

# Exposure Stacks of Live Scenes with Hand-held Cameras

Jun Hu<sup>1</sup>, Orazio Gallo<sup>2</sup>, and Kari Pulli<sup>2</sup>

<sup>1</sup> Department of Computer Science, Duke University

<sup>2</sup> NVIDIA Research, Santa Clara, CA

**Abstract.** Many computational photography applications require the user to take multiple pictures of the same scene with different camera settings. While this allows to capture more information about the scene than what is possible with a single image, the approach is limited by the requirement that the images be perfectly registered. In a typical scenario the camera is hand-held and is therefore prone to moving during the capture of an image burst, while the scene is likely to contain moving objects. Combining such images without careful registration introduces annoying artifacts in the final image. This paper presents a method to register exposure stacks in the presence of both camera motion and scene changes. Our approach warps and modifies the content of the images in the stack to match that of a reference image. Even in the presence of large, highly non-rigid displacements we show that the images are correctly registered to the reference.

## 1 Introduction

The use of image stacks is gaining popularity among casual and amateur photographers because it allows the user to overcome some of the intrinsic limitations of standard cameras. One of the most common uses of such approaches is high-dynamic-range (HDR) imaging: when the contrast of the scene exceeds what the camera can capture, the user can take multiple pictures with different exposure settings, thus effectively acquiring different samples of the scene’s irradiance. Despite the advantages offered by these techniques, and the proliferation of dedicated apps for both point-and-shoot cameras and camera phones, the applicability of stack-based photography is still limited to a set of somewhat constrained scenarios: all these methods generally require that the images be perfectly registered. A small displacement of the camera between shots, a branch blown by the wind, the ocean’s waves, or people walking in the scene generally cause artifacts that can easily void the benefits of combining multiple images.

Several methods have been proposed that can deal with either camera motion [1–3] *or* dynamic scenes [4–6]—often at the cost of discarding some of the information—but produce sub-optimal results when both sources of artifacts

are present, as would happen with the stack in Figure 1<sup>3</sup>. Our objective is to preserve most of the available information from the different shots: instead of disregarding misaligned pixels, our approach warps and modifies the content of each image in the stack to better align it with an image of the stack that we pick as a reference.

We first attempt to map each pixel of each image in the stack to the corresponding pixel in the reference image. From the initial incomplete matches we learn the color transfer function, which captures and describes how the red, green, and blue intensities change across different shots, and use it to verify the matches. If we cannot find a match to every reference image pixel in the other images, possibly because they are occluded, or we do not want to use the matching pixels because they are too noisy or saturated, we copy the structure of the missing pixels either from the image we are registering or the reference, and blend them.

Our algorithm outputs a set of images aligned to each other at the pixel level. It is as if we froze the time at the reference frame, and created new versions of that same time instant with different camera settings. The results show that our method can easily deal with large object motions, significant camera displacement, and occlusions. Moreover it can deal with what can be referred to as the “branch nightmare”: the common, usually unavoidable problem of natural bodies, such as water or foliage, that move non-rigidly in the scene. In such regions, most approaches simply disregard all the images in the stack that do not match the reference; on the contrary, our method still combines information from as many images of the stack as possible.



**Fig. 1.** A stack of 4 images of a non-static scene (bottom row). The stack is affected by both camera and scene motion, as shown in (a), where the images are directly averaged together. The result obtained fusing the images after the proposed registration captures the whole range and is free of artifacts (b).

<sup>3</sup> The displacements between the original images, even when large, are often difficult to appreciate when performing a side-by-side comparison. We refer the reader to the paper’s website for comparisons between the images:  
[http://www.cs.duke.edu/~junhu/hdr\\_registration.html](http://www.cs.duke.edu/~junhu/hdr_registration.html).

## 2 Related Work

The creation of HDR images from sets of low-dynamic-range (LDR) shots, while seemingly straightforward, presents several challenges, in particular when the user is a casual photographer attempting to take multiple shots of a dynamic scene with a hand-held device such as a camera phone. The artifacts caused by misalignment of either moving camera or moving objects can be more unpleasant than the over- and under-saturation of any of the LDR shots.

