Multiscopic HDR Image sequence generation

Raissel Ramirez Orozco
Group of Geometry and Graphics
Universitat de Girona (UdG), Spain
rramirez@ima.udg.edu

Céline Loscos
CReSTIC-SIC
Université de Reims Champagne-Ardenne (URCA), France

Ignacio Martin
Group of Geometry and Graphics
Universitat de Girona (UdG), Spain

Alessandro Artusi
Graphics & Imaging Laboratory
Universitat de Girona (UdG), Spain

ABSTRACT

Creating High Dynamic Range (HDR) images of static scenes by combining several Low Dynamic Range (LDR) images is a common procedure nowadays. However, 3D HDR video acquisition hardware barely exist. Limitations in acquisition, processing, and display make it an active, unsolved research topic. This work analyzes the latest advances in 3D HDR imaging and proposes a method to build multiscopic HDR images from LDR multi-exposure images. Our method is based on a patch match algorithm which has been adapted and improved to take advantage of epipolar geometry constraints of stereo images. Up to our knowledge, it is the first time that an approach different than traditional stereo matching is used to obtain accurate matching between the stereo images. Experimental results show accurate registration and HDR generation for each LDR view.

Keywords

High Dynamic Range, Stereoscopic HDR, Stereo Matching, Image Deghosting

1 INTRODUCTION

High Dynamic Range (HDR) imaging is an increasing area of interest at academic and industrial level, and one of its crucial aspects is the reliable and easy content creation with existing digital camera hardware.

Digital cameras with the ability to capture extended dynamic range, are appearing into the consumer market. They either use a sensor capable of capturing an intensity range larger than the one captured by traditional 8-10 bit sensors, or integrate hardware and software improvements to largely increase the acquired intensity range. However, due to their high costs, their use is very limited [BADC11].

Traditional low dynamic range (LDR) camera sensors provide an auto-exposure feature that can be used to increase the dynamic range of light captured from the scene. The main idea is to capture the same scene at different exposure levels, and then to combine them to reconstruct the full dynamic range.

To achieve this, different approaches have been presented [MP95, DM97, RBS99, MN99, RBS03], but they are not exempt of drawbacks. Ghosting effects may appear in the reconstructed HDR image, when the pixels in the source images are not perfectly aligned [TA+14]. This is due to two main reasons: either camera movement or objects movement in the scene. Several solutions for general image alignment exist [ZF03]. However, it is not straightforward to consider such methods because exposures in the image sequence are different, making alignment a difficult problem.

High Dynamic Range content creation is lately moving from the 2D to 3D imaging domain introducing a series of open problems that need to be solved. 3D images are displayed in two main different ways: either from two views for monoscopic displays with glasses or from multiple views for auto-stereoscopic displays. Most of current auto-stereoscopic displays accept from five to nine different views [LLR13]. To our knowledge, HDR auto-stereoscopic displays do not exist yet. We can feed LDR auto-stereoscopic displays with tone-mapped HDR, but we will need at least five different views.

Some of the techniques used for 2D applications have been recently extended for multiscopic images [TKS06, LC09, SMW10, BRR11, BLV+12, OMLA13, OMLA14, BRG+14, SDBRC14]. However, most of these solutions suffer from a common limitation: they need to rely on accurate dense stereo matching between images which may fail in case of different brightness between exposures [BVNL14]. Thus, more robust and faster solutions for matching different exposure images that allow an easy and reliable acquisition of multiscopic HDR content are highly needed.

(a) Non aligned   (b) Bätz *et al.* [BRG⁺14]   (c) Our Result
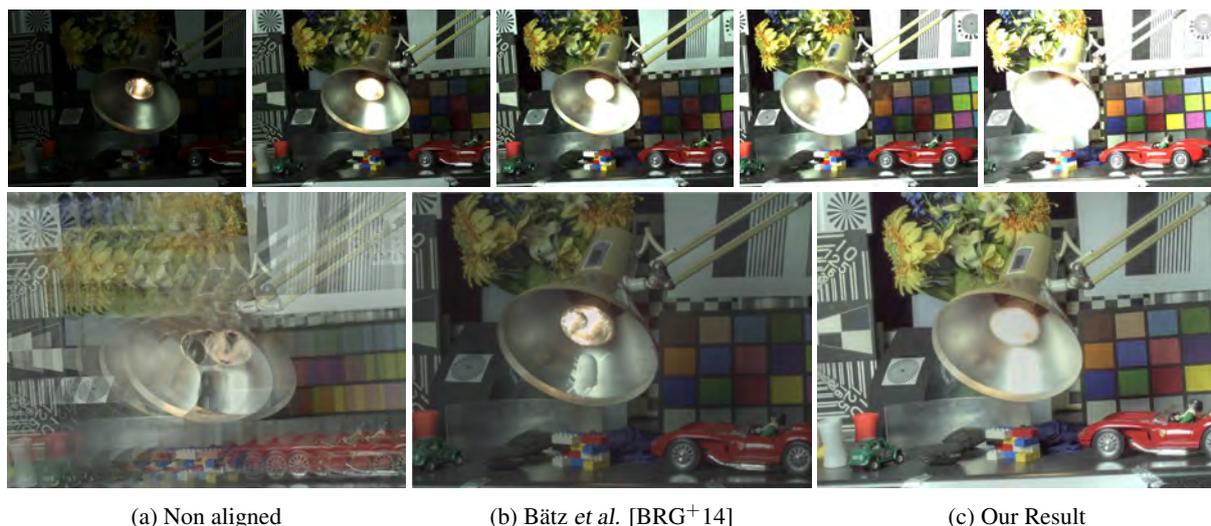
Figure 1: Set of LDR multiview images from the IIS Jumble data-set, courtesy of Bätz [BRG⁺14]. The top row shows five multiview exposure images, one exposure per view. The bottom row shows HDR images obtained without alignment (a), using Bätz's method (b) and using our proposed patch-match method (c).

