Seamless Image Stitching of Scenes with Large Motions and Exposure Differences

Ashley Eden
Computer Science Division
U.C. Berkeley
eden@cs.berkeley.edu

Matthew Uyttendaele
Interactive Visual Media Group
Microsoft Research
mattu@microsoft.com

Richard Szeliski
Interactive Visual Media Group
Microsoft Research
szeliski@microsoft.com

Abstract

This paper presents a technique to automatically stitch multiple images at varying orientations and exposures to create a composite panorama that preserves the angular extent and dynamic range of the inputs. The main contribution of our method is that it allows for large exposure differences, large scene motion or other misregistrations between frames and requires no extra camera hardware. To do this, we introduce a two-step graph cut approach. The purpose of the first step is to fix the positions of moving objects in the scene. In the second step, we fill in the entire available dynamic range. We introduce data costs that encourage consistency and higher signal-to-noise ratios, and seam costs that encourage smooth transitions. Our method is simple to implement and effective. We demonstrate the effectiveness of our approach on several input sets with varying exposures and camera orientations.

1. Introduction

In this paper, we address the problem of creating a high dynamic range (HDR) panorama from an image set captured by an amateur photographer. This requirement means that we must allow for our image set to: 1) have large misregistrations, *e.g*. due to scene motion, parallax, or calibration errors, 2) encompass a large range of exposures, and 3) have been taken with a standard hand-held camera. Large misregistrations are even more likely to occur given our allowance of panoramic image sets, *i.e*. sets that span a wide field of view. There are many good previous methods for creating HDR images and panoramas, but none of them deal with all aspects of our problem, nor can they be trivially extended to do so.

As discussed in section 3, saturated pixels are problematic when computing an HDR image. It is worth noting that if one ignores the first requirement of our problem, *i.e*. assumes that there are no large misregistrations in an image set, it is easy to deal with saturated pixels when creating an HDR image. In the case that there are no large misregistrations, we are guaranteed that each point in the scene will be captured at a wide range of exposures. (Capturing information at multiple exposures is a common requirement for all methods of HDR image construction.) Having no large misregistrations also guarantees that overlapping pixels always correspond to the same scene point. Thus, when there are no large misregistrations, we may simply avoid using saturated pixels in any sort of calculation. However, in the presence of large misregistrations, it may be that the configuration of the scene we wish to display is only captured in images with saturated values. Thus, for our problem, saturated pixels are much more of an issue than in previous HDR methods that relax our first requirement.

Our basic pipeline is to align the input images using a feature based registration algorithm that is invariant to exposure differences, and then convert these images to radiance maps. We then use a novel two step approach to select radiance values from these maps for the final composite. This approach is built on the graph-cut method used by [1]. Their method smoothly stitches together images in the presence of large misregistrations, but does not work with multiple exposure or panoramic inputs.

1.1. Previous Work

HDR Image Creation (Non-Hardware): There is much previous work in the creation of HDR images. Several methods work on input images taken at a wide range of exposures with a standard camera, *i.e*. with no additional camera hardware [18, 6, 19, 28, 13]. The inputs are used to determine the radiometric response function of the camera, and combined into a single HDR composite by calculating a weighted average of overlapping pixels. Mitsunaga and Nayar [19] determine the radiometric response function from areas of spatial flatness, and are thus insensitive to small misregistrations in the inputs.

None of these methods, however, account for large misregistrations, and they are thus ill-suited for our specific

problem. They also cannot be easily extended. Taking a weighted average of overlapping pixels is inaccurate when overlapping pixels do not correspond to the same scene point, which is the case with large misregistrations. Kang *et al*. [13] address the problem inherent to averaging by assigning weights to pixels based on the plausibility that they correspond to a reference configuration. This method is well suited for its application to video or rapidly acquired stills, but still has problems with large misregistrations (see Figure 6). In particular, there is no spatial smoothness to the pixel selection. It also does not address panoramic imagery.