Much research has focused on the combination of the LDR images, assuming perfect registration and completely static scenes. The pioneering work of Mann and Picard [7] and Debevec and Malik [8] has been refined and extended by recent work aiming at defining optimal weights based on the noise characteristics of the sensor (e.g., Akyüz and Reinhard [9], Hasinoff et al. [10], Granados et al. [11], Robertson et al. [12]). When the same pixel across the stack captures irradiance from different objects in the scene, whether because of camera motion or changes in the scene, it generates artifacts commonly referred to as *ghosting* because the averaging process produces transparent copies of the moving objects.

To address the problems caused by camera motion only, a solution as simple as that of using a tripod while acquiring the stack would work. However, having to carry a tripod strips the convenience of carrying compact cameras or camera phones fitting in a pocket. A more convenient solution is to register the LDR images after their capture. The registration algorithms need to be tuned specifically for HDR stacks as they have to be robust to loss of contrast in the regions that are not correctly exposed. A few methods have been proposed to address this problem under the assumption that the scene is static and that some kind of rigid transformation suffices. The original method proposed by Ward [1] focuses on translations only and was later extended to account for rotations [4]; Tzimiropoulos et al. find a similarity transformation working in the Fourier domain [2]. Tomaszewska and Mantiuk use SIFT descriptors and robust matching to find a global homography matrix [3]. Even for static scenes, all these methods are only approximate: a homography can only explain pure rotations of the camera around its optical center, or different views of a completely planar scene.

A different approach assumes that the LDR images are registered for camera motion and tackles only the problem of scene changes. A first class of algorithms detects regions that are likely to generate ghosting by looking at the variance of a pixel across the stack [6], looking at the entropy [4], or binarizing the images and isolating the pixels whose binary value changes [13]. They then replace such regions with a single exposure from the stack, thus limiting the effective captured range. Gallo et al. [5] attempt to overcome this problem by defining a reference image and blending together, for each region, only the LDR images that are consistent with it. Eden et al. [14] avoid the ghosting problem altogether by using a single LDR image for each part of the scene. Some of these approaches fail to completely remove the artifacts and, more importantly, they may discard valuable information.

A second class of algorithms modifies the fusion algorithm to account for potential sources of ghosting artifacts during the weighting process: Khan et al.

[15] use kernel estimation, while Zhang and Cham [16] use information about gradient magnitude and phase. These methods are likely to capture more information, but they may also just *reduce* the artifacts, as opposed to completely removing them. Moreover, Zhang and Cham rely on the assumption that a given region is affected by motion only in a small subset of the images in the stack: in other words, their method can fail if a pixel captures slightly different parts of the scene in each of the images of the stack, a common case in the presence of large or continuous, non-rigid motions.

A more elegant approach consists in searching for a non-rigid transformation between LDR images to address both camera motion and scene changes *at the same time*; this is usually performed by means of dense correspondences between the images. However, standard dense optical flow methods tend to fail due to the change of exposure between the different images in the stack, in particular when this causes saturation. Kang et al. [17] address this problem by first boosting the luminance values of the different images to compensate for the exposure change; they then rely on the brightness constancy assumption and use Lucas-Kanade method to compute optical flow. Zimmer et al. [18] take a similar approach but assume gradient constancy instead. We found that approaches based on optical flow are limited to cases where the motion between images is small. Moreover, unless the registration is performed in the linear irradiance domain, the assumption that the gradient is preserved across images in the stack is often too strong: aside from the obvious case of saturated regions, the non-linear transformation of the values of the pixels between different images in the stack violates it.

To overcome this limitation, our approach combines the advantage of warping based on dense correspondences, and that of blending small patches transformed and pasted from the original images. In Section 4 we show the benefit of this strategy by means of comparison with the state-of-the-art approaches.

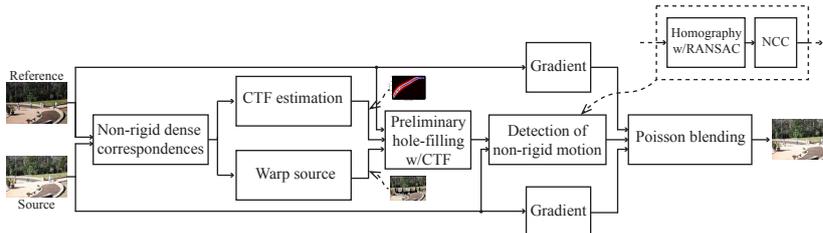
### 3 Method

Our algorithm attempts to preserve as much information from each image in the stack as possible. It does not require radiometric calibration and is thus suitable for approaches that fuse the images in the luminance domain (e.g., Exposure Fusion by Mertens et al. [19], which we use for the examples in this paper); it can however be easily adapted to work in the linear irradiance domain.