In response to this need, we propose in this paper a solution to combine sets of multiscopic LDR images into HDR content using image correspondences based on the Patch Match algorithm [BSFG09]. This algorithm has been used recently by Sen *et al.* [SKY⁺12] to build HDR images that are free of ghosting effects. The need of improving the coherence of neighbour patches was already presented in [FP10].The results were promising for multi-exposure sequences where the reference image is moderately under exposed or saturated but it fails when the reference image has large under exposed or saturated areas.

We propose to adapt this approach for multiscopic image sequences (Figure 1), that answer to a simplified epipolar geometry obtained by parallel optical axes (images not originally taken with this geometric configuration can be later rectified). In particular, we reduce the search space in the matching process and improving the incoherence problem of the patch-match. Each image in the set of multi-exposed images is used as a reference; we look for matches in all the remaining images. These accurate matches allow to synthesize images corresponding to each view which are merged into one HDR image per view.

Our contributions into the field can be summarized as follows:

- We provide an efficient solution to multiscopic HDR image generation.

- Traditional stereo matching produce several artifacts when directly applied on images with different exposures. We introduce the use of an improved version of patch-match to solve these drawbacks.

- Patch-match algorithm was adapted to take advantage of the epipolar geometry reducing its computational costs while improves its matching coherence drawbacks.

2   RELATED WORK

Two main areas were considered in this work. The following section presents the main state of the art related to stereo HDR acquisition and multi-exposed image alignment for HDR generation.

2.1   Stereo HDR Acquisition

Some prototypes have been proposed to acquire stereo HDR content from multi-exposure views. Most approaches [TKS06, LC09, SMW10, Ruf11, BRG⁺14, AKCG14] are based on a rig of two cameras placed like a conventional stereo configuration that captures differently exposed images. Troccoli *et al.* [TKS06] propose to use cross correlation stereo matching to get a primary disparity match. The correspondences are used to calculate the camera response function (CRF) to convert pixel values to radiance space. Stereo matching is executed again but now in radiance space to extract the depth maps.

Lin and Chang [LC09] use SIFT descriptors to find correspondences. The best correspondences are selected using epipolar constrains and used to calculate the CRF. The stereo matching algorithm is based on belief propagation to derive the disparity map. A ghost removal technique is used to avoid artifacts due to noise or stereo mismatches. Even though, disparity maps are not accurate in large areas that are under exposed or saturated.

Rüfenacht[Ruf11] compares two different approaches to obtain stereoscopic HDR video content: a temporal

approach, where exposures are captured by temporally changing the exposure time of two synchronized cameras to get two frames of the same exposure per shot, and a spatial approach, where cameras have different exposure times for all shots so that two frames of the same shot are exposed differently.

Bonnard *et al.* [BLV⁺12] propose a methodology to create content that combines depth (3D) and HDR video for auto-stereoscopic displays. They use reconstructed depth information from epipolar geometry to drive the pixel match procedure. The matching method lacks of robustness especially on under exposed or saturated areas. Akhavan *et al.* [AYG13, AKCG14] offer a useful comparison of the difference between disparity maps obtained from HDR, LDR and tone-mapped images.

Selmanovic *et al.* [SDBRC14] propose to generate Stereo HDR video from a pair HDR-LDR, using an HDR camera and a traditional digital camera. In this case, one HDR view needs to be reconstructed. Three methods are proposed to generate an HDR image: (1) to warp the existing one using a disparity map, (2) to increase the range of the LDR view using an expansion operator and (3) an hybrid of the two methods which provides the best results.

Bätz *et al.* [BRG⁺14] present a framework with two LDR cameras, the input images are rectified before the disparity estimation. Their stereo matcher is exposure invariant and use Zero-Mean Normalized Cross Correlation (ZNCC) as a matching cost. The matching is performed on the gray-scale radiance space image followed by local optimization and disparities refinement. Some artifacts may persist in the saturated areas.

2.2 Multi-exposed Image Alignment

In the HDR context, most of methods on image alignment focus on movement between images caused by hand-held capture, small movement of tripods or matching moving pixels from dynamic objects in the scene. One of the main drawbacks for HDR video acquisition is the lack of robust algorithms for deghosting. Hadziabdic *et al.* [HTM13], Srikantha *et al.* [SS12] and Tursun *et al.* [TA⁺14] provide good reviews and comparisons between recent methods.

Kang *et al.* [KUWS03] proposed to capture video sequences alternating long and short exposure times. Adjacent frames are warped and registered to finally generate an HDR frame. Sand and Teller [ST04] combine feature matching and optical flow for spatio-temporal alignment of different exposed videos. They search for frames that best match with the reference frame using locally weighted regression to interpolate and extrapolate image correspondences. This method is robust to changes in exposure and lighting, but it is slow and artifacts may appear if there are objects moving at high speed.

Mangiat and Gibson [MG10] propose to use a method of block-based motion estimation and refine the motion vectors in saturated regions using color similarity in the adjacent frames of an alternating multi-exposed sequence.

Sun *et al.* [SMW10] assume that the disparity map between two rectified images can be modeled as a Markov random field. The matching problem is then posed as a Bayesian labeling problem in which the optimal values are obtained minimizing an energy function. The energy function is composed of a pixel dissimilarity term (using NCC as similarity measure) and a smoothness term which corresponds respectively to the MRF likelihood and the MRF prior.

Sen *et al.* [SKY⁺12] present a method based on a patch-based energy-minimization formulation that integrates alignment and reconstruction in a joint optimization. This allows to produce an HDR result that is aligned to one of the exposures and contains information from all the rest. Artifacts may appear when there are large under exposed or saturated areas in the reference image.

2.3 Discussion

Stereo matching is a mature research field; very accurate algorithms are available for images taken under the same lighting conditions and exposure. However, most of such algorithms are not accurate for images with important lighting variations. We propose a novel framework inspired by Barnes *et al.* [BSFG09] and Sen *et al.* [SKY⁺12]. We adapt the matching process to the multiscopic context resulting in a more robust solution.

