Image Stitching with Efficient Brightness Fusion and Automatic Content Awareness

Yu Tang and Jungpil Shin

Graduate School of Computer Science and Engineering, The University of Aizu, Tsuruga, Ikki-machi, Aizuwakamatsu City, 965-8580, Fukushima, Japan

Keywords: Image Stitching, Brightness Fusion, Content Awareness.

Abstract: Image Stitching, also be called photo stitching, is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image. Image stitching is challenging in two fields. First, the sequenced photos taken from various angles will have different brightness. This will certainly lead to a un-nature stitched result with no harmony of brightness. Second, ghosting artifact due to the moving objects is also a common problem and the elimination of it is not an easy task. This paper presents several novel techniques that make the process of addressing the two difficulties significantly less labor-intensive while also efficient. For the brightness problem, each input image is blended by several images with different brightness. For the ghosting problem, we propose an intuitive technique according to a stitching line based on a novel energy map which is essentially a combination of gradient map which indicates the presence of structures and prominence map which determines the attractiveness of a region. The stitching line can easily skirt around the moving objects or salient parts based on the philosophy that human eyes mostly notice only the salient features of an image. We compare result of our method to those of 4 state-of-the-art image stitching methods and it turns out that our method outperforms the 4 methods in removing ghosting artifacts.

1 INTRODUCTION

Image stitching could be widely used in a lot of fields. One striking application of image stitching is that, it has been used extensively in panoramic photography. Image stitching method helps the panoramic photography to capture images with elongated fields of view. Photographers need to assemble multiple images of a view into a single wide image. Crude form of image stitching could be only the image matching and then registration (Baumberg, 2000; Brown, 1992; Brown, et al., 2004; Brown et al., 2005), the stitched result would be very unsatisfactory since the seam area will always be blurring. So image blending is crucial in solving this problem. The purposes of image blending as reflected in the pertinent literatures can be grouped as follows in our opinion: (1) de-blurring in the seam area. (Azzari and Bevilacqua, 2006; Jia and Tang, 2008; Brown and Lowe, 2007). Blurring is usually resulted from mismatching during the image registration process or by parallax. (2) de-ghosting. Ghosting artifacts is due to the moving objects or scene movements (Tang and Jiang, 2009; Uyttedaele and Eden, 2001; Yeh and Che, 2008; Yingen, 2009; Yao, 2008; Tang and Shin, 2010). (3) Eliminating visible seam. Visible seam is due to the various image brightness (Jia and Tang, 2008; Allene and Pons, 2008; Yeh and Che, 2008; Yao, 2008; Levin and Zomet, 2004). For the problem of brightness fusion, Burt (Burt et al., 1993) proposed to use image fusion to create a high quality image from bracketed exposures. But the measures could not be adjusted as flexible as our method. Tone mapping operators is applied spatially uniform remapping of intensity to compress the dynamic range (DiCarlo and Wandell, 2000; Drago et al., 2003; Fattal et al., 2002; Reinhard et al., 2002; Tumblin and Rushmeier, 1993). Their main merits are speed but difficult to produce a satisfactory image. A pyramid image decomposition method has been proposed by Li (Li et al., 2005) and attenuate the coefficients of the different exposures at each level to compress the dynamic range. Our method is also based on pyramid decomposition but works on the coefficients of the different brightness.

Tang Y. and Shin J..
Image Stitching with Efficient Brightness Fusion and Automatic Content Awareness.
DOI: 10.5220/0005087200600066
In *Proceedings of the 11th International Conference on Signal Processing and Multimedia Applications* (SIGMAP-2014), pages 60-66
ISBN: 978-989-758-046-8

Figure1: General architecture of the proposed method.