HDR Image Creation (Hardware): There are also several good hardware based approaches to the creation of HDR images. Some cameras split incoming light to multiple image detectors, each with a different exposure [7, 24, 25, 14, 12]. This has the advantage that only one split image is necessary to create the HDR image, which can be created in real-time. The disadvantage is that it requires precision optics and calibration, and much additional hardware, and is thus very expensive. Nayar and Mitsunaga [20] introduce a far less expensive method using only one detector fitted with a mosaic of neutral density filters. Although all of these methods directly generate HDR images, they all also require internal hardware modifications to the camera. Because we wish to provide a solution for a standard camera, they are thus not appropriate for our problem. In addition, [7, 24, 25, 14, 12] do not deal with the problem of misregistrations if panoramic images are desired.

Several hardware approaches address the creation of HDR panoramas [26, 27, 2]. They rigidly attach simple filters to the camera. As the camera is rotated, the same part of the scene will be seen by different parts of the filter, effectively capturing each scene point under multiple exposures. Because the filter does not have to be placed directly on the sensor, these methods could be considered more flexible and cheaper hardware options than an internal modification. In addition, they are simple to apply, and [26, 27] can be applied generally to a wide range of scene information. None of these methods, however, are appropriate for our purposes. For one, they still require that the images be taken with an enhanced camera. They also employ averaging, and so cannot be used in the presence of large misregistrations.

Non-HDR Image Stitching: A wide variety of prior techniques are available to create seamless non-HDR panoramas in the presence of large misregistrations. In Davis [5] and Agarwala *et al*. [1], each pixel in the stitched image comes from one and only one of the inputs. The transitions between contribution from one input to another are made in areas of strong agreement for smoothness. These techniques, however, are intended to work on single exposure images. Uyttendaele *et al*. [29] allow for inputs of different exposures, first using a block-based exposure adjustment to make each input look more like its neighbors, and then performing a weighted vertex cover extraction to avoid seams. This method, however, is not intended for inputs with huge exposure differences.

Litvinov and Schechner [17] create seamless mosaics in the presence of camera nonuniformities (*e.g*. due to vignetting). By recovering the radiometric response function and camera nonuniformity at the same time, they have no need for later seam-feathering. However, they work on single exposure images of static scenes.

Our solution is inspired by [1]. Their system does not address the case of inputs with varying exposure. It would be a simple extension to the system to have it operate in radiance space, however this wouldn't avoid choosing source pixels with poor signal to noise or missing detail. In addition, the presence of such pixels means there will not be a common radiance mapping across different exposure images. Our solution addresses these issues.

1.2. Overview

Our system takes as input multiple images at varying orientations and exposures. We use a feature based registration technique similar to [4] to automatically align the inputs. This registration technique is tolerant to exposure differences and thus suitable for our purpose. Once the images are geometrically aligned, we map them to the same global radiance space. To do this, we use a pre-calibrated camera, the camera settings extracted from EXIF tags, and solve for an unknown per-color channel gain. Once we have the radiance values for each input image, we create a *final composite*, also in radiance space, where each pixel is set with the value from one of the input images. This is done in two steps.

We first use a subset of the radiometrically and geometrically aligned input images to create a *reference panorama* that covers the full angular extent of all the inputs, but not necessarily the full dynamic range. We create the reference using the graph cut technique in [1]. Each pixel in the reference is given the radiance value from one of the overlapping subset images. The reference will have optimally smooth transitions even when neighboring pixel values are chosen from different input images. As in [1], however, the creation of the reference does not avoid using pixels with a poor signal-to-noise ratio or missing detail. This is done in the second step, where the dynamic range of the reference is extended to that available in the full set of inputs. (Even if all of the inputs were considered in the creation of the reference, we are only guaranteed that all of the dynamic information is used after the second step.) We introduce cost functions that prefer choosing values with larger signal to noise ratios while keeping smooth transitions. These costs are minimized via a max flow graph cut. Once we have the final composite, we can optionally smooth any seams with an image blending step, and optionally apply tone-mapping