3 PATCH-BASED MULTISCOPIC HDR GENERATION

Our method takes as input a sequence of LDR images (RAW or not). We transform the input images to radiance space, all the rest of steps are performed using radiance space values instead of RGB pixels. For 8-bits LDR images a CRF per camera needs to be estimated. An overview of our framework is shown in the diagram of the Figure 2. The first step is to recover the correspondences between the **n** images of the set. We propose to use a nearest neighbor search algorithm (see section 3.1) instead of a full stereo matching approach. Each image acts like a reference for the matching process. The output of this step is **n-1** warped images for each exposure. Which then are combined into an output HDR image for each view through a second step (see section 3.2).

3.1 Nearest Neighbor Search

For a pair of images I_r and I_s, we compute a Nearest Neighbor Field (NNF) from I_r to I_s using an improved version of the method presented by Barnes *et*
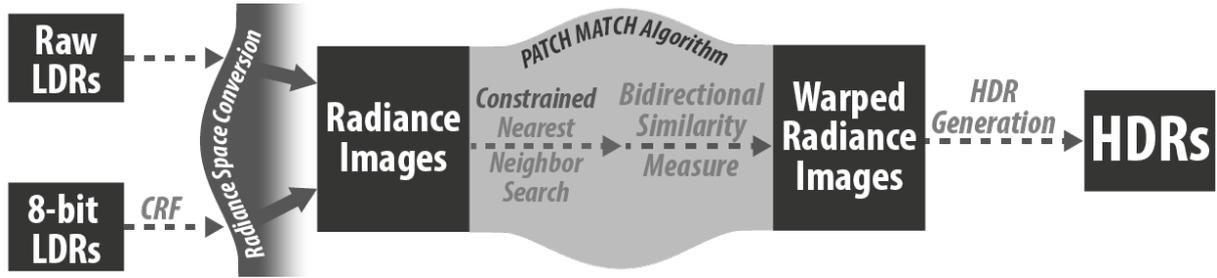
Figure 2: Proposed framework for multiscopic HDR Generation. It is composed by three main steps: (1) radiance space conversion, (2) patch match correspondences search and (3) HDR generation

al. [BSFG09]. NNF is defined over patches around every pixel coordinate in image I_r for a cost function **D** between two patches of images I_r and I_s. Given a patch coordinate $\mathbf{r} \in I_r$ and its corresponding nearest neighbor $\mathbf{s} \in I_s$, $NNF(\mathbf{r}) = \mathbf{s}$. The values of NNF for all coordinates are stored in an array with the same dimensions as I_r.

We start initializing the NNFs using random transformation values within a maximal disparity range on the same epipolar line. Consequently the NNF is improved by minimizing **D** until convergence or a maximum number of iterations is reached. Two candidate sets are used in the search phase as suggested by [BSFG09]: .

(1) *Propagation* uses the known adjacent nearest neighbor patches to improve NNF. It converges fast but it may fall in a local minima.

(2) *Random search* introduces a second set of random candidates that are used to avoid local minima. For each patch centered in pixel v_0, the candidates u_i are sampled at an exponentially decreasing distance from v_i:

$$u_i = v_0 + w\alpha^i R_i \qquad (1)$$

where R_i is a uniform random value $\in$ [-1,1], w is the maximum value for disparity search and α is a fixed ratio (1/2 is suggested).

Taking advantage of the epipolar geometry both search accuracy and computational performances are improved. Geometrically calibrated images allow to reduce the search space from 2D to 1D domain, consequently reducing the search domain. As an example, using random search we only look for matches in the range of maximal disparity in the same epipolar line (1D domain), avoiding to search in 2D space. This reduces significantly the number of samples to find a valid match.

Typical drawback of the original NNFs approach [BSFG09], used in the patch match algorithm, is the non geometrically coherency of its search results. This problem is illustrated in Figures 3 and 4. Two static neighbor pixels, in the reference image, match two separated pixels in the source image (Figure 3).
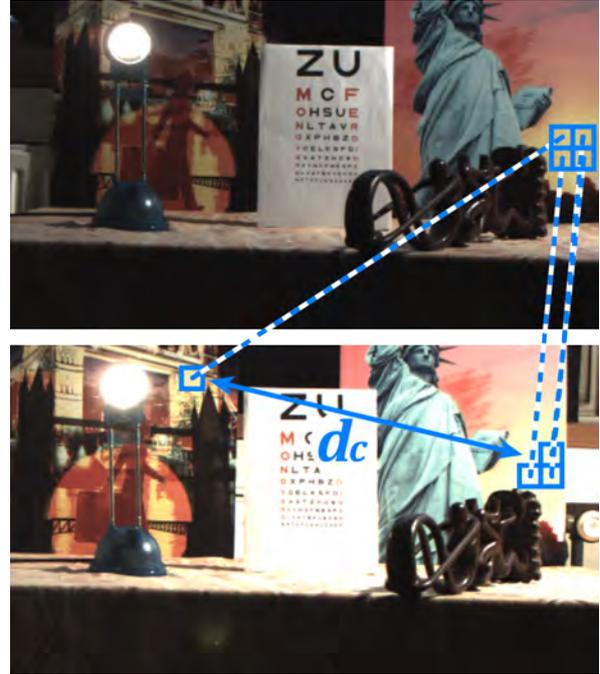

Figure 3: Patches from the reference image (Up) look for their NN in the source image (Down). Even when destination patches are similar in terms of color, matches may be wrong because of geometric coherency problems.