The workflow starts with the selection of a reference image, to which the other images in the stack are then aligned. We select it as the picture with the fewest over- and under-saturated pixels [17, 5]. We first describe the case of two images and we then explain how to extend it to larger stacks.

#### 3.1 Registering an image pair

Our method comprises several steps, which we explain thoroughly in the remainder of this section. Before diving into the details, however, we give an overview

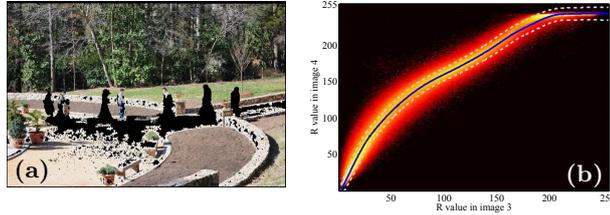


**Fig. 2.** The flow of the algorithm to register the source and reference images, see text.

of the method for additional clarity. With reference to Figure 2, given two images, source and reference, we first attempt to find dense correspondences using the technique by HaCohen et al. [20]; we then warp the pixels of the source for which a correspondence is found, and determine how the colors change between the two images. We use this color transfer function to detect wrong correspondences. The erroneous and incomplete matches result in holes in the warped source. We solve this problem in two steps: first, we use the mapping between the colors of the two images to synthesize the pixels’ values in such holes. This temporary image serves as a starting point for the rest of the computation. Then, we blend the gradients from the corresponding regions in the source image; we do so after checking that a homography can account for the motion of the pixels between the two images, and verifying that the region is consistent with the one synthesized before using normalized cross-correlation (NCC). For the regions in which pasting from the source is not possible, we blend the gradients from the reference. The resulting image is perfectly aligned with the reference by construction, and captures most of the information available in the source. We now proceed to explain the details of our approach.

The method proposed by HaCohen et al. is a powerful tool that can find matches across a wide range of appearance changes by applying geometric and photometric (from RGB to *Lab* + gradient magnitude) transformations to image patches, and by looking for local consistency of each match. We found that we can further improve its outcome by pre-multiplying one of the two images by the ratio of the exposure times of the pair. The algorithm by HaCohen et al. outputs the correspondences for many but not all of the source image pixels. Given a map of correspondences, we forward-warp source pixels to their reference destinations [21]. Most approaches for registration based on optical flow take the opposite route and perform backward warping [18, 22, 23]. Backward warping has the apparent benefit of not producing gaps when the flow field of the source pixels diverges, and can be done using the texturing hardware of GPUs. However, we do not want to automatically fill any gaps with potentially inconsistent data; at this stage we want them to remain explicitly open, as illustrated in Figure 3(a). The black areas failed to find a match because of occlusions and disocclusions due to moving objects, change of view point, or because they did not have a

sufficient amount texture to produce reliable matches. We address those areas in Section 3.2.



**Fig. 3.** (a) The points that are matched for one of the images from the stack in Figure 1 are warped to the corresponding locations in the reference image. Note that despite the robustness of the approach by HaCohen et al., not all of the image is matched. (b) The scattergram of the red channel for images 3 and 4 in Figure 1. The fitted  $\tau_r$  is shown in blue on top of the scattergram. The dotted lines show the threshold used to determine inliers (Equation 3).