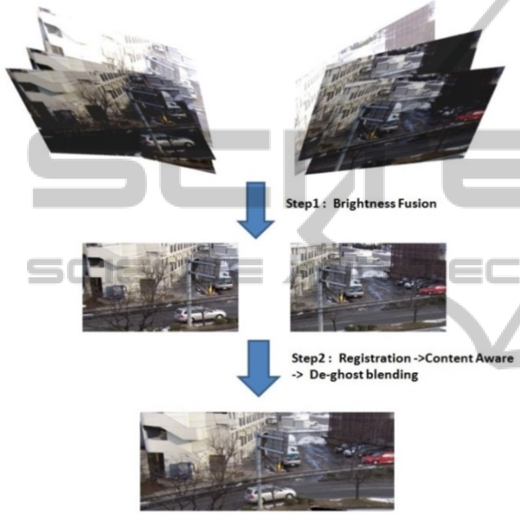

Figure 2: 1st row and 2nd row are a group of images taken under different exposure for left scene and right scene respectively. The last row is about the fused results for the sequenced images in the 1st row and 2nd row.

For the problem of de-ghosting, although existing image studies are various and have different focuses but not many specialized in eliminating the ghosting. A tonal registration method (Azzari and Bevilacqua, 2006) which is robust to moving, however, this works only when the presence of same moving objects in consequent frames which will not alter the overall histogram.

Uyttedaele (Uyttedaele and Eden, 2001) presents a weighted vertex cover algorithm to remove ghost effects, yet, this method is not failsave. Since the algorithm only prefers removing regions that are near the edge of the image because the vertex weights are computed by summing the feather weights in the ROD (Uyttedaele and Eden, 2001). Multi-blending (Allene and Pons, 2008; Brown and Lowe, 2007; Yao, 2008) is effective for de-ghosting but not in the case of presence of so many moving objects. Stitching line method conceived in (Tang and Jiang, 2009; Han and Lin, 2006) is good but cannot always find a satisfactory stitching line. Our paper still adopts stitching line method but the distinct feature is that our method can detect the prominent objects with automatic-awareness and thus can always search out an optimal stitching line to remove ghosting with least distortion.

2 BRIEF GENERAL SCHEME

The scheme presented here is quite a simple while efficient architecture for tackling the two tough problems: brightness variance and ghosting artifacts.

Step1: Brightness fusion. Considering that brightness variance always occurs when taking multiple photos, especially in outdoor because if the camera is back to Sunshine, image will be so dark. Otherwise, image will be so bright. We take multiple photos with different exposure (automatically set by camera with different exposure parameters) for each scene and then fuse the multiple photos to one image with appropriate brightness.

Step2: In this paper, we skip explaining the image matching part and registration part. SIFT feature based image matching (Baumberg, 2000; Brown, 1992; Brown et al, 2004; Brown et al., 2005) is already a very mature and efficient method for image registration. For blending and deghosting (In the left image, there is a car. While in right image, there is no car. We need to eliminate the ghosting artifact) in the seam area, we present an intuitive technique for finding an optimal stitching line which is automatically aware of the content and thus skirts around the salient objects. And then stitch the images according to this stitching line. The preciseness of the salient object awareness is satisfactory. We could see this in the experiment section.

3 BRIGHTNESS FUSION

There are multiple images taken with different exposure for each same scene. And we assume that the images are perfectly aligned. To achieve that, we need to be sure that the position of the camera should be fixed or there will be no any parallax movement of the camera. If it is unavoidable, we should possibly use a registration algorithm (Ward, 2003).

The fusion methods we propose will kick out bad parts and keep only the “best” parts in the multiple-brightness image sequence.