To overcome this drawback we propose a new distance cost function D by incorporating a coherence term to penalize matches that are not coherent with the transformation of their neighbors. Both Barnes *et al.* [BSFG09] and Sen *et al.* [SKY$^+$12] use the Sum of Squared Differences (SSD), described in equation 3 where **T** represents the transformation between patches of **N** pixels in images I_r and I_s. We propose to penalize matches with transformations that differ significantly form it neighbors by adding the coherence term **C** defined in equation 4. The variable d_c represents the Euclidean distance to the closest neighbor's match and Max_{disp} is the maximum disparity value. This new cost function forces pixels to preserve coherent transformations with their neighbors.

$$D = SSD(r,s)/C(r,s) \qquad (2)$$

$$SSD = \sum_{n=1}^{N} (I_r - T(I_s))^2 \qquad (3)$$

$$C(r,s) = 1 - d_c(r,s)/Max_{disp} \qquad (4)$$

(a) Src Image　　　　　　(b) Ref Image

(c) PM NNF　　　　　　(d) Ours NNF

(e) PM synthesized　　　　(f) Ours synthesized
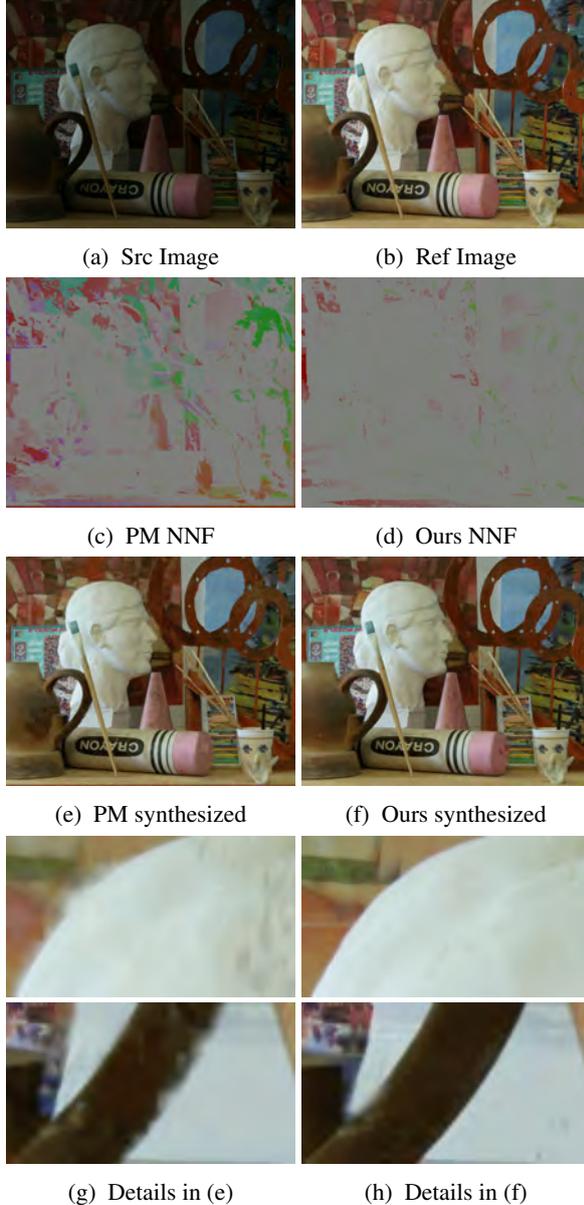
(g) Details in (e)　　　　(h) Details in (f)

Figure 4: Matching results using original Patch Match [BSFG09] (Left) and our implementation (right) for two iterations using 7x7 patches. Images in the 'Art' dataset courtesy of [vis06]

Figures 4c and 4e show the influence of the coherence problems described in Figure 3 in the matching results. Figures 4d and 4f correspond to the results including the improvements presented in this section. Figures 4c and 4d show a color representation of the NNFs using HSV color space, magnitude of the transformation vector is visualized in the saturation channel and the angle in the hue channel. Areas represented with the same color in the NNF color representation mean similar transformation. Objects in the same depth may have similar transformation. Notice that the original Patch Match [BSFG09] finds very different transformations for neighbor pixels of the same objects and produces artifacts in the synthesized image.

3.2 Warping Images and HDR Generation

The warping images are generated as an average of the patches that contribute to a certain pixel. Direct warping from the NNFs is possible, but it may generate visible artifacts as shown in Figure 5. This is due mainly to incoherent matches between the I_r and I_s images. To solve this problems we use Bidirectional Similarity Measure (BDSM) (Equation 5), proposed by Simakov *et al.* [SCSI08] and used by Barnes *et al.* [BSFG09], which measure similarity between pairs of images. It is defined for every patch $\mathbf{Q} \subset I_r$ and $\mathbf{P} \subset I_s$, and a number $\mathbf{N}$ of patches in each image respectively. It consists of two terms: *coherence* that ensures that the output is geometrically coherent with the reference and *completeness* that ensures that the output image maximizes the amount of information from the source image:

$$d(I_r,I_s) = \overbrace{\frac{1}{N_{I_r}} \sum_{Q \subset I_r} \min_{P \subset I_s} D(Q,P)}^{d_{completeness}} + \overbrace{\frac{1}{N_{I_s}} \sum_{P \subset I_s} \min_{Q \subset I_r} D(P,Q)}^{d_{coherence}}$$

$$(5)$$

(a) Direct warping　　　　(b) Using BDSM
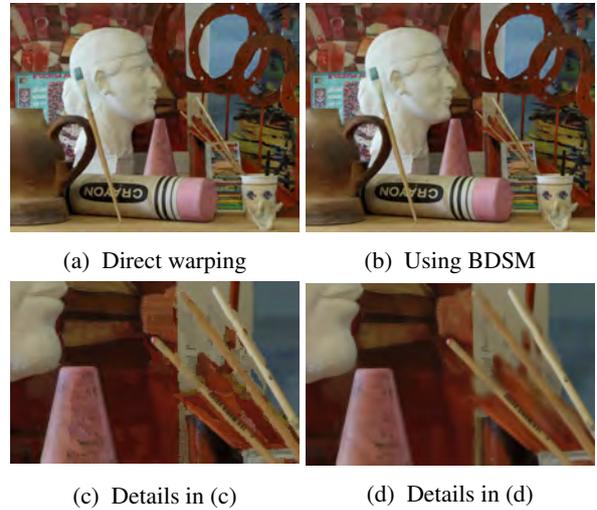
(c) Details in (c)　　　　(d) Details in (d)

Figure 5: Images 5a and 5b are both synthesized from the pair in Figure 4. Image 5a was directly warped using values only from the NNF of Figure 4c, which corresponds to matching 4a to 4b. Image 5b was warped using the BDSM of Equation 5 which implies both NNFs of Figures 4c and 4d.