*Estimation of the color transfer function.* Barring noise, the change of RGB values of all the correctly matched pixels between source and reference should follow the same transformation; however, because we work in the compressed domain, the mapping of the colors is not linear. We approximate the mapping between images taken with different exposure times via three color transfer functions (CTF)  $\tau_c (c \in \{r, g, b\})$ , one per channel. A simple parametric fitting, such as a line or a gamma function, would generally fail because cameras use custom transformations, including non-linear gamut mapping, as they attempt to create a visually pleasant picture instead of a faithful reproduction of the subject [24]. Ideally, converting the images to the *Lab* or *YCrCb* colorspaces and computing only the CTF  $\tau$  for the luminance channel should result in lower chroma noise, but we found that the accuracy is poor. Therefore, we work directly in RGB colorspace. This task has been tackled before: Tico and Pulli [25] adapt the comparagrams proposed by Mann [26] to increase their robustness in the context of fusion of blurry/noisy image pairs. HaCohen et al. [20] use cubic splines to fit the color transformation between images of different scenes. We adapt the fitting process to account for the noise added by wrong matches, which are in large part a consequence of the very nature of exposure stacks: because the images are acquired with the intent of sampling different parts of the range, they might have severely over- and under-exposed regions. Essentially, the fitting consists in finding the optimal solution for the following problem:

$$\begin{aligned} \tau_c^* &= \underset{\tau_c}{\operatorname{argmin}} \sum_p (\tau_c(R_c(p)) - S_c^w(p))^2 \\ &s.t. \quad \tau_c'(B) > 0, \forall B \in [0, 1], \quad c \in \{r, g, b\} \end{aligned} \quad (1)$$

where  $R_c$  and  $S_c^w$  denote the  $c$  color channel of reference and warped source images, respectively, and  $p$  is the pixel's index. To address the presence of outliers, which may bias the fitting, we use RANSAC. The monotonicity is a common

requirement [8, 27]. Instead of using cubic splines, we use cubic Hermite splines, which are globally more stable with respect to small perturbations of the samples and more importantly, preserve *monotonicity* of the samples being interpolated. To achieve a good fit, the selection of the control points that define the Hermite spline in Figure 3(b) requires some care. First, we divide the full range of values (x-axis in Figure 3(b)) into  $k$  adjacent, non-overlapping intervals and, at each RANSAC iteration, we take only one sample from each interval. While too few samples may not capture the behavior of the data, using too many samples may cause over-fitting of certain regions. In our experiment we found  $k = 7$  to provide a good compromise. Additionally, there might be parts of the range affected by too much noise (see Figure 3(b)). RANSAC derives its robustness from the fact enough samples support the correct model; in sparsely populated regions, however, this assumption fails and the quality of the fit is largely affected. To address such cases, at each iteration we accept a set of samples only if all the points in it are monotonically increasing. Figure 3(b) shows the scattergram and the fitted Hermite spline  $\tau_r$  for images 3 and 4 in the bottom row of Figure 1.

*Detection of mismatched pixels.* To detect false matches, we compare the RGB values of all the matched pixels in the warped source  $S^w$  with the value predicted by  $\tau_c$  from the corresponding pixel in the reference  $R$ . Assuming that pixel  $p$  has been matched, for each channel  $c = \{r, g, b\}$  we accept the match if

$$\frac{|S_c^w(p) - \tau_c(R_c(p))|}{\sqrt{1 + (\tau_c'(R_c(p)))^2}} < \delta, \quad (2)$$

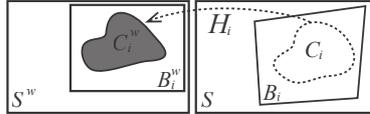
where  $1/\sqrt{1 + (\tau_c'(R_c(p)))^2}$  projects the vertical distance to the distance along a direction perpendicular to the fitted color transfer curve. We found that using a constant threshold  $\delta$  over the whole range is a poor choice: a larger error should be tolerated for larger color values. To reflect this, we define

$$\delta = \sigma_{med} \cdot S_c^w(p)^{0.25}, \quad (3)$$

where  $\sigma_{med}$  is median of the standard deviation of the data in the different parts of the scattergram. We consider all the points not satisfying Equation 2 to be false positives and add them to the set  $U$  of unmatched pixels. The set  $U$  can then be decomposed into a set of disjoint connected components  $\{C_i^w\}$ .