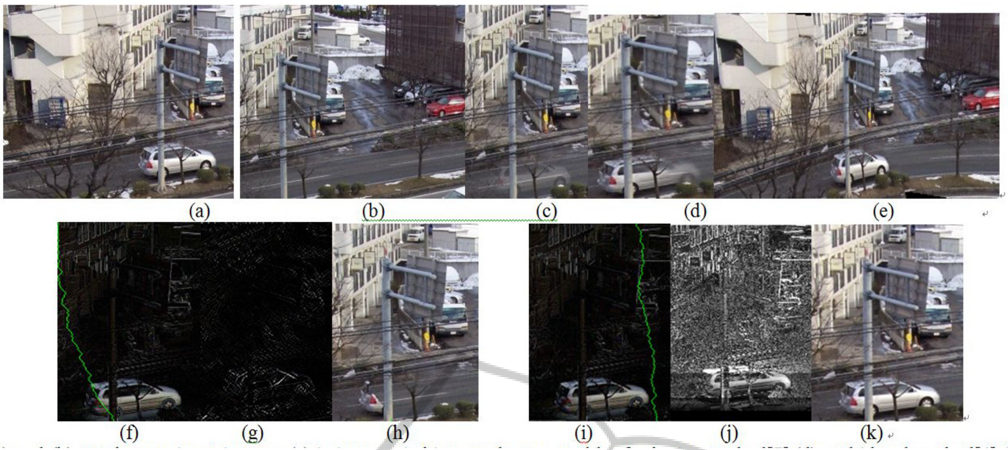

Figure 3: (a) and (b) are the two input images, (c) is image stitching result generated by feathering method (Uyttedaele and Eden, 2001). (d) multi-band method (Brown et al., 2007) € structure deformation method (Jia and Chi-Keung, 2008). The orange line indicates the shape of the traffic way is bent to curve. (f) to (h) is for our previous work (Yu and Huiyan, 2009)where subtracted image (f) shows the stitching line calculated by gradient map (g) and (h) is the stitching result. (i) to (k) is for the method of this work where subtracted image (i) shows the stitching line calculated by improved energy map (j) and (k) is the stitching result. Our result (k) shows the best.

We set three Brightness measures and make a weight map. Based on this weight map, we do the weighted blending of the multiple images.

3.1 Brightness Measures

To measure the images in the stack is well exposed or not, we apply the following measures to assign the weight. The regions in the image is under or overexposed, should receive less weight, while the area containing bright color should be preserved.

Exposedness (Measure E): Within the a channel we weight each intensity i based on how close it is to 0.5 using a Gauss curve:

$$\exp(-(i-0.5)^2/(2\sigma^2))$$

We apply the Gauss curve to each channel separately.

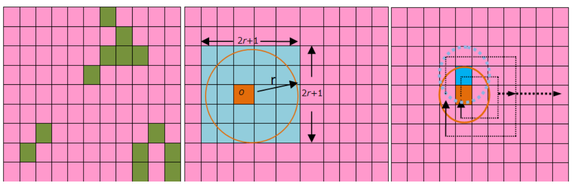

Figure 4: (a) Tentative detection based on global contrast. (b) The prominence detection based on local contrast of an image region with respect to its neighborhood is defined within a circle of radius *r*. (c) Filtering the images at one of the scales in rotate-expanding manner.

Contrast (Measure C): Each image will be applied a Laplacian filter and take the absolute value of the filter response. Because of Laplacian filter, it will assign a high weight to salient parts such as edges.

Saturation (Measure S): According to the photographing experience, the color will be desaturated if the photo is exposed for a long time. We get a saturation measure by computing the standard deviation within the R, G and B channel at each pixel.

We combine the measures in a linear form using a power function to get a weight map:

$$W_{ij,k} = (C_{ij,k})^{w_C}(S_{ij,k})^{w_S}(E_{ij,k})^{w_E} \quad (1)$$

$\omega_C, \omega_S, \omega_E$ is the weighting exponents of measure C, S, E. The subscript i, j, k refers to pixel i, j in the k-th image. If the exponent ω is zero, this measure will not be taken into account.

