High-dynamic range video acquisition with a multiview camera

Jennifer Bonnard*, Céline Loscos, Gilles Valette, Jean-Michel Nourrit and Laurent Lucas

CReSTIC SIC, Université de Reims Champagne-Ardenne, 51100 Reims, France

ABSTRACT

We propose a new methodology to acquire HDR video content for autostereoscopic displays by adapting and augmenting an eight view video camera with standard sensors. To augment the intensity capacity of the sensors, we combine images taken at different exposures. Since the exposure has to be the same for all objectives of our camera, we fix the exposure variation by applying neutral density filters on each objective. Such an approach has two advantages: several exposures are known for each video frame and we do not need to worry about synchronization. For each pixel of each view, an HDR value is computed by a weighted average function applied to the values of matching pixels from all views. The building of the pixel match list is simplified by the property of our camera which has eight aligned, equally distributed objectives. At each frame, this results in an individual HDR image for each view while only one exposition per view was taken. The final eight HDR images are tone-mapped and interleaved for autostereoscopic display.

Keywords: HDR images, HDR videos, multiview, pixel matching, autostereoscopic displays.

1. INTRODUCTION

It is common nowadays to see adverts about high-definition, 3D displays and television, stereo cameras, or cameras with enhanced intensities and colors. While acquisition equipment and display technologies are getting more and more sophisticated, it is still difficult to find technology that combines all of the innovative components. We do one step towards this by proposing a methodology to create content that combines depth (3D) and high-dynamic range (HDR) video content.

The 3D video content is intended to be viewed without wearing glasses on autostereoscopic displays. This type of screens requires input coming from several viewpoints (e.g. five, eight and nine for our lab equipment), calibrated in their geometrical distances and positions and in their color.[1] It is possible to create virtual content from computer graphics software, but technology is not yet commercialised to acquire real scenes for these displays. Experimental dedicated multiview cameras[2] were also created to allow video stream input for these displays. They are built with eight aligned and synchronized objectives, controlled by a unique command. Content coming from these multiview cameras can also be used for stereo displays, by choosing two consecutive views for input. Each sensor of this camera acquires a 10 bits per component value, which is not enough to represent the perceived light intensity (above 20 bits) by a human eye.

HDR acquisition has been focused for a long time on static photographic cameras. HDR content[3] extends the traditional range of 8 bits per color component to reach up to perceptual range (20 bits) or beyond if the physical range (e.g., sun intensity) is needed. HDR acquisition devices and sensors are still very confidential and HDR content is still most of the time created by combining different input information into a single HDR image. Input data often come from images taken at different exposure times. Acquiring different exposures for a video sequence is a difficult problem. This was addressed differently by two different approaches. Kang et al.[4] modified the exposure time of their video camera at each frame, alternating darker and brighter images, interpolating HDR values between 3 consecutive frames. Nayar and Mistunaga[5] used a different approach acquiring different exposures in one frame by placing a mask with four different grey levels between the optics and the sensors of the camera. HDR values are then reconstructed by interpolation between neighboring pixels.

Implementing one of the previous solutions on our eight view cameras proved to be impossible. First, we did not have automatic control of the exposure time between two consecutive frames because we can not change

* jennifer.bonnard@etudiant.univ-reims.fr

Optics, Photonics, and Digital Technologies for Multimedia Applications II, edited by
Peter Schelkens, Touradj Ebrahimi, Gabriel Cristóbal, Frédéric Truchetet, Pasi Saarikko,
Proc. of SPIE Vol. 8436, 84360A · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.922789

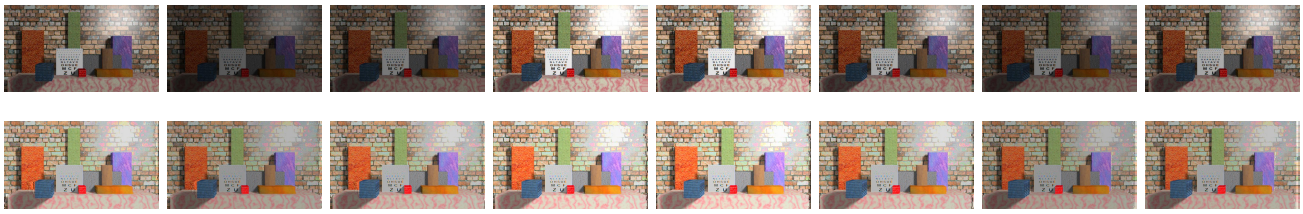


Figure 1. 1st row: original images with different exposures on each view, 2nd row: HDR images generated for each view.

exposure time and keep the video function of our camera. Second, we have no access to the hardware, preventing us from inserting masks within the camera devices. Instead, we physically constrain the exposure of each objective by a neutral filter, thus fixing for each objective the percentage of light arriving to the sensors. For each acquired frame, we process the input images to create eight HDR images. In order to be displayed on an autostereoscopic display, these images are tone mapped so that every intensity tone can be viewed and are interleaved to produce a 3D HDR image sequence. With this methodology, we believe to be the first to propose a solution to create 3D HDR video input (see Fig. 1).