### 3.2 Correcting for missing correspondences

To fill the holes  $C_i^w$  in  $S^w$  caused by missing or rejected matches, we use  $\tau_c$  and synthesize the missing RGB values. This gives us a first approximation of the final result. Then, we attempt to recover the pixels' values from the source: if the region is small, it is often possible to approximate its transformation from the source with a local homography  $H_i$ . For a synthesized region  $C_i^w$ , we define  $B_i^w$  as the minimum bounding box extended by 4 pixels in width and height. Note that, by construction,  $B_i^w$  contains enough reliable pixels to estimate  $H_i$ , even in the unlikely case of a perfectly rectangular  $C_i^w$ . Using the notation indicated in



**Fig. 4.** The warped source contains holes  $C_i^w$  corresponding to unmatched or mismatched pixels, as detected by Equation 2. For such regions, we first attempt blending the gradients from the source: we take a box  $B_i^w$  around the hole and, if the motion from  $B_i$  to  $B_i^w$  appears to be rigid, we paste the gradients from the source using a local homography  $H_i$ . Otherwise we blend the gradients from the reference image instead.

Figure 4, with the set of pixels  $M_i^w = B_i^w \setminus C_i^w$  and their known correspondences, we can use RANSAC to estimate  $H_i$  from  $B_i$  to  $B_i^w$ .

The assumption that the motion can be approximated by means of a homography needs to be verified even for small patches, as they might contain small non-rigid bodies. To address this problem, we look at the set of inliers  $I_i$  determined at RANSAC’s best iteration and check if  $\frac{\#I_i}{\#M_i} < 0.7$ , where ‘#’ indicates the cardinality of a set. The rationale is that the outlier pixels have likely moved under a different transformation, suggesting a non-rigid motion within the patch. For increased robustness, we perform an additional test: using normalized cross-correlation, we check that this region matches the destination previously synthesized. For successful candidates, we paste the gradient of  $S(C_i)$  transformed with the homography  $H_i$  and blend it in using Poisson blending [28]. If, on the other hand, the result of these tests indicate a non-rigid motion, we blend the gradients from the reference, as they provide a better approximation than transforming the pixels directly using  $\tau_c$ .

### 3.3 Registering multiple images

The larger the difference in exposure time between two shots, the more difficult it is to establish reliable matches; for this reason, instead of directly aligning all the images to the reference, we align pictures that are adjacent in terms of exposure time. Consider a four-image stack ordered from brightest (image 1) to darkest (image 4), and where image 3 has been selected as the reference. We start by warping the image 2 to 3. We then register image 1 to the already registered and warped image 2 instead of image 3. The pixels of image 2 should be at the same locations as those of image 3, but their colors are more similar to those in image 1 than the colors of image 3 are, thus the matching process will produce a larger set of correspondences. The result of this stage is a new stack of LDR images that are perfectly aligned: Figure 5 shows the stack from Figure 1 aligned.

## 4 Results

In this section, we show the performance of the proposed algorithm on a number of challenging stacks. All our results use the Exposure Fusion method by Mertens



**Fig. 5.** The registered images from the stack in Figure 1. Note our algorithm correctly compensated for camera motion, and that all the people have been aligned to those in the reference image. The fused image is shown in Figure 1(b).

et al. [19]. We start by providing a comparison with state-of-the-art approaches for globally pre-registered images; while this is an arguably simpler problem, and there exist algorithms showing good results in such a case, it is an important benchmark and allows to illustrate some strengths of the proposed approach.

Gallo et al. [5] choose a reference image from the stack, divide it in patches, and combine patches from the other images only if they are consistent with the reference itself, without attempting to perform any registration. While their results are visually pleasing, we claim that such a strategy discards valuable information. Figure 6 shows a comparison with their method. It also compares how many of the original pictures have been used for each pixel of the final result. Note that for large portions of the image Gallo et al. use only one or two images, discarding blocks both due to motion and over- or under-saturation, while the proposed algorithm combines more pictures from the stack, only discarding too bright or dark pixels. This provides a higher contrast in some areas, such as the dead tree trunk on the left of the image (see electronic version of the paper).

Zhang and Cham [16] address the ghosting problem by weighting corresponding pixels in the stack based on the alignment of their gradients. They show that their method works well on a number of examples, but some artifacts are not completely removed, as is evident in the clouds in Figure 8. Rather than non-rigidly registering the images in the stack, they attenuate or discard pixels that would produce artifacts: this is the reason why the trees look washed-out. It should be said that, for stacks larger than 3-4 images, their method allows to remove objects appearing in one location in only one image, sometimes a desirable feature.