3.2 Laplacian Pyramid Fusion

Given the sequenced images $\hat{I}_{ij,k}$, we could get the fused image I by a weighted blending of the input images:

$$I_{i,j} = \sum_{k=1}^{N} W_{ij,k}\hat{I}_{ij,k} \quad (2)$$

with $\hat{I}_k$ the k-th input image in the sequence. However, applying formula 2 immediately will lead to disturbing seams. It is because the images we are combining contain different absolute intensities due to their exposure times. To address the seam problem, we use the technique inspired by Burt and Adelson, their technique blends two images seamlessly by an alpha mask and use a pyramid image decomposition. This multi-resolution blending is quite effective in avoiding seams. We also adapt this technique to our cases.

As seen from the Fig.2, we take three images of different exposure for the left scene and right scene in the first row and second row. We take the images with highly exposure and middle exposure and low exposure. We take the images in winter and it is a dark day, so even under middle exposure, the image is darker than usual. Using the fusion method described above, we get the fused results I1,I2 in the last row for left scene and right scene respectively. We could see that the fused results are quite satisfactory.

4 PRELIMINARIES FOR STITCHING AND BLENDING

Given two input images I_1, I_2 in Fig.3(a) and (b), after registration we get an aligned image. We define the part of Image I_1 in the overlap region as θ_1(the part of I_2 as θ_2 ,respectively). The subtraction between two images $\theta_1(x, y)$ and $\theta_2(x, y)$ in R, G, B channel, expressed as:

$$\theta^i(x,y) = \theta_1^i(x,y) - \theta_2^i(x,y), \quad i \in \{R,G,B\} \tag{3}$$

The subtracted image θ is shown in Fig.1(f) and (i) with no green stitching line. In order to remove the ghosting effects we target on searching a stitching line in θ which can intelligently go around the contour of the car. And stitch the two images according to this line.

