement during the Winter shoot.

This work was partly supported by "Kooperatives Promotionskolleg Digital Media" at Stuttgart Media University, University of Tuebingen and the University of Stuttgart.

#### REFERENCES

[2] Kang, S.B., Uyttendaele, M., Winder, S., and Szeliski R., "High dynamic range video," ACM SIGGRAPH, (2003).
[3] Krawczyk, G., "HDR Video Environment Maps Samples," Max-Planck-Institut für Informatik, 20 June 2006, <http://www.mpi-inf.mpg.de/resources/hdr/video/> (27 July 2013).

[5] Arnold & Richter Cine Technik GmbH, "Arri ALEXA Camera Technical Specification,"

<a href="http://www.arri.com/camera/digital">http://www.arri.com/camera/digital</a> cameras/cameras/camera details.html?product=251> (03 Jan 2014).

[6] Brendel, H., "ALEXA LogC Curve – Usage in VFX – Reference Manual", Arnold & Richter Cine Technik GmbH, Munich, Germany, (2012).

[7] ITU-R, "Recommendation ITU-R BT.1886. Reference electro-optical transfer function for flat panel displays used in HDTV studio production," International Television Union, Geneva, (2011).

[8] SMPTE, "RP 431-2, D-Cinema Quality – Reference Projector and Environment for the Display of DCDM in Review Rooms and Theaters," The Society of Motion Picture and Television Engineers, New York, (2011).

[9] Dolby Laboratories Inc., "Dolby PRM-4200 Professional Reference Monitor Whitepaper", Dolby Laboratories Inc., San Francisco, (2012).

[10] Seetzen, H., Makki, S., Ip, H., Wan, T., Kwong, V., Ward, G., Heidrich, W., Whitehead, L., "Self-calibrating wide color gamut high dynamic range display," Human Vision and Electronic Imaging XII, (2007)

[11] Debevec, P. E., and Malik, J., "Recovering high dynamic range radiance maps from photographs," ACM SIGGRAPH, (1997)

<sup>[1]</sup> Tocci, M. D., Kiser, C., Tocci, N., and Sen, P., "A versatile HDR video production system," ACM SIGGRAPH, (2011).

<sup>[4]</sup> Thorpe, L. J.. "Sony HDCAM: Exposure Latitude - Issues of Dynamic Range 24P Technical Seminar #3", Sony Electronics Inc (2001).

<sup>[12]</sup> Red Digital Cinema, "High Dynamic Range Video with HDRx," <a href="http://www.red.com/learn/red-101/hdrx-high-dynamic-range-video">http://www.red.com/learn/red-101/hdrx-high-dynamic-range-video</a> (3 Jan 2014).

<sup>[13]</sup> Andriani, S., Brendel, H., Seybold, T., Goldstone, J., "Beyond the Kodak image set: a new reference set of color image sequences," IEEE International Conference on Image Processing, (2013)

<sup>[14]</sup> Apple Inc., "Apple ProRes White Paper," Apple Inc., (2012)

<sup>[15]</sup> The Foundry, "Occula", <http://www.thefoundry.co.uk> (27 July 2013).

<sup>[16]</sup> Froehlich, J., et al., "HdM-HDR-2014 Project," < http://www.hdm-stuttgart.de/~froehlichj/hdm-hdr-2014/> (5 Feb 2014).