Kang et al. [17] and Zimmer et al. [18] are more similar to our approach in that they attempt to recover the non-rigid pixel transformations between shots. They both propose elegant solutions to the problem, however, they largely rely on the quality of the optical flow, whereas we carefully evaluate when the optical flow works and have a recovery strategy for when it fails. Figure 9 shows a comparison with Kang et al. While their results are visually pleasant, some artifacts remain that are caused by mistakes of the optical flow, as shown in the blow-ups in Figure 9(b). Similar considerations apply to the comparison with Zimmer et al. (Figure 10): when small objects or people are moving, the optical flow estimation may fail, in which case the final image is affected by ghosting.

A particularly difficult case for HDR imaging is that of non-rigid objects changing their appearance. For example, a typical situation with which amateur photographers often struggle is that of scenes containing water. Waves and ripples on the surface of the water are extremely difficult to register to begin with, and the different exposures exacerbate the problem. In Figure 11 we show one such scene and the result of using our algorithm. We also show the artifacts that the motion of the water causes by showing the result of fusing the three LDR images after aligning them with a rigid, global registration.



**Fig. 6.** A comparison with Gallo et al. [5], here shown in (a). The images on the right show the number of pictures from the stack used in the fusion (color-coded one through four). Our algorithm (b) uses more information; this provides a better contrast, particularly visible on the dark part of the trunk on the left, which is also shown in Figure 7 for an easier comparison.



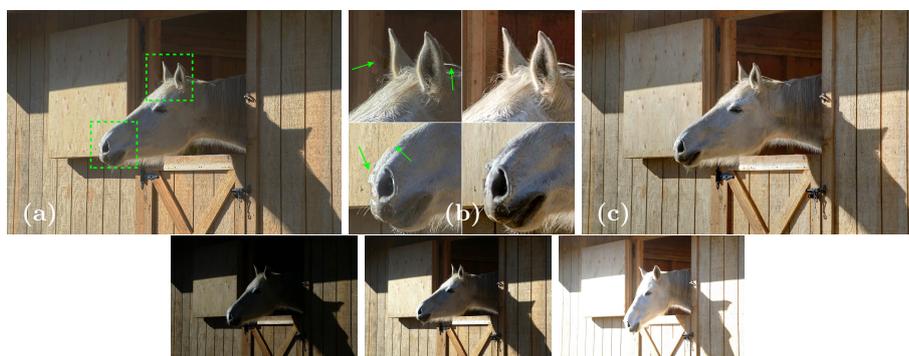
**Fig. 7.** A close-up comparison of the method by Gallo et al. (left) and ours (right) for the inset marked with the red box in Figure 6(a). Note how our method produces a better contrast in the dark area of the trunk, thanks to the fact that we use a larger set of images for each pixel: Gallo et al. discard a whole patch if enough of the pixels it contains are detected as outliers, see text.

## 5 Limitations and Future Work

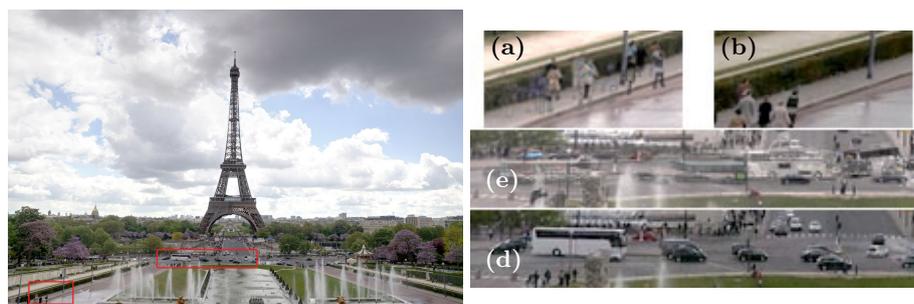
The results we presented show that our algorithm can deal with many of the situations in which standard methods may fail. Our method also tolerates a small amount of blur, as in the case of the mouth of the horse in the bright LDR image of Figure 9. However, sampling the whole irradiance range often requires fairly long exposures, and can therefore introduce a large amount of blur, thus potentially causing the proposed algorithm to fail. We plan to address the case



**Fig. 8.** A comparison between the method by Zhang and Cham (on the left) and our algorithm. Note the artifacts due to ghosting in the clouds. Image stack courtesy of Wei Zhang.