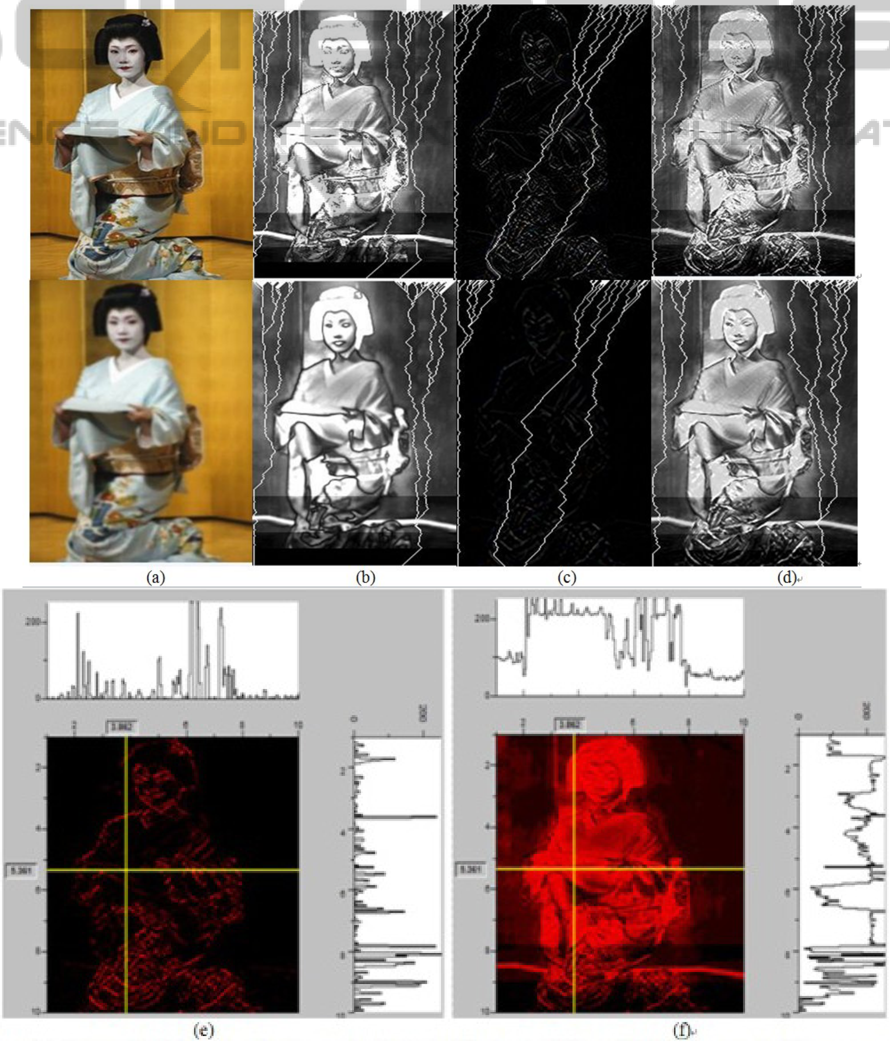

Figure 5: Column (a) shows original (above and noise version (below) of the image. Column (b)(c)(d) shows several less energy stitching lines calculated on prominence map, gradient map, improved energy map. (e) and (f) are profiles of gradient map and improved energy map respectively.

5 IMPROVED ENERGY MAP

Although (Tang and Jiang, 2009; Xiong, 2009; Levin and Zomet; 2004) shows that the gradient enjoys the advantage of respecting structures within the image thanks to assigning higher value on edges, the demerit lies in the gradient magnitude can be misled by trivial and repeated structures and salient objects are not well detected. In our improved energy map, we combine it with prominence map which considers the regions that are attractive but homogeneous as salient. Based on that, an optimal stitching line can always be found out for de-ghosting.

5.1 Prominence Map

1) Tentative Detection: To begin with, we roughly determine the prominence by evaluating the Euclidean distance of the average Lab vector value of an input image with each pixel vector value as:

$$E_{tent} = \|G_{\mu} - G(x, y)\| \quad (4)$$

where $G(x, y)$ is the input image, G_{μ} is the average of all Lab pixel vectors of the image, $E_{tent(x,y)}$ is the pixel prominence at position *(x,y)*. We use the Lab space instead of RGB space since RGB space does not take lightness of the color into consideration. Also Lab space also has advantage of approximating human vision and aspiring to perceptual uniformity. When $E_{tent}(x, y) > T$ (a threshold value), we impose a mark (shown as the green blocks in Fig.2(a) on that pixel which signifies that this pixel might be of importance since it keeps a large distance from the other pixels in the whole image. So in this stage, the prominent pixels are only tentatively identified.

2) Circular Scanning: Last stage, the tentative detection is based on the global contrast of the image. In this stage, we determine the prominent pixels on the basis of the local contrast of an image region with respect to its neighborhood. Thereby, we define a pixel as origin O(x, y) and designate its neighborhood as a circular region with radius r around it shown as Fig4.(b). The orange block is the origin O(x, y) while the blue blocks enclosed or passed by the orange circle are the neighborhood of origin O. The neighborhood is (2r+1)×(2r+1) block region in practical terms. The prominence of the origin O is evaluated as:

As a result, we observed how much educational effect about learning dictation and stroke order was emerged from this experiment and how much differences by learning process can be found.

$$E_{pro}^{i}(O) = \|N_{\mu}^{r} - v(O)\| \ , \quad i \in \{1,2,3\} \quad (5)$$

where N_{μ}^{r} represents the mean vector value of the neighboring pixels within the circular region with radius *r* in Lab space while *v(O)* denotes the vector value of origin. $E_{pro}^{i}(O)$ is still the L_2 -norm measured by Euclidean distance. Parameter *i* indicates that the prominence of the pixel is in which scale. We will calculate the prominence at three scales in total. N_{μ}^{r} is simply computed as :

$$N_{\mu}^{r} = \frac{1}{N}\sum_{n=1}^{N} v_n \quad (6)$$

where N is the total pixels of the neighborhood. The filtering is performed in a rotate-expanding manner as Fig.4(c) shown.