Our methodology covers any step from the acquisition to the image restitution. This paper explains each step, while focusing mainly on how the eight input images are processed for creating HDR content, based on pixel matching algorithms. The contributions of this paper are the following:

- adapt a multiview video camera to acquire a larger range of intensities by placing neutral filters;
- adapt a known pixel matching and disparity mapping method to accept exposure varying input;
- propose a new equation to create HDR values coming from several non-identical views.

The paper is organised as follows. Section 2 details related work in HDR acquisition methods and pixel matching methods. Section 3 presents the method overview and details pixel matching adaptation and HDR content reconstruction. Results are presented in section 4 before concluding.

2. RELATED WORK

The goal of our work is to generate 3D HDR videos. While work has been done independently on generating video content for 3D displays and HDR displays, to our knowledge, only one other method exists today that combines HDR and stereo video acquisition. Lin and Chang[6] take two images acquired with different exposure times at two different viewpoints and modify these images with the camera response curve to obtain images with the same exposure time. Then, they use a pixel matching method and construct the disparity map with a belief propagation method. This method limits itself to the generation of stereo HDR videos not allowing to create multiview HDR videos. Our approach is similar: we make compatible HDR creation with 3D extraction on multiview video frames. In the following we first review related work on HDR image acquisition, before explaining pixel matching and disparity mapping methods in more details.

2.1 HDR acquisition

For a long time, HDR acquisition focused on single image acquisition. We first review methods to build static HDR images from static and dynamic content, then we consider HDR video acquisition.

2.1.1 HDR image acquisition

Building an HDR static image from non-HDR static images is possible using the Debevec and Malik's[7] method that consists of combining several images acquired from the same viewpoint with different exposure times. A camera response function can be derived to associate the value of the incoming light to the value attributed to the pixel in order to ensure coherent color values between images.[7–9] Another way of proceeding is by fusing exposures,[10] selecting only "good" pixels instead of performing a weighted average.[7]

Moving people or objects in the scene during the single viewpoint multi exposure acquisition process is a very important problem. Approaches oppose themselves on wether or not the dynamic elements are in the final HDR image. For example Granados et al[11] detect moving objects or persons in the scene and reconstruct the entire background of the scene with all details. The final HDR image do not include objects in movement.

Other images may include part or all of the dynamic objects. Gallo et al[12] identify movement according to a reference image chosen among all the acquired images by computing an error estimate. The reconstructed HDR image accurately includes elements in movement. Grosch[13] presents a similar approach but with a different error estimate algorithm. Effort is made in this method to include all dynamic elements in the final HDR image. Other methods[14,15] use the likeliness of a pixel to belong to an image to include it in the final HDR value computation. A drawback of these methods is that there is no guarantee that the dynamic elements are in the final image. Movement detection can be used to identify regions in movement. In Jacobs et al[16] are presented two different approaches, one based on variance estimation, the other using local entropy more appropriate when little contrast exists between the elements in movement and the background. The detected dynamic elements are inserted in the final HDR image. A recent approach[17] allows movement detection based on mutual information and local warping to reconstruct HDR information for moving regions. Most of the previous methods make the assumption that a large part of the image is fixed and aligned (by aligned we mean that a point in the scene projects at the same pixel coordinates in all images). In our case, none of the pixels are aligned and we need a different method.

2.1.2 HDR video acquisition

Making HDR video by adapting standard video cameras was tried. Kang et al[4] capture one video in which the exposure time between two consecutive frames changes. This method could relate to ours, although movement is continuous from one frame to another. In their case optical flow is used to match pixels. Optical flow to match content on video frames was also used by Sand et Teller,[18] but only if content is acquired several times from identical viewpoints at different exposure times.

Experimental devices were proposed to acquire HDR videos directly as HDR cameras.[19,20] The camera developed by Tocci et al[19] can acquire three different exposure times at once. For this, prisms are placed inside the camera. The system is built such that each prism diffracts the incoming light at different intensities, thus simultaneously acquiring three images with different exposure times. The resulting HDR video is thus the combination of information from three different exposures taken simultaneously from a same viewpoint. The Spheron video camera[20] is actually built using sensors acquiring directly the HDR range. This camera is definitely the most performant HDR video camera nowadays delivering a 20-bit precision raw video stream. Unfortunately, this camera is still at an experimental stage, and is not yet ready for public use.