This allows to improve both coherence and consistency by using bidirectional NNFs (from I_r to I_s and backward). It is more accurate to generate images using three iterations in each direction than only six from I_r to

I_s. Using BDSM also prevents artifacts in the occluded areas.

Since the matching is totally independent for pairs of images, it was implemented in parallel. Each image matches all other views. This produces **n-1** NNFs for each view. The NNFs are in fact the two components of the BDSM of equation 5. The new image is the result of accumulating pixel colors of each overlapping neighbor patch and averaging them.

The final HDR image per view is generated using a weighted average [MP95, DM97, MN99] as defined in Equation 6 and the weighting function of Equation 7 proposed by Khan *et al.* [KAR06]:

$$E(i,j) = \frac{\sum_{n=1}^{N} w(I_n(i,j))\left(\frac{f^{-1}(I_n(i,j))}{\Delta t_n}\right)}{\sum_{n=1}^{N} w(I_n(i,j))} \qquad (6)$$

$$w(I_n) = 1 - (2\frac{I_n}{255} - 1)^{12} \qquad (7)$$

where I_n represents each image in the sequence, w corresponds to the weight, f is the CRF, Δt_n is the exposure time for the Ith image of the sequence.

4   EXPERIMENTAL RESULTS

Five data-sets were selected in order to demonstrate the robustness of our results. For the set 'Octo-cam' all the objectives capture the scene at the same time and synchronized shutter speed. For the rest of data-sets the scenes are static. This avoids the ghosting problem due to dynamic objects in the scene. In all figures of this paper we use the different LDR exposures for display purposes only, the actual matching is done in radiance space.

The 'Octo-cam' data-set are eight RAW images with 10-bit of color depth per channel. They were acquired simultaneously using the Octo-cam [PCPD+10] with a resolution of 748x422 pixels. The Octo-cam is a multi-view camera prototype composed by eight objectives horizontally disposed. All images are taken at the same shutter speed (40 ms) but we use three pairs of neutral density filters that reduce the exposure dividing by 2, 4 and 8 respectively. The exposure times for the input sequence are equivalent to 5, 10, 20 and 40 ms respectively [BLV+12]. The objectives are synchronized so all images corresponds to the same time instant.

The sets 'Aloe', 'Art' and 'Dwarves' are from the Middlebury web site [vis06]. We selected images that were acquired under fixed illumination conditions with shutter speed values of 125, 500 and 2000 ms for 'Aloe'and 'Art' and values of 250, 1000 and 4000 ms for 'Dwarves'. They have a resolution of 1390 x 1110 pixels and were taken from three different views. Even if we have only 3 different exposures we can use the seven available views by alternating the exposures like shown in Figure 9.

The last two data-sets were acquired from two of the state of the art papers. Bätz *et al.* [BRG+14] shared their image data set (IIS Jumble) at a resolution of 2560x1920 pixels. We selected five different views from their images. They where acquired at shutter speeds of 5, 30, 61, 122 and 280 ms respectively. Pairs of HDR images like the one in Figure 6, both acquired from a scene and synthetic examples come from Selmanovic *et al.* [SDBRC14]. For 8-bit LDR data sets, the CRF is recovered using a set of multiple exposure of a static scene. All LDR images are also transformed to radiance space for fair comparison with other algorithms.

4.1   Results and discussion

(a) Src Image

(b) Ref Image

(c) PM NNF

(d) Ours NNF

(e) PM synthesized

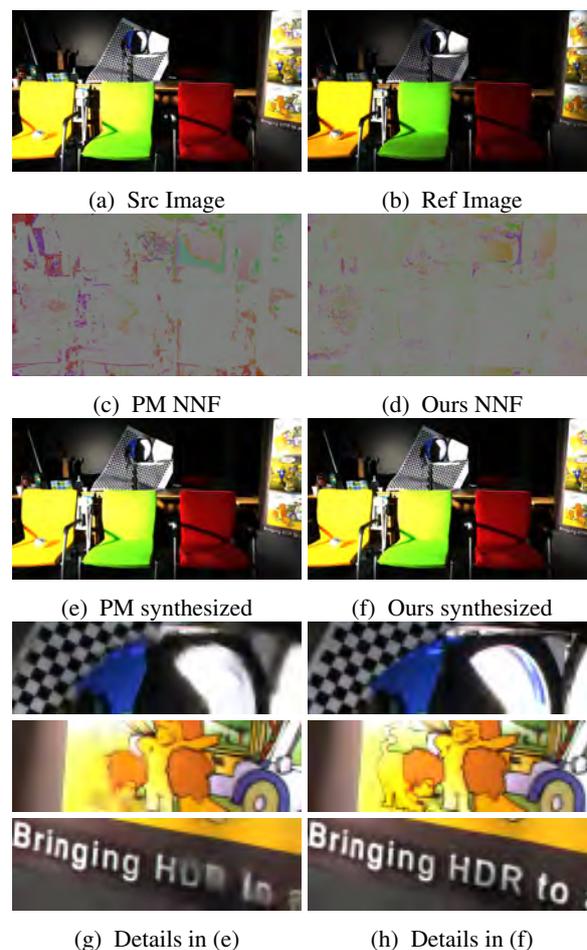(f) Ours synthesized

(g) Details in (e)