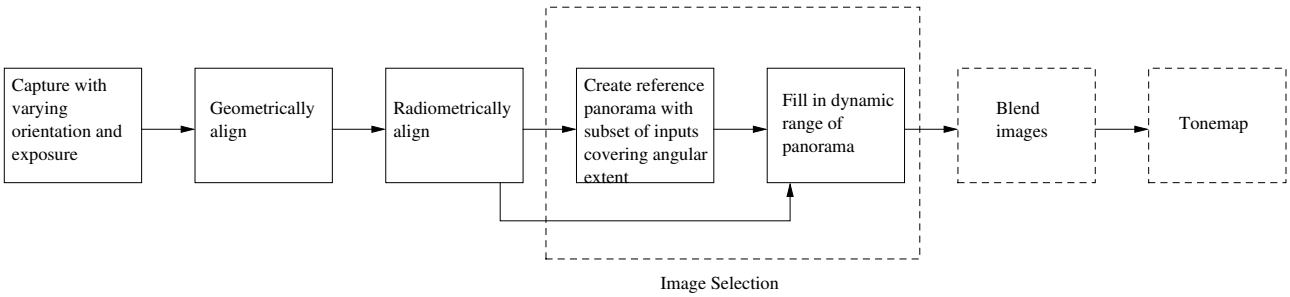


Figure 1. The processing stages to create an HDR panorama. The last two stages are optional, and are not used in this paper.

to the result to turn the high dynamic range of the composite into a displayable image. A block diagram of this process can be seen in Figure 1.

2. Computing Radiance Maps

To register the input images, we use a feature based image alignment technique that is invariant to affine changes in intensity [4]. See Figure 4a for an example of geometrically (as well as radiometrically) aligned images.

Once the images are geometrically aligned, we radiometrically align them by computing the radiance value at each pixel, $L_{p,i}$. Each pixel's measured intensity is mapped inversely through the camera model to a value proportional to the true scene radiance. Note that with perfect sub-pixel registration, $L_{p,i}$ and $L_{p,j}$ would be the same for overlapping pixel p in images i and j.

We use the camera model shown in Figure 2. Here, the camera settings—shutter speed, aperture, ISO, and white balance—are multiplicative factors of the radiance that may vary between images. The camera sensor is assumed to have a linear response and a nonlinear function, $f()$, is introduced at the camera output. The following equation shows the relationship between the radiance and measured intensity values:

$$L_{p,i} = \frac{g(M_{p,i})}{t_i d_i ISO_i WB_i} \tag{1}$$

where $g()$ is the inverse of the camera response function, M is measured intensity, and t_i, d_i, ISO_i, and WB_i are the respective shutter speed, aperture, ISO, and white balance of image i. We can extract the shutter speed, aperture, and ISO from the EXIF tags of most modern digital cameras, so we assume they are known.

Obtaining $g()$ is possible using a variety of methods, including:

a Pre-calibrate the camera using a published technique [18, 6, 19, 28, 21, 11].

b Use an International Color Consortium (ICC) profile for the camera.

c Solve for $g()$ directly on the inputs.

For this paper we use approach a, in particular a technique similar to [19]. Thus, the only unknown is the white balance.

To calculate white balance, we first get a rough estimate of the radiance with the following equation:

$$L'_{p,i} = \frac{g(M_{p,i})}{t_i d_i ISO_i} \tag{2}$$

where $L'_{p,i}$ is calculated per RGB color channel. We use RGB because this is the space in which cameras do their color correction. We choose one of the images as a reference image whose color balance we want all the other images to match. An example is shown in Figure 3 of a set of registered images with one labeled as the reference, R. We now want to compute the per color channel gain such that each non-reference image matches the color balance of the reference. To do this, we first consider a graph of how the images relate to each other, where each image is a node, and an edge connects overlapping images. The graph corresponding to Figure 3(left) is in Figure 3(right).

Along each edge in the graph, we construct a sum of squared difference error term between overlapping pixels, multiplied by the unknown gain. If the gain between image i and the reference is denoted by G_{iR}, the total error for our example is:

$$\begin{aligned} E \quad = \quad & \sum_{p \in \text{valid overlap}(1,R)} (L'_{p,1} G_{1R} - L'_{p,R})^2 + \\ & \sum_{p \in \text{valid overlap}(1,3)} (L'_{p,1} G_{1R} - L'_{p,3} G_{3R})^2 + \cdots \end{aligned} \tag{3}$$

i.e., we sum the difference between the scaled values in the overlap between images 1 and R, then 1 and 3, etc. Valid pixels p in overlapping images are chosen to avoid underexposed and saturated values. We solve for the unknown gains via least squares. Once we have the gains, we apply them inversely to L', thus normalizing out the color balance differences to obtain the final radiance value for each pixel in each image.