HDR acquisition with dynamic content or video input has been addressed with many different approaches in the last decade. However, besides Lin and Chang[6] who proposed a method for stereo pairs, multiview HDR video (more than two views) has not yet been addressed.

2.2 Pixel matching and disparity maps

In a multiview context, a major problem is to find matching pixels in all images. To apply methods previously presented, some pixel matching methods exist and their goal is to find all pixels in each image that represent the same 3D point of the scene. These methods are classified into two ways: global or local methods. In global methods, a cost function is calculated, so that its minimization improves the correspondence.

Several constraints are applied during the stereo matching to reduce the number of potential corresponding pixels, to choose between two plausible corresponding pixels, or to verify and even to eliminate correspondences. For example the epipolar constraint permits a 1D research on the epipolar line, uniqueness constraint forbids a pixel to belong to two different correspondents and order rank constraint makes it possible the choice between two candidates.

The minimization of the cost function requires the use of optimization methods such as dynamic programming,[21,22] simulated annealing, relaxation, graph cuts,[23,24] belief propagation,[25] Markov fields or genetic algorithms. Scharstein and Szeliski[26] have elaborated an evaluation of stereo matching methods and published a classification by result quality.

Figure 2. Our pipeline for 3D HDR video generation.

Niquin et al[24] is one of the few to have worked on a multiscopic camera that allows to explore the multiview domain. From input coming from a video camera equipped with eight objectives, they reconstruct a 3D scene using a pixel matching method based on graph cuts.

3. PROPOSED METHOD

3.1 Overview

Our method aims at visualizing a real scene with a high dynamic range on an autostereoscopic display from video input. Fig. 2 shows the different steps of our method:

1. **Camera calibration:** we use a specific camera with eight objectives. In order to use the properties of a simplified epipolar geometry, these objectives are aligned and equally distributed. However, it remains necessary to calibrate them in their position and color characterization. For that purpose we use methods available in the literature.[27]

2. **Multi acquisition:** the camera was built for a LDR acquisition of eight view in real time. Thus we had to find a solution to allow a multi exposure acquisition in order to calculate an HDR video. We have chosen to apply different neutral intensity filters on each objective, as detailed in section 3.2.

3. **Stereo matching:** to compute the HDR value of a pixel we need to identify the pixels representing the same 3D point of the scene on the other views. For that step, we adapted an approach from Niquin et al[24] as explained in section 3.3.

4. **Multi HDR images generation:** for each view we have to calculate an HDR image, based on the list of matching pixels. We reformulate the well known method of Debevec and Malik[3] to take into account the multiview aspect of the images. This reformulation is detailed in section 3.4.

5. **Multi LDR images generation:** due to the limitations of current displays, it is not possible to use the HDR images for visualization. It is first necessary to transform the eight HDR images into eight LDR images. For the moment we simply use one of the tone mapping algorithms available in the literature.[28,29]

6. **LDR compositing:** in order to be rendered on an autostereoscopic display, individual LDR images must be interleaved in a way depending of the characteristics of this display. We used Drago's algorithm[30] but others tone mappers could be considered.[31]

7. **3D visualization:** finally the calculated LDR composite image from tone-mapped HDR images of step 6 is displayed on the 3D screen.

3.2 Acquisition

As mentioned in section 2.1, one of the most popular methods to compose a high dynamic range image is to calculate an average image from several images taken at different exposure times from a same viewpoint.[3,7] However, in our case, depth reproduction requires the simultaneous acquisition of multiple shifted images, each corresponding to a different viewpoint. In order to calculate a 3D HDR image, the principle of our method is to add a different exposure time to this perspective shift. Thus, each image contributes to recreate not only the depth information, but also a higher dynamic range.

Instead of varying the exposure for each view, we have fixed the exposure variation by applying different neutral intensity filters on each objective. In practice, we use three pairs of 0.3, 0.6 and 0.9 filters, plus two views

without filters. The advantage of such an approach is twofold. First, for each video frame several exposures are known. Second, by using filters, we do not need to worry about synchronization issues since each frame takes the same time to be captured by the eight objectives of the camera. In order to improve the results of the stereo matching, the filters are positioned to minimize the intensity difference between two adjacent images as shown in Table 1. In particular, a 0.9 density filter next to an objective without filter was avoided.

Table 1. Allocation of the different filters to each objective of the camera with the corresponding simulated exposure time.

View k	1	2	3	4	5	6	7	8
Filter density	0.3	0.9	0.6	-	-	0.3	0.9	0.6
Percentage of irradiance transmission	50%	12.5%	25%	100%	100%	50%	12.5%	25%
Real exposure time (ms)	40	40	40	40	40	40	40	40
Simulated exposure time Δt_k (ms)	20	5	10	40	40	20	5	10

3.3 Stereo matching

In classical multi exposure techniques, the values of pixels at the same coordinates in different images are averaged to calculate the radiance value of the corresponding pixel in the HDR image. Because of the multiview context of our acquisition, we cannot simply use the pixels at the same coordinates. Indeed, because of parallax due to the difference between viewpoints, they may not correspond to the same 3D point of the scene. Thus we have to find the corresponding pixels before the HDR generation step, which is a typical stereovision problem.