**Fig. 9.** A comparison with the method by Kang et al. [17] (a) and our algorithm (c). The regions within the green boxes are blown-up for comparison in (b), where the arrows point to the artifacts produced by the method by Kang et al. (left column). The bottom row shows the original images from the stack. Image stack courtesy of Sing Bing Kang.



**Fig. 10.** A comparison with the method by Zimmer et al. On the left the full result of our algorithm, with the red blocks indicating the locations of the insets on the right. (a) and (e) are their results while (b) and (d) are ours. Notice how our method solves the ghosting problems apparent in the results by Zimmer et al. (insets (a) and (e) are from the project website [29]). Image stack courtesy of Henning Zimmer.



**Fig. 11.** High-dynamic-range images of scenes containing water are notoriously difficult to capture due to the non-rigid motion of the water, as in the case of the LDR shots shown in the bottom row. A single, rigid homography only succeeds in registering the static parts of the scene (a), whereas our method (b) creates a picture that is as crisp as if the three LDR images used to generate it were taken at the same time.

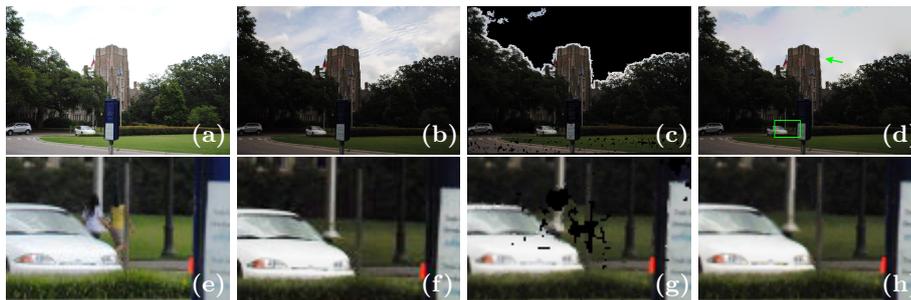
of stacks affected by strong blur with a combination of an increased robustness to noise for the dense correspondences step, and a more sophisticated completion scheme.

Figure 12 summarizes the potential failure cases of the proposed approach. Our method may fail when a large region of the reference image is completely saturated (or under-exposed), as in the sky of this example. In such cases, in addition to failing in capturing the whole irradiance range, Poisson blending may introduce small bleeding artifacts (Figure 12(d)). Finally, very small objects might cause problems as in the region indicated by the square in Figure 12(d), where the two pedestrians are completely included in the bounding box around the un-matched regions in Figure 12(g); for this reason, the patch from the source image happens to be consistent with the reference image and is therefore used, thus causing local inconsistency between reference and warped source. Our experiments indicate that such situations are extremely uncommon.

In our future work, we plan on addressing these issues by combining information from other images in the stack instead of limiting the search to either the source or the reference.

## 6 Conclusions

We have presented a method to process a stack of differently exposed pictures taken with a hand-held camera, and in the presence of motion in the scene. Most research in this area either focuses on rigid registration of the images or assumes registration and tries to address the fact that the scene is dynamic. We solve for both problems at the same time, and show that our algorithm produces registered



**Fig. 12.** (a): reference image, (b): source image, (c) dense correspondences, (d) final results. Below each image is the respective blowup from the region marked in (d). (Note that the contrast of (f) through (h) has been adjusted to make the problem look more obvious.) See Section 5.

images which, when fused together, are virtually artifact-free. We demonstrate that the robustness of our algorithm is superior to the previous work by means of comparisons. We have applied our method in the context of HDR imaging using exposure stacks, but the same approach can make other applications of burst photography, such as capturing focus stacks and denoising of low-light images by combining several exposures, more practical.