(h) Details in (f)

Figure 6: Comparison between original Patch Match and our implementation for two iterations using 7x7 patches. Images 6c and 6d show the improvement on the coherence of the NNF using our method. Images cortesy of [SDBRC14]

Figure 6 shows a pair of images linearized from HDR images courtesy of Selmanovic *et al.* [SDBRC14] and the comparison between the original PM from Barnes *et al.* [BSFG09] and our method including the coherence term and epipolar constrains. The images in Figures 6c and 6d represent the NNF. They are codified into an
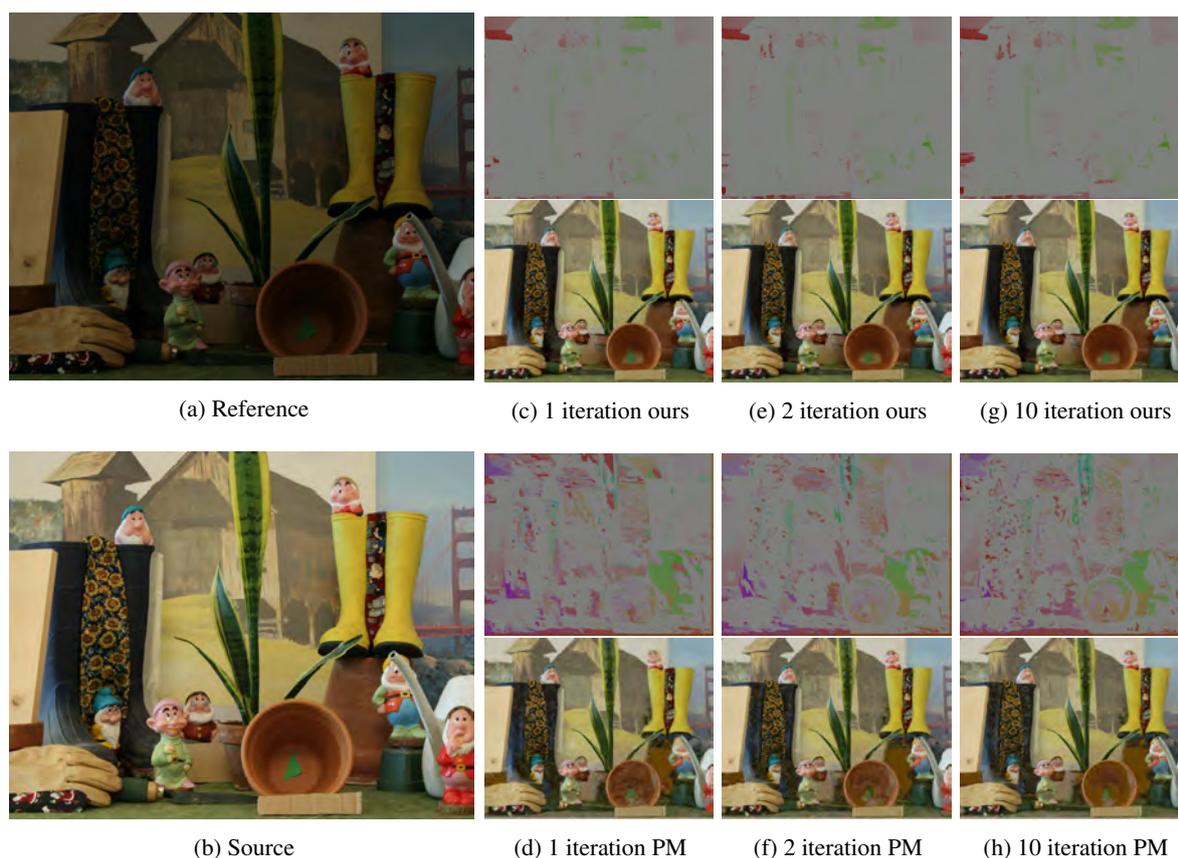
(a) Reference   (c) 1 iteration ours   (e) 2 iteration ours   (g) 10 iteration ours

(b) Source   (d) 1 iteration PM   (f) 2 iteration PM   (h) 10 iteration PM

Figure 7: Two images from the 'Dwarves' set of LDR multi-view images form Middlebury [vis06]. Our method with only two iterations achieve very accurate matches. Notice that the original patch match requires more iterations to achieve good results in fine details of the image.

image in HSV color space. Magnitude of the transformation vector is visualized in the saturation channel and the angle in the hue channel. Notice that our result represent more homogeneous transformations, represented in gray color. Images in Figure 6e and 6f are synthesized result images for the **Ref** image obtained using pixels only from the **Src** image. The results correspond to the same number of iterations (2 in this case). Our implementation converges faster producing accurate results in less iterations than the original method.

All the matching and synthesizing process are performed in radiance space. They were converted to LDR using the corresponding exposure times and the CRF for display purposes only. The use of an image synthesis method like the BDSM instead of traditional stereo matching allows us to synthesize values for occluded areas too.

Figure 7 shows the NNFs and the images synthesized for different iterations of both our method and the original patch match. Our method converges faster and produce more coherent results than [BSFG09]. In occluded areas the matches may not be accurate in terms of geometry due to the lack of information. Even in such cases, the result is accurate in terms of color. After

several tests, only two iterations of our method were enough to get good results while five iterations were recommended for previous approaches.

Figure 8 shows one example of the generated HDR corresponding to the lowest exposure LDR view in the IIS Jumble data-set. It is the result of merging all synthesized images obtained with the first view as reference. The darker image is also the one that contains more noisy and under-exposed areas. HDR values were recovered even for such areas and no visible artifacts appears. On the contrary, the problem of recovering HDR values for saturated areas in the reference image remains unsolved. When the dynamic range differences are extreme the algorithm does not provide accurate results. Future work must provide new techniques because the lack of information inside saturated areas does not allow patches to find good matches. The CRFs for the LDR images were calculated in a set of aligned multi-exposed images using the software RASCAL, provided by Mitsunaga and Nayar [MN99]. Figure 9 shows the result of our method for a whole set of LDR multi-view and differently exposed images. All obtained images are accurate in terms of contours, no visible artifacts comparing to the LDR were obtained.