We use the geometry-based approach developped by Niquin et al[24] which aims to calculate disparity maps in a multiview context with a simplified epipolar geometry. For that purpose they first set up an energy function E of the form

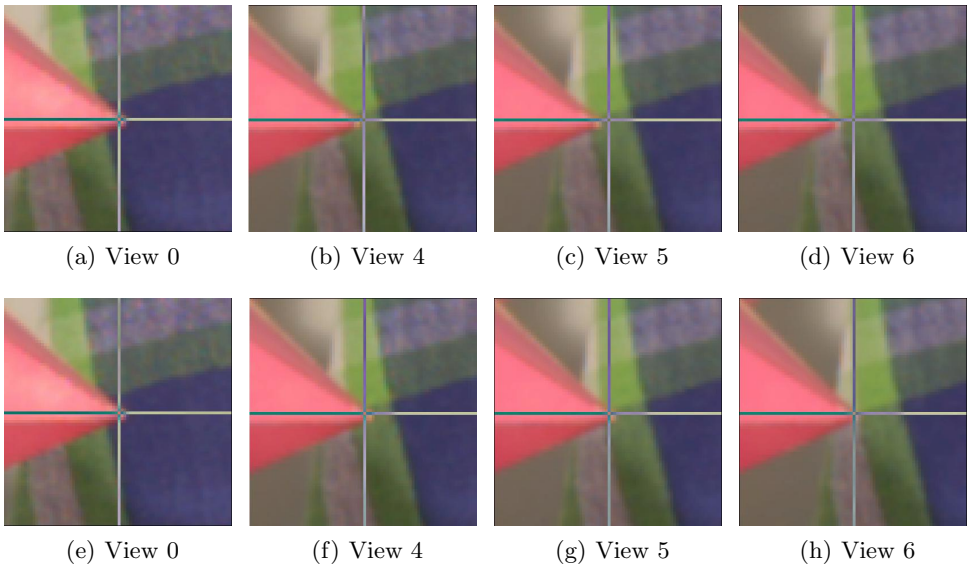
$$E(\mathcal{D}) = E^m(\mathcal{D}) + \sum_{k=1}^{N} E_k^s(d_k), \tag{1}$$

where N is the number of views and $\mathcal{D}$ is the set of all disparity functions d_k. E_k^s is the smoothness term, applied to each disparity function d_k, and it penalizes neighbouring pixels that are assigned to different disparities. E^m contains the dissimilarity cost, which measures the color dissimilarity between corresponding pixels, and the occlusion and consistency penalties, which are defined with the following rules. Let us consider a pixel p in image $\mathcal{I}_i$ and its corresponding pixel q in image $\mathcal{I}_j$. If $d_i(p) = d_j(q)$, p and q match, there is no penalty. If $d_i(p) < d_j(q)$, p is occluded and there is a penalty. Finally, if $d_i(p) > d_j(q)$, there is a consistency error and thus a penalty.

To obtain the best disparity map, the energy function defined in Equation (1) is subjected to the optimization achieved via graph-cuts, implemented on GPU for higher speed. One of the advantages of the method proposed by Niquin et al[24] is that it takes advantage of information from all views to ensure the geometric consistency of their disparity maps.

During the calculation of disparity maps, the pixels which are the projection of the same 3D point on the different images are grouped in a set called a "match". The list $\mathcal{M}$ of all the matches forms a partition of the set of pixels of all the images. This property will be useful to reformulate the equation of radiance estimation in the next section.

It is worth noticing that this stereo matching method is based on color similarity. Thus, prior to the computation of the list of matches, a factor based on the percentage of light reaching the sensor is applied to the images so that color intensity is similar on all images. Moreover, to measure the difference between two colors, we use the L_1 distance and we change the color space from RGB, which gives results of average quality, to LUV, which is one of the best-performing color space in global stereo matching, according to the work of Bleyer et al.[32] We have effectively noticed that the use of the LUV color space improves the pixel matching as shown in Fig. 3.



(a) View 0 (b) View 4 (c) View 5 (d) View 6

(e) View 0 (f) View 4 (g) View 5 (h) View 6

Figure 3. We have computed the matches for the "Moebius" scene (from Middlebury website[33]), with the RGB color space (1st row) and the LUV color space (2nd row). This figure shows at the intersection of the horizontal and vertical lines (added to the images for better visibility) some of the 7 pixels included in the same match for different views. With the LUV color space, the pixels are all at the tip of the branch of the star, which is not the case with the RGB color space.