## References

1. Ward, G.: Fast, robust image registration for compositing high-dynamic range photographs from handheld exposures. *Journal of Graphics Tools* **8** (2003) 17–30
2. Tzimiropoulos, G., Argyriou, V., Zafeiriou, S., Stathaki, T.: Robust FFT-based scale-invariant image registration with image gradients. *IEEE Trans. on PAMI* **32** (2010) 1899–1906
3. Tomaszewska, A., Mantiuk, R.: Image registration for multi-exposure high dynamic range image acquisition. In: *Proc. of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*. (2007) 49–56
4. Jacobs, K., Loscos, C., Ward, G.: Automatic high-dynamic range image generation for dynamic scenes. *IEEE CG&A* **28** (2008) 84–93
5. Gallo, O., Gelfand, N., Chen, W., Tico, M., Pulli, K.: Artifact-free high dynamic range imaging. In: *Proc. of IEEE ICCP*. (2009) 1–7
6. Reinhard, E., Heidrich, W., Pattanaik, S., Debevec, P., Ward, G., Myszkowski, K.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann (2010)
7. Mann, S., Picard, R.: Being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. In: *Proc. of IS&T*. (1995) 422–428
8. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: *ACM ToG (Proceedings of SIGGRAPH)*. (1997) 369–378
9. Akyüz, A.O., Reinhard, E.: Noise reduction in high dynamic range imaging. *Journal of Visual Communication and Image Representation* **18** (2007) 366–376
10. Hasinoff, S.W., Durand, F., Freeman, W.T.: Noise-optimal capture for high dynamic range photography. In: *Proc. of IEEE CVPR*. (2010) 553–560

11. Granados, M., Ajdin, B., Wand, M., Theobalt, C., Seidel, H., Lensch, H.: Optimal HDR reconstruction with linear digital cameras. In: Proc. of IEEE CVPR. (2010) 215–222
12. Robertson, M.A., Borman, S., Stevenson, R.L.: Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging* **12** (2003) 219–228
13. Pece, F., Kautz, J.: HDR for dynamic scenes. In: Proc. of Conference for Visual Media Production. (2010)
14. Eden, A., Uyttendaele, M., Szeliski, R.: Seamless image stitching of scenes with large motions and exposure differences. *Proc. of IEEE CVPR* **2** (2006) 2498–2505
15. Khan, E., Akyüz, A., Reinhard, E.: Ghost removal in high dynamic range images. In: Proc. of IEEE ICIP. (2006) 2005–2008
16. Zhang, W., Cham, W.K.: Gradient-directed composition of multi-exposure images. In: Proc. of IEEE CVPR. (2010) 530–536
17. Kang, S.B., Uyttendaele, M., Winder, S., Szeliski, R.: High dynamic range video. In: ACM ToG (Proceedings of SIGGRAPH). (2003) 319–325
18. Zimmer, H., Bruhn, A., Weickert, J.: Freehand HDR imaging of moving scenes with simultaneous resolution enhancement. In: CGF (Proc. of Eurographics). Number 2 (2011) 405–414
19. Mertens, T., Kautz, J., Reeth, F.V.: Exposure fusion: A simple and practical alternative to high dynamic range photography. In: CGF. (2008) 161–171
20. HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D.: Non-rigid dense correspondence with applications for image enhancement. In: ACM ToG (Proceedings of SIGGRAPH). Volume 30. (2011) 70:1–70:9
21. Zwicker, M., Pfister, H., van Baar, J., Gross, M.: Surface splatting. In: ACM ToG (Proceedings of SIGGRAPH). (2001) 371–378
22. Liu, C., Yuen, J., Torralba, A.: SIFT flow: Dense correspondence across scenes and its applications. In: IEEE Trans. on PAMI. (2011) 978–994
23. Brox, T., Malik, J.: Large displacement optical flow: descriptor matching in variational motion estimation. In: IEEE Trans. on PAMI. (2011) 500–513
24. Kim, S.J., Lin, H.T., Lu, Z., Susstrunk, S., Lin, S., BrownHai, M.S.: A new in-camera imaging model for color computer vision and its application. In: IEEE Trans. on PAMI. (2012)
25. Tico, M., Pulli, K.: Image enhancement method via blur and noisy image fusion. In: Proc. of ICIP. (2009) 1521–1524
26. Mann, S.: Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. on Image Processing* **9** (2000) 1389–1406
27. Grossberg, M., Nayar, S.: Determining the Camera Response from Images: What is Knowable? *IEEE Trans. on PAMI* **25** (2003) 1455–1467
28. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. *ACM ToG* **22** (2003) 313–318
29. <http://www.mia.uni-saarland.de/Research/SR-HDR/limitations.shtml> (Accessed on March 3<sup>rd</sup>, 2012)