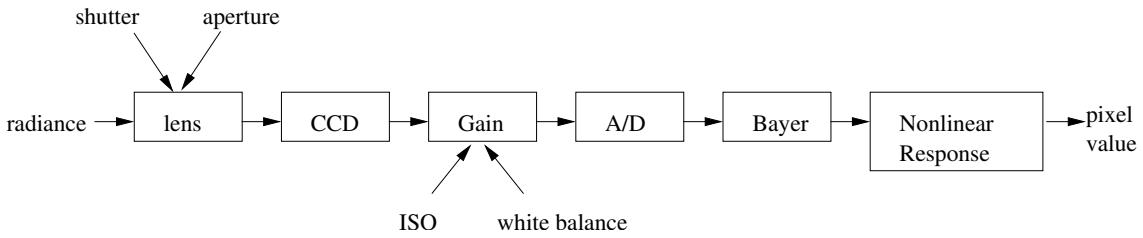


Figure 2. Block diagram of camera.

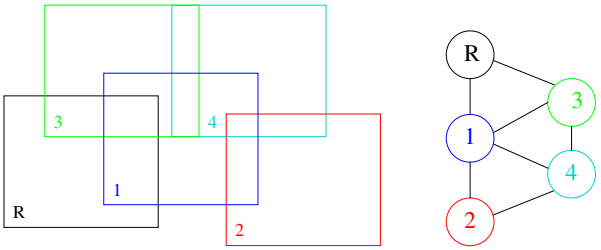


Figure 3. A set of overlapping images and their corresponding overlap graph.

3. Image Selection

In this section we present our image selection technique that allows for some scene motion between inputs while handling exposure differences. This is done in two passes. In our first pass, we create a reference panorama from a subset of the radiometrically and geometrically aligned inputs. The reference covers the field of view spanned by all the inputs, but not necessarily the full dynamic range.

Because of possible scene motion, not all of the input images necessarily represent the same scene at overlapping pixels. The reference panorama, and thus the subset used to create it, defines the desired position of moving objects in the scene. The inputs to this step can be chosen manually, if specific object positions are desired, or automatically. We create the reference panorama using a graph-cut technique similar to [1], except that we work in radiance space. At each pixel p that falls within the full angular extent, we want to assign a radiance value to the reference. By definition, at least one of the images in the subset overlaps at p. The radiance value assigned to p in the reference will be the radiance value at p from one of those overlapping images.

As in [1], during the generation of the reference, there is no penalty for choosing radiance values that come from underexposed or saturated measured intensities. It is the job of the second pass to fill in the complete available dynamic range from the full set of inputs, using the reference as a substrate. We add detail to regions if it is available and if it is consistent with the scene defined by the reference. To do this, we define a cost for using the radiance value from image i at pixel p in the final composite.

The radiance value at p in the final composite comes from one of the inputs, like in the first step. If the radiance value from image i is used, we may think of i as being the *label* of the composite at p. When determining the labeling Λ_p for any pixel p in the composite, we want a high probability P of referring to the same scene point as in the reference and a high signal to noise ratio, W. We use the following equation to compute the signal to noise ratio weighting for a pixel with a measured intensity M,

$$W(M) = \begin{cases} M & \text{if } M \text{ not saturated,} \\ W_{min} & \text{otherwise.} \end{cases} \quad (4)$$

Note that as the measured intensity increases, the signal to noise ratio increases. This makes sense because the signal is larger at brighter intensities. When M is saturated, however, we down-weight the pixel by assigning it a minimal weight.