3.4 HDR Generation

In order to compute the HDR images, we use the method of Debevec and Malik.[3] Assuming that the capturing device is perfectly linear, the radiance value $L_c(p)$ of a pixel p with spatial coordinates (x, y) in image $\mathcal{I}$ for each color channel $c \in \{R, G, B\}$ is given by the following equation:

$$\forall\ p = (i,j) \in \mathcal{I}, \quad L_c(p) = \frac{\sum_{k=1}^{N} w\left(Z_{c,k}(p)\right) Z_{c,k}(p)/\Delta t_k}{\sum_{k=1}^{N} w\left(Z_{c,k}(p)\right)}, \tag{2}$$

where N is the number of exposures, Δt_k is the exposure time of image k, $Z_{c,k}(p)$ is the value of the pixel p for the color channel c in the image k, and w a weighting function. The role of this function is to give higher weight to exposures in which the pixel's value is closer to the middle of the response function. We use the weighting function defined as:

$$w(v) = 1 - \left(\frac{2v}{255} - 1\right)^{12}, \tag{3}$$

where v is the pixel's value.

As we work with RAW images and we assume that the linearity property of the sensors is verified, we could apply Equation (2). Nevertheless, we have to modify this equation in order to take the multiview aspect of our images into account. Thus we do not consider a pixel in a given position (i, j), because this position in other images may not correspond to the same 3D point of the scene. Instead, we compute a radiance value for the pixels of a match, i.e. a set of corresponding pixels representing the same 3D point in the scene and thus having the same radiance value (under the assumption of a perfect Lambertian scene). Then the radiance value $L_c(p)$ of the pixels p of a match m for each color channel $c \in \{R, G, B\}$ is given by the following equation:

$$\forall\ m \in \mathcal{M}, \quad \forall\ p = (i,j,k) \in m, \quad L_c(p) = \frac{\sum_{q \in m} w(Z_c(q)) Z_c(q)/\Delta t(q)}{\sum_{q \in m} w(Z_c(q))}. \tag{4}$$

It is worth noticing that in this reformulation the pixel p contains not only its spatial coordinates (i, j) but also the index k of the image it belongs to. Thus, the exposure time $\Delta t(p)$ in2 this case represents Δt_k. Moreover,

as the radiance value is the same for all the pixels of a match, it is computed only once for each match. Finally, as the set of matches $\mathcal{M}$ is a partition of all the pixels of the N views, it is sufficient to compute the radiance values for all the matches for the three color channels to be able to compose N HDR images.

4. RESULTS

The method that we propose allows the generation of 3D HDR images but it is obviously not possible to print this kind of image thus we show results for only one view. We tested our method on several sets of images. We evaluated the quality of our result visually and by using the following objective criteria: the Similitude SIMilarity index (SSIM) and the Peak Signal to Noise Ratio (PSNR). For these criteria the comparison is made between one HDR image obtained with our method and one reference HDR image. The reference HDR image is generated applying a weighted average function on four images acquired with different exposure times from a same viewpoint of the same scene.

We used three different sets of images. The first set is composed by 7 "Moebius" images of size 800 x 750 pixels that can be found on the Middlebury's website.[33] The second set contains 8 images of size 748 x 422 pixels generated using the Persistence of Vision Raytracer (POV-Ray). Finally, the third set consists of 8 images of size 748 x 422 pixels acquired with our camera with filters applied such as shown in Table 1. We artificially modified the exposition in the two first sets of images so that is it corresponds to the exposure mentioned in Table 1. In the case od "Moebius" images, only one image expresses the full exposure (no filter applied).

We show in Fig. 4 some results obtained with our method. Fig. 4 (a)-(c) represent original view number 4 of the Middlebury's images and number 5 for the two other sets. Fig. 4 (d)-(f) present depth maps generated with the modified method of Niquin et al.[24] Fig. 4 (g)-(i) show HDR images for the view 4/5 generated using the proposed method detailed in section 3 and after applying the Drago's tone mapping algorithm[30] with identical parameters. We can visually observe in this figure that the results obtained for Moebius and POV-Ray images are of similar quality but problems arise on the set of camera images.

We calculated the SSIM index on each R, G and B channel between the pairs of images presented in the Fig. 5 tone-mapped by the Drago's algorithm[30] with identical parameters. We computed the PSNR directly on HDR images. Results are presented in Tab. 2. These results show that for calibrated images ("Moebius"), our method comes close to an image generated from one viewpoint. For the POV-Ray generated image, we intentionally added complexity to see how our system would perform. While visually results are convincing, measures show that the generated image is fairly different from the reference one. The challenge was similar for the camera images. Similarity is even lower but we suspect a poor calibration and a rough pixel matching. Moreover we do not yet treat the particular case of matches with only under- or overexposed pixels.