(a) IIS Jumble data-set

(b) Lower exposure LDR   (c) Tone-mapped HDR

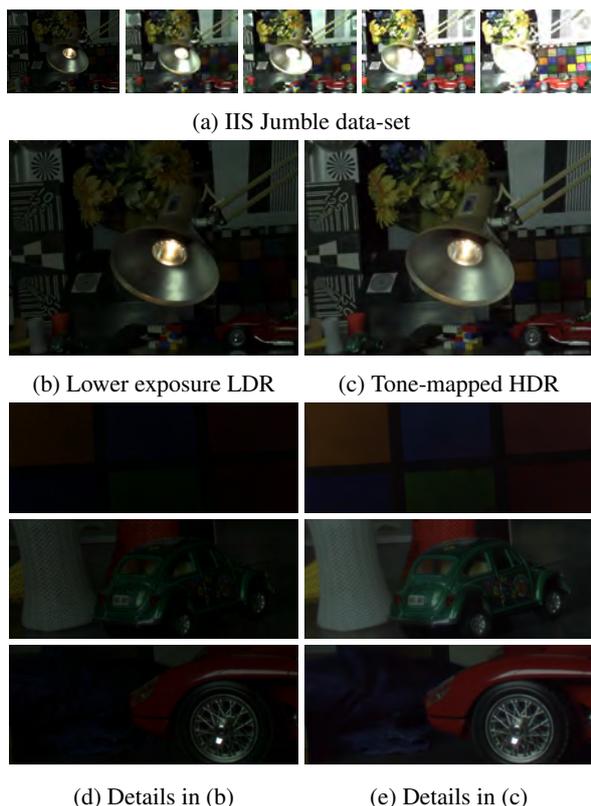(d) Details in (b)   (e) Details in (c)

Figure 8: Details of the generated HDR image corresponding to a dark exposure. Notice that under-exposed areas, traditionally difficult to recover, are successfully generated without visible noise or misaligned artifacts.

Figures 10 show the result of the proposed method in a scene with important lighting variances. The presence of the light spot introduce extreme lighting differences between the different exposures. For bigger exposures the light glows from the spot and saturate pixels not only inside the spot but also around it. There is not information in saturated areas and the matching algorithm does not find good correspondences. The dynamic range is then compromised in such areas and they remain saturated. Our method is not only accurate but faster than previous solutions. [SKY$^+$12] mention that their method takes less than 3 minutes for a sequence of 7 images of 1350x900 pixels. The combination of a reduced search space and the coherence term effectively implies a reduction of the processing time. In a Intel Core i7-2620M 2,70 GHz with 8 GB of memory, our method takes less than 2 minutes (103 $\pm$ 10 seconds) for the Aloe data set with a resolution of 1282x1110 pixels.

5 CONCLUSIONS

This paper presented a framework for auto-stereoscopic 3D HDR content creation that combines sets of multiscopic LDR images into HDR content using image dense correspondences. Methods that, when used for 2D domain cannot be used for 3D HDR content creation without introducing visible artifacts. Our novel approach is extending the well known Patch Match algorithm, introducing an improved random search function that takes advantage of the epipolar geometry. Also a coherence term is used for improving the matching process. These modifications allow to extend the original approach to work for HDR stereo matching, while improving its computational performances. We have presented a series of experimental results showing the robustness of our approach, in the matching process, when compared with the original approach and its qualitative results.

6 ACKNOWLEDGMENTS

REFERENCES

[AKCG14] Tara Akhavan, Christian Kapeller, Ji-Ho Cho, and Margrit Gelautz. Stereo hdr disparity map computation using structured light. In *HDRi2014 Second International Conference and SME Workshop on HDR imaging*, 2014.

[AYG13] Tara Akhavan, Hyunjin Yoo, and Margrit Gelautz. A framework for hdr stereo matching using multi-exposed images. In *Proceedings of HDRi2013 First International Conference and SME Workshop on HDR imaging*, Paper no. 8, Oxford/Malden, 2013. The Eurographics Association and Blackwell Publishing Ltd.

[BADC11] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers. *Advanced High Dynamic Range Imaging: Theory and Practice*. AK Peters (CRC Press), Natick, MA, USA, 2011.

[BLV$^+$12] Jennifer Bonnard, Celine Loscos, Gilles Valette, Jean-Michel Nourrit, and Laurent Lucas. High-dynamic range video acquisition with a multiview camera. *Optics, Photonics, and Digital Technologies for Multimedia Applications II*, pages 84360A–84360A–11, 2012.

[BRG$^+$14] Michel Bätz, Thomas Richter, Jens-Uwe Garbas, Anton Papst, Jürgen Seiler, and André Kaup. High dynamic range video reconstruction from a stereo camera setup. *Signal Processing: Image Communication*, 29(2):191 – 202,

Figure 9: Up: 'Aloe' set of LDR multi-view images from Middlebury web page [vis06]. Down: the resulting tone mapped HDR taking each LDR as reference respectively. Notice the coherence between all generated images.

Figure 10: Up: Set of LDR multi-view images acquired using the Octo-cam [PCPD+10]. Down: the resulting tone mapped HDR taking each LDR as reference respectively. Despite the important exposure differences of the LDR sequence, coherent HDR results are obtained. It is important to mention that highly saturated areas remain saturated in the resulting HDR.

2014. Special Issue on Advances in High Dynamic Range Video Research.

[BRR11] Michael Bleyer, Christoph Rhemann, and Carsten Rother. Patchmatch stereo - stereo matching with slanted support windows. In *Proceedings of the British Machine Vision Conference*, pages 14.1–14.11. BMVA Press, 2011. http://dx.doi.org/10.5244/C.25.14.