The scene consistency for a pixel p in image i (with radiance $L_{p,i}$) is given by

$$P(L_{p,i}, L_{p,ref}) = \begin{cases} 1 & \text{if } M_{p,ref} \text{ is saturated} \\ & \text{and } L_{p,i} > L_{p,ref}, \\ e^{-\frac{(L_{p,i}-L_{p,ref})^2}{2\delta^2}} & \text{otherwise,} \end{cases} \quad (5)$$

where δ is the standard deviation of the noise in the imaging pipeline. In equation 5, there are two cases. If a pixel in the reference panorama is not saturated, we use a Gaussian noise model to determine the probability that an overlapping image's radiance value corresponds to the same scene. If a pixel in the reference panorama is saturated, all we have is a lower bound on its radiance value. In this case, if an overlapping image's radiance value is greater than the reference value, it is declared to be consistent and P is set to 1. If it is less than the reference radiance, it is treated in a similar manner to the reference not being saturated. Having a continuous function makes the optimization process more likely to find a coherent solution.

Using equations 4 and 5, we determine the data cost $DC(p, i)$ for selecting image i as the label of the composite at pixel p:

$$DC(p, i) = \frac{\lambda}{W(M_{p,i})P(L_{p,i}, L_{p,ref})} \quad (6)$$

where λ is a parameter used to regularize W and P. Since DC is a cost, higher values signify less desirable labelings,

so we make it inversely proportional to W and P. Thus, we favor higher intensity, consistent pixels.

The data cost is a penalty for selecting a given pixel based on information only at that pixel. Because a single input image does not generally cover the entire spatial or dynamic range of the output, we also need a criterion for a good transition from one image to another. As in [1], and in our creation of the reference panorama, we have a separate smoothness penalty. To determine the penalty for labeling pixel p in the composite with image i, i.e. setting $\Lambda_p = i$, we must calculate a seam cost C over each neighbor of p. For pixel p at location (x, y), consider one of its neighbors, p' at location $(x + 1, y)$. The current labeling in the final composite at p' is $\Lambda_{p'} = j$. The seam cost over p and p' is thus given by

$$C(p, p', i, j) = \frac{|L_{p,i} - L_{p,j}| + |L_{p',i} - L_{p',j}|}{L_{p,i} + L_{p,j}}. \quad (7)$$

Notice that since our inputs may have a very wide range of radiance values, for each neighbor we normalize the seam cost by the average of the two radiance values at the current pixel.

The goal of the second step of image selection is to find the labeling of the final composite that minimizes both the data and seam costs. Determining an optimal labeling using these criteria is done using an energy minimizing graph cut [3, 15].

3.1. Image Blending and Tone Mapping

Once the labeling of each pixel in the final composite has been computed, we can either directly copy the corresponding radiance value into the final high dynamic range image, or we can apply Poisson Blending [22, 16, 1] to merge the gradients from the original images. The latter approach can often help hide any inconsistencies in the radiance estimates between the different input images, but, in our experience, can also lead to artifacts if strong intensity edges from different images are not selected consistently.

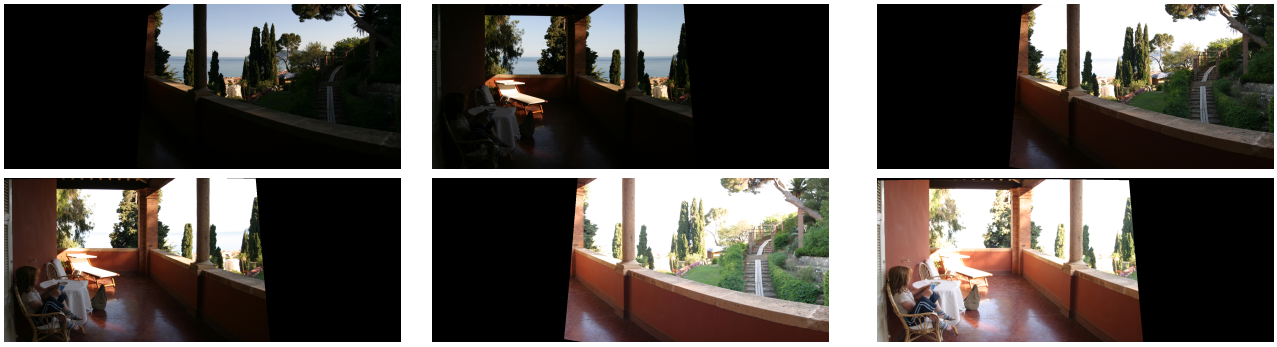
To display the final HDR image, a variety of tone mapping algorithms can be used [8, 9, 23]. In the pictures shown in the results section, we do not apply gradient domain blending but do use tone mapping. The radiance images can also be displayed after manually adjusting the contrast and brightness settings to show both the dark and bright regions.