Table 2. Comparison of our method with the generated HDR image on a unique point of view.

	SSIM R	SSIM G	SSIM B	SSIM Mean	PSNR (dB)
Moebius images	82%	88%	83%	84%	60
POV-Ray images	63%	73%	63%	66%	56
Camera images	58%	58%	57%	58%	58

5. CONCLUSION

In this paper, we presented a new method for the generation of 3D HDR videos. To achieve our goal, we first fix a density filter on each objective of the camera to acquire images with different exposures which are used to compute an HDR image. We then construct lists of matches, which contain pixels that represent the same 3D point of a scene. Finally, we adapt the Debevec and Malik's method to construct one HDR image for each view using these matches. To evaluate our method, we tested different sets of images, including images acquired with our eight view camera and we have obtained promising results.

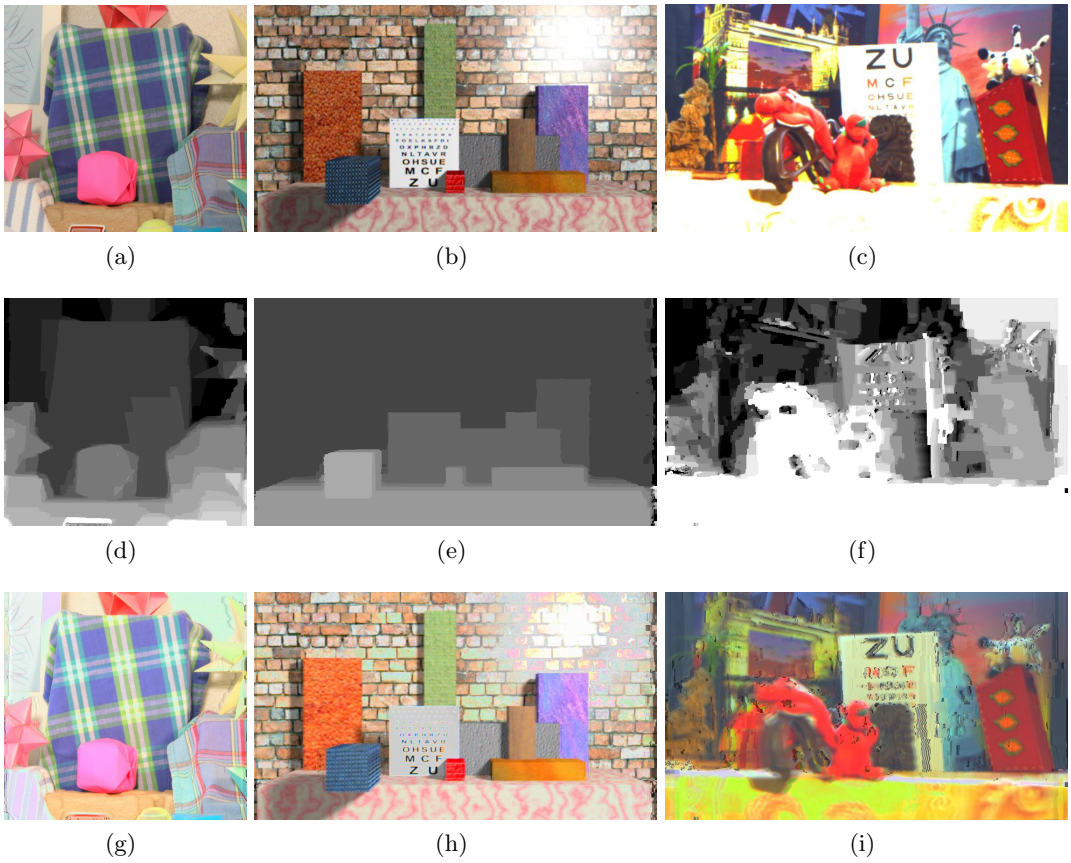


Figure 4. Results of our method for view 4 (1st column) and view 5 (2nd and 3rd columns). (a) Moebius[33] (b) POV-Ray image (c) Our camera. (a) - (c) reference images, (d) - (f) depth maps, (g) - (i) tone-mapped HDR images generated with our method.

This helped us in identifying cases where our method fails for which we intend to propose solutions in the future. We definitely need to handle lists of unique matches in order to provide valid HDR information for the final interleaved HDR image. We intend to explore interpolation methods to reconstruct the HDR information on over- or underexposed pixels for which we could not find correspondences. We are also still measuring the accuracy of the chosen approach for pixel matching. We do not exclude in the future to explore other ways of building the list of matches.