[BSFG09] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. Patchmatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), aug 2009.

[BVNL14] Jennifer Bonnard, Gilles Valette, Jean-Michel Nourrit, and Celine Loscos. Analysis of the Consequences of Data Quality and Calibration on 3D HDR Image Generation. In *European Signal Processing Conference (EUSIPCO)*, Lisbonne, Portugal, 2014.

[DM97] Paul Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *In proceedings of ACM SIGGRAPH (Computer Graphics)*, volume 31, pages 369–378, 1997.

[FP10] Yasutaka Furukawa and Jean Ponce. Accurate, dense, and robust multiview stereopsis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(8):1362–1376, Aug 2010.

[HTM13] Kanita Karaduzovic Hadziabdic, Jasminka Hasic Telalovic, and Rafal Mantiuk. Comparison of deghosting algorithms for multi-exposure high dynamic range imaging. In *Proceedings of the 29th Spring Conference on Computer Graphics*, SCCG '13, pages 021:21–021:28, New York, NY, USA, 2013. ACM.

[KAR06] Erum Arif Khan, Ahmet Oğuz Akyüz, and Erik Reinhard. Ghost removal in high dynamic range images. In *Image Processing, 2006 IEEE International Conference on*, pages 2005–2008, Oct 2006.

[KUWS03] Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. High dynamic range video. *ACM Trans. Graph.*, 22(3):319–325, 2003.

[LC09] Huei-Yung Lin and Wei-Zhe Chang. High dynamic range imaging for stereoscopic scene representation. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 4305–4308, 2009.

[LLR13] Laurent Lucas, Celine Loscos, and Yannick Remion. *3D Video from Capture to Diffusion*. Wiley-ISTE, October 2013.

[MG10] Stephen Mangiat and Jerry Gibson. High dynamic range video with ghost removal. *Proc. SPIE*, 7798:779812–779812–8, 2010.

[MN99] Tomoo Mitsunaga and Shree K. Nayar. Radiometric self calibration. *IEEE International Conf. Computer Vision and Pattern Recognition*, 1:374–380, 1999.

[MP95] Steve Mann and R. W. Picard. *On Being undigital With Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures*. Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology, 1995.

[OMLA13] Raissel Ramirez Orozco, Ignacio Martin, Celine Loscos, and Alessandro Artusi. Patch-based registration for auto-stereoscopic hdr content creation. In *HDRi2013 - First International*

Conference and SME Workshop on HDR imaging, Oporto Portugal, April 2013.

[OMLA14]   Raissel Ramirez Orozco, Ignacio Martin, Celine Loscos, and Alessandro Artusi. Génération de séquences d'images multivues hdr: vers la vidéo hdr. In *27es journées de l'Association française d'informatique graphique et du chapitre français d'Eurographics*, Reims, France, November 2014.

[PCPD⁺10]   Jessica Prévoteau, Sylvia Chalençon-Piotin, Didier Debons, Laurent Lucas, and Yannick Remion. Multi-view shooting geometry for multiscopic rendering with controlled distortion. *International Journal of Digital Multimedia Broadcasting (IJDMB), special issue Advances in 3DTV: Theory and Practice*, 2010:1–11, March 2010.

[RBS99]   Mark A. Robertson, Sean Borman, and Robert L Stevenson. Dynamic range improvement through multiple exposures. In *In Proc. of the Int. Conf. on Image Processing (ICIP 1999)*, pages 159–163. IEEE, IEEE, 1999.

[RBS03]   Mark A. Robertson, Sean Borman, and Robert L. Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12(2):219–228, 2003.

[Ruf11]   Dominic Rufenacht. *Stereoscopic High Dynamic Range Video*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, Agost 2011.

[SCSI08]   Denis Simakov, Yaron Caspi, Eli Shechtman, and Michal Irani. Summarizing visual data using bidirectional similarity. *IEEE Conference on Computer Vision and Pattern Recognition 2008 (CVPR'08)*, 2008.

[SDBRC14]   Elmedin Selmanovic, Kurt Debattista, Thomas Bashford-Rogers, and Alan Chalmers. Enabling stereoscopic high dynamic range video. *Signal Processing: Image Communication*, 29(2):216 – 228, 2014. Special Issue on Advances in High Dynamic Range Video Research.

[SKY⁺12]   Pradeep Sen, Nima Khademi Kalantari, Maziar Yaesoubi, Soheil Darabi, Dan B. Goldman, and Eli Shechtman. Robust patch-based HDR reconstruction of dynamic scenes. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2012)*, 31(6):203:1–203:11, November 2012.

[SMW10]   Ning Sun, H. Mansour, and R. Ward. Hdr image construction from multi-exposed stereo ldr images. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Hong Kong, 2010.

[SS12]   Abhilash Srikantha and Désiré Sidibé. Ghost detection and removal for high dynamic range images: Recent advances. *Signal Processing: Image Communication*, page 10.1016/j.image.2012.02.001, February 2012. 23 pages.

[ST04]   Peter Sand and Seth Teller. Video matching. *ACM Transactions on Graphics*, 23(3):592–599, August 2004.

[TA⁺14]   Okan Tarhan Tursun, Ahmet Oğuz Akyüz, , Aykut Erdem, and Erkut Erdem. Evaluating deghosting algorithms for hdr images. In *Signal Processing and Communications Applications Conference (SIU), 2014 22nd*, pages 1275–1278, April 2014.

[TKS06]   A. Troccoli, Sing Bing Kang, and S. Seitz. Multi-view multi-exposure stereo. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on*, pages 861–868, June 2006.

[vis06]   vision.middlebury.edu. Middlebury stereo datasets. http://vision.middlebury.edu/stereo/data/, 2006.

[ZF03]   Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003.