4. Results

In this section we show results generated by our system on three different input sets. In figure 6 the input images were taken with a handheld camera, and have very different exposures and significant scene motion. In this case the reference image is simply the middle exposure image. Because of this, the moving child's position is defined by that image. In static parts of the scene, image detail is added to the reference in the second pass of image selection as appropriate: the darker portions of the indoors are taken from the longer exposure, while the brighter portions are taken from the shortest exposure. The labeling corresponding to this is shown in figure 6c. Our final result after tone-mapping is show in figure 6d. We also compare our result to that of [13] in figure 6e. Notice the artifacts in their result due to optical flow errors and the per-pixel image selection decision. In particular, the scene outside the window looks distorted, and around the ice cream cone (to the right of the image) looks mottled.

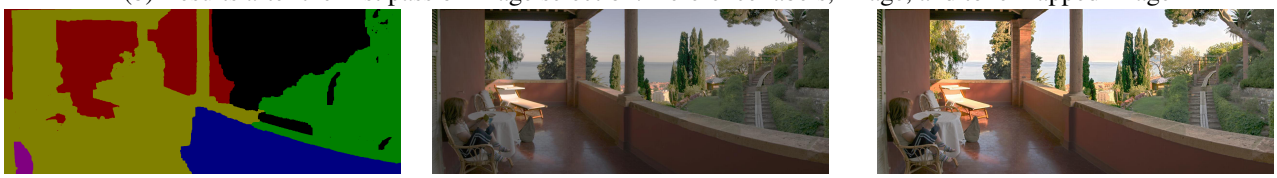
In figure 5 we show a typical panorama case of a rotating camera with variable gain and some moving people in the scene. The registered inputs are shown in figure 5a,d. The labeling and tone-mapped reference panorama, generated by the first pass of the image selection process, are shown figure 5b,c. Here, the reference panorama is essentially what would have been generated by [1] if they worked in radiance space instead of pixel space. As in [1] the generation of the reference does not avoid saturated sections of the inputs. The results of the second pass are shown in figure 5e,f. Notice that where possible, saturated pixels are replaced after comparison to the reference. In places where saturated pixels aren't replaced, because their radiance values are clipped, they show up as gray. Most of the saturated pixels after the second pass, however, are in a region that was only covered by one input image, which happened to have saturated values. Given this lack of further information, there is no way the second pass could have extended the dynamic range in these areas.

The final result we show is a hand-held rotating camera in auto-bracket mode. At each camera position, three exposures are captured. We show the images after geometric registration in figure 4a. Notice that there is a fair amount of parallax in the inputs, as can be seen by the pole at the center of the images relative to the trees. There is also a person moving slightly in the scene. The reference panorama, generated from the input images of shortest exposures, is shown in figure 4b. We show a tone-mapped result for the reference, but since the reference doesn't contain the full dynamic range, this is done mainly to illustrate the noise in the darker portions of the scene. In figure 4c the final result after the second pass is shown. Notice from the label and tone-mapped images there are many places where more detail has been added. The second pass greatly improves on the result of the noisy reference image, while containing minimal artifacts due to parallax and scene motion. The tone-mapper we used did introduce some banding artifacts, and as an alternative output we also show a result (center of figure 4b,c) of a global range compression curve applied to our radiance images.



(a) Registered bracketed input images

(b) Results after the first pass of image selection: Reference labels, image, and tone mapped image



(c) Results after the second pass of image selection: Final result labels, compressed HDR image, and tone mapped image

Figure 4. Results from a bracketed panoramic image sequence. The middle images of b and c are generated from the floating point radiance images but have a global compression curve applied to them instead of the local compression tone-mapping does.