During our experiments, we found out that the camera sensors were very sensitive to intensity variations and it was necessary to often recalibrate the system. We would like to explore automatic calibration procedures to reduce errors due to poor color calibration between the camera objectives. Finally, the use of filters was guided by the type of equipment available in our lab. Cameras for which exposure time could be individually controlled could help adapting to a larger type of scenes, in particular scenes presenting a significantly larger or smaller range than the one covered by our filters. This may lead however to synchronisation problems that would need to be taken care of.

ACKNOWLEDGEMENTS

This work is funded by the regional project ESSAIMAGE 3D-HDR. Authors of this paper participate within the COST action HDRI IC1005.

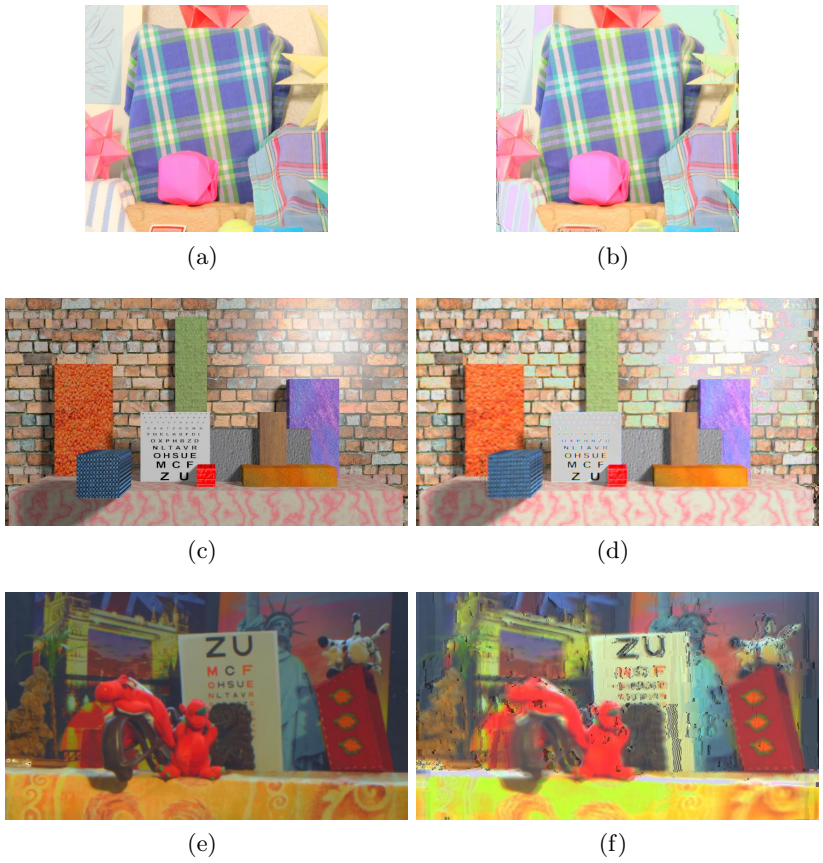


Figure 5. Tone-mapped HDR images. 1st column is the reference images and 2nd column images compute with our method. These images are used to calculate SSIM index and PSNR. (a)-(b) Moebius, (c)-(d) POV-Ray images and (e)-(f) camera.

REFERENCES

[1] Dodgson, N. A., "Analysis of the viewing zone of multi-view autostereoscopic displays," in [*Proceedings of SPIE - The International Society for Optical Engineering*], **4660**, 254–265 (2002).

[2] Prévoteau, J., Chalenon-Piotin, S., Debons, D., Lucas, L., and Remion, Y., "Multiview shooting geometry for multiscopic rendering with controlled distortion," *International Journal of Digital Multimedia Broadcasting* **2010** (2010).

[3] Reinhard, E., Ward, G., Pattanaik, S., Debevec, P., Heidrich, W., and Myszkowski, K., [*High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*], The Morgan Kaufmann series in Computer Graphics, Elsevier (Morgan Kaufmann), Burlington, MA, 2nd ed. (2010).

[4] Kang, S. B., Uyttendaele, M., Winder, S., and Szeliski, R., "High dynamic range video," *ACM Trans. Graph.* **22**, 319–325 (July 2003).

[5] Nayar, S. and Mitsunaga, T., "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], **1**, 472–479 (Jun 2000).

[6] Lin, H.-Y. and Chang, W.-Z., "High dynamic range imaging for stereoscopic scene representation," in [*Proceedings of the 16th IEEE International Conference on Image Processing*], *ICIP'09*, 4249–4252, IEEE Press (2009).

[7] Debevec, P. E. and Malik, J., "Recovering High Dynamic Range Radiance Maps from Photographs," *Computer Graphics* **31**(Annual Conference Series), 369–378 (1997).