5. Summary and Future Work

Because of the angular and dynamic limitations of standard cameras, the generation of a HDR panorama requires several inputs of varying exposures and orientations. In practice, these inputs are often of a non-static scene, captured with a handheld camera, leading to scene motion and parallax. These factors, in addition to imperfect camera calibration, make perfect sub-pixel registration of panoramic inputs nearly impossible. We have presented a novel non-hardware based method for seamless HDR image stitching assuming large scene motions and other misregistrations.

The main part of our algorithm is a two pass graph-cut based approach that first defines the position of the objects in the scene and then fills in the dynamic range when possible and consistent. We create data costs to encourage radiance values that are both consistent with object placement defined by the first pass and of a higher signal-to-noise ratio. Seam costs ensure that transitions occur in regions of consistent radiances.

Our results show that our method is able to deal with three different types of input sets: three image auto-bracket, auto-gain, and bracketed panorama. This encompasses most ways a photographer captures inputs.

In the future, it may be possible to obtain better results with blending and tonemapping algorithms better suited to our technique. Also, it would be nice to have an automatic way to choose which images define the placement of the objects in the creation of the reference panorama. A more accurate modeling of the camera response function would give more accurate corresponding radiance values, and thus a more accurate data cost. Where low detail pixels (saturated or underexposed) must be used, it may be possible to apply a synthesis technique to fill in the missing data. Finally, we could take into account vignetting [10], which is common in many images captured with low quality cameras such as camera phones.

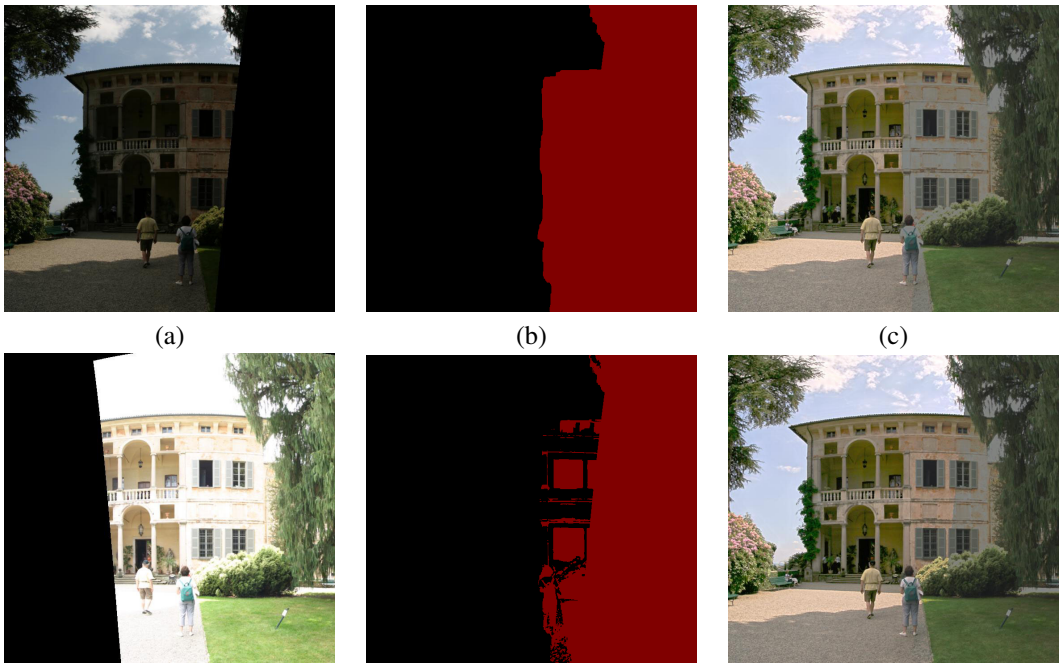
References

[1] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. *ACM Trans. Graph.*, 23(3):294–302, 2004.

[2] M. Aggarwal and N. Ahuja. High dynamic range panoramic imaging. In *Proceedings of the Eighth IEEE International Conference on Computer Vision*, volume 1, pages 2–9, 2001.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.