[8] Mitsunaga, T. and Nayar, S., "Radiometric Self Calibration," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], **1**, 374–380 (Jun 1999).

[9] Mann, S. and Picard, R. W., "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in [*Proceedings of IS&T*], 442–448 (1995).
[10] Mertens, T., Kautz, J., and Reeth, F. V., "Exposure fusion," *Computer Graphics and Applications, Pacific Conference* , 382–390 (2007).
[11] Granados, M., Seidel, H.-P., and Lensch, H. P. A., "Background estimation from non-time sequence images," in [*Proceedings of Graphics Interface 2008*], *GI '08*, 33–40, Canadian Information Processing Society (2008).
[12] Gallo, O., Gelfand, N., Chen, W., Tico, M., and Pulli, K., "Artifact-free high dynamic range imaging," in [*Proceedings of the IEEE International Conference of Computational Photography (ICCP)*], (2009).
[13] Grosch, T., "Fast and robust high dynamic range image generation with camera and object movement," in [*Vision, Modeling and Visualization, RWTH Aachen*], 277–284 (2006).
[14] Pedone, M. and Heikkilä, J., "Constrain propagation for ghost removal in high dynamic range images," in [*VISAPP (1)*], 36–41 (2008).
[15] Khan, E. A., Akyz, A. O., and Reinhard, E., "Ghost removal in high dynamic range images," in [*IEEE International Conference on Image Processing*], 2005–2008 (2006).
[16] Jacobs, K., Loscos, C., and Ward, G., "Automatic High-Dynamic Range Image Generation for Dynamic Scenes," *Computer Graphics and Applications, IEEE* **28**(2), 84–93 (2008).
[17] Orozco, R. R., Martin, I., Loscos, C., and Vasquez, P. P., "Full high-dynamic range images for dynamic scenes," in [*Proceedings of SPIE vol. 8436, SPIE Photonics Europe: Optics, Photonics and Digital Technologies for Multimedia Applications, Session 3: High Dynamic Range Imaging*], (17-18 April 2012).
[18] Sand, P. and Teller, S., "Video matching," in [*ACM SIGGRAPH 2004 Papers*], *SIGGRAPH '04*, 592–599, ACM (2004).
[19] Tocci, M. D., Kiser, C., Tocci, N., and Sen, P., "A Versatile HDR Video Production System," *ACM Transactions on Graphics (TOG) (Proceedings of SIGGRAPH 2011)* **30**(4) (2011).
[20] "http://www.spheron.com."
[21] Wang, L., Liao, M., Gong, M., Yang, R., and Nister, D., "High-quality real-time stereo using adaptive cost aggregation and dynamic programming," in [*Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*], 798–805, IEEE Computer Society, Washington, DC, USA (2006).
[22] Zach, C., Sormann, M., and Karner, K., "Scanline optimization for stereo on graphics hardware," in [*Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*], 512–518, IEEE Computer Society, Washington, DC, USA (2006).
[23] Boykov, Y., Veksler, O., and Zabih, R., "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 1222–1239 (Nov. 2001).
[24] Niquin, C., Prévost, S., and Remion, Y., "An occlusion approach with consistency constraint for multiscopic depth extraction," *International Journal of Digital Multimedia Broadcasting (IJDMB), special issue Advances in 3DTV: Theory and Practice* **2010**, 1–8 (Feb. 2010).
[25] Sun, J., Zheng, N.-N., and Shum, H.-Y., "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 787–800 (July 2003).
[26] Scharstein, D. and Szeliski, R., "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision* **47**, 7–42 (2002).
[27] "http://opencv.willowgarage.com/documentation/camera_calibration_and_3d_reconstruction.html."
[28] Ledda, P., Chalmers, A., Troscianko, T., and Seetzen, H., "Evaluation of tone mapping operators using a high dynamic range display," *ACM Trans. Graph.* **24**, 640–648 (July 2005).
[29] Yoshida, A., Blanz, V., Myszkowski, K., and peter Seidel, H., "Perceptual evaluation of tone mapping operators with real-world scenes," in [*Human Vision & Electronic Imaging X, SPIE*], (2005).
[30] Drago, F., Myszkowski, K., Annen, T., and Chiba, N., "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum* **22**, 419–426 (2003).
[31] Kooima, R. L., Peterka, T., Girado, J. I., Ge, J., Sandin, D. J., and DeFanti, T. A., "A gpu sub-pixel algorithm for autostereoscopic virtual reality," *Virtual Reality Conference, IEEE* **0**, 131–137 (2007).

[32] Bleyer, M., Chambon, S., Poppe, U., and Gelautz, M., "Evaluation of different methods for using colour information in global stereo matching approaches," in [*Congress of the International Society for Photogrammetry and Remote Sensing*], 63–68 (2008).
[33] "http://vision.middlebury.edu/stereo."