[4] M. Brown and D. G. Lowe. Recognising panoramas. In *Proceedings of the Ninth IEEE International Conference on*



(a) (b) (c)

(d) (e) (f)

Figure 5. Results from a panoramic image sequence with variable gain given only two inputs: (a & d): warped input images; (b–c): labels and tonemapped reference image; (e–f): labels and tonemapped final image. The grayish tones on the right-hand side of the composite image are caused by one of the color channels being saturated in the right-hand input. Notice that the final result (after the second pass) contains much less saturated areas than the reference panorama (after the first pass). Most of the saturated pixels after the second pass are in a region that was only covered by one input image, which happened to have saturated values. Given this lack of further information, there is no way the second pass could have extended the dynamic range.

Computer Vision, pages 1218–1225. IEEE Computer Society, 2003.

[5] J. Davis. Mosaics of scenes with moving objects. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 354–360, 1998.

[6] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *Computer Graphics*, 31(Annual Conference Series):369–378, 1997.

[7] H. Doi, Y. Hara, Y. Kenbo, and M. Shiba. Image sensor. *Japanese Patent 08-223491*, August 1986.

[8] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics (TOG)*, 21(3):257–266, 2002.

[9] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM Transactions on Graphics (TOG)*, 21(3):249–256, 2002.

[10] D. B. Goldman and J.-H. Chen. Vignette and exposure calibration and compensation. In *Tenth International Conference on Computer Vision (ICCV 2005)*, pages 899–906, Beijing, China, October 2005.

[11] M. Grossberg and S. Nayar. What is the Space of Camera Response Functions? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume II, pages 602–609, Jun 2003.

[12] E. Ikeda. Image data processing apparatus for processing combined image signals in order to extend dynamic range. *U.S. Patent 5801773*, September 1998.

[13] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High dynamic range video. *ACM Trans. Graph.*, 22(3):319–325, 2003.

[14] T. Kimura. Image pickup device. *Japanese Patent 10-069011*, March 1998.

[15] V. Kolmogorov and R. Zabih. What energy functions can be minimizedvia graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):147–159, 2004.

[16] A. Levin, A. Zomet, S. Peleg, and Y. Weiss. Seamless image stitching in the gradient domain. In *Eighth European Conference on Computer Vision (ECCV 2004)*, volume IV, pages 377–389, Prague, May 2004. Springer-Verlag.

[17] A. Litvinov and Y. Y. Schechner. Radiometric framework for image mosaicking. *Optical Society of America - A*, 22(5):839–848, 2005.

[18] S. Mann and R. Picard. Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proceedings of IS & T 46th annual conference*, pages 422–428, 1995.

[19] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 374–380, June 1999.

[20] S. K. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 472–479, 2000.

(a) Bracket input image triplet

(b) Reference image (after the first pass)

(c) Label image after the second pass

(d) Final result (after the second pass)

(e) HDR Video result

Figure 6. Results from a bracketed image triplet. The results are good despite the large amount of scene motion, and are better than those produced by the "HDR Video" technique, especially outside the window and near the ice cream cone.

[21] C. Pal, R. Szeliski, M. Uyttendaele, and N. Jojic. Probability models for high dynamic range imaging. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 173–180, 2004.

[22] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Transactions on Graphics*, 22(3):313–318, July 2003.

[23] E. Reinhard et al. Photographic tone reproduction for digital images. *ACM Transactions on Graphics (TOG)*, 21(3):267–276, 2002.

[24] K. Saito. Electronic image pickup device. *Japanese Patent 07-254965*, February 1995.

[25] K. Saito. Electronic image pickup device. *Japanese Patent 08-340486*, December 1996.

[26] Y. Schechner and S. Nayar. Generalized mosaicing. In *Proceedings of the Eighth IEEE International Conference on Computer Vision*, volume 1, pages 17–24. IEEE Computer Society, 2001.

[27] Y. Schechner and S. Nayar. Generalized mosaicing: High dynamic range in a wide field of view. *IJCV*, 53(3):245–267, July 2003.

[28] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of the CCD imaging process. In *IEEE International Conference on Computer Vision*, pages 480–487, July 2001.

[29] M. Uyttendaele, A. Eden, and R. Szeliski. Eliminating ghosting and exposure artifacts in image mosaics. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 509–516, December 2001.