5.2 Comparison and Analysis

As shown in Fig. 5 column (a), given an original image (first row) and its Gaussian twice-blurred version(using 3×3 binomial kernel) in the second row. We explore the stitching lines within prominence map, gradient map, improved energy map and in the column (b)(c)(d), respectively.

It is obvious that prominence map and improved energy map protect the whole body of geisha against the passing of stitching lines very well. What is more, the two maps verify their robustness to the very noise. In contrast, the gradient map shows far worse results in the two versions since the stitching lines are either clustered to one side or just cross the body of geisha. Let us identify the reason for gradient map's failure of protecting the salient objects by comparing the profiles of the three maps in both local and global manner as Fig. 5(e) to (f). First, the colored energy maps of prominence map and improved energy map present that quite a bit of pixels with high visual significance (prominent pixels are marked in red) are distributed in the region of geisha evenly, while gradient map only assigns visual importance to the edges or trivial structures.

In other words, the amount of significant pixels is too few to protect the salient object. Assuming our stitching lines are horizontal or vertical straight lines as shown in the colored energy map. The top and right line charts showing the value of each pixel of the corresponding position in the horizontal and vertical stitching lines, respectively. Apparently the energy of stitching lines in the prominence map and improved energy map are much higher and more uniform than that in the gradient map. To fix the intractable problem of ghosting, safeguarding single salient object is far insufficient, our improved energy map is also very effective for protecting multi-objects.

6 EXPERIMENTAL RESULTS

We demonstrate our method capable of generating natural image stitching results for presence of either single moving object or multi-moving-objects. Comparison with other methods is also given.

In, Fig.3, (a) and (b) show two overlapped images of an automobile ambulation scene. Ghosting artifact is obvious in (c) using the Feathering algorithm (Uyttedaele and Eden, 2001). (d) use the multi-blend algorithm and presents a much better result than (c). Nevertheless, apparent blurring displays in the vincinity of car's head. At first sight, (e) is a good stitched mosaic, however, it is not a true reflection of input pictures since the original straight shape of the motor way is bent to a curve due to the deformation method in (Jia and Tang, 2008). (h) is the result of (Tang and Jiang, 2009) based on gradient map. Due to the stitching line's failure of avoiding the prominent object, a small fraction of car is remained. (k) shows the best result using our improved energy map(j) and stitching line (i).

7 CONCLUSION

In this work, we propose a simple while efficient framework to solve two difficult problems, brightness overexposed or underexposed and ghosting artifacts. To solve the problem of brightness, we take 3 pictures of the scene under different exposure, and pick good parts of each picture to be fused to the final result. The criteria for good or not is flexible and adjustable. It is mainly based on the measure of color, saturation, contrast. The number of input pictures is not limited to 3. More is welcome to enrich the fused details. We could see that our method works effectively in brightness fusion in experiment section. To eliminate ghosting artifacts, we present a novel energy map for finding an optimal stitching line which is automatically aware of the content and thus skirts around the salient objects. Since the energy map is essentially a combination of gradient map and prominence map which assigns higher importance to whole visually prominent regions (not only edges), the stitching line can easily skirt around the moving objects. The result section demonstrates that our method is better than the other four state-of-the-art (Azzari and Bevilacqua, 2006; Jia and Tang, 2008; Tang and Jiang, 2009; Brown and Lowe, 2007) techniques for de-ghosting.

REFERENCES

P. Azzari and A. Bevilacqua. Joint spatial and tonal mosaic alignment for motion. Proceeding on AVSS 2006, pp.89–102, Nov. 2006.

Jia jia, Chi-Keung Tang. Image stitching using structure deformation. *Pattern Analysis and Machine Intelligence IEEE Transactions.* 30(4):617-631, Apr. 2008.

Yu Tang, Huiyan Jiang. Highly efficient image stitching based on energy map. CISP'09 , pp.1-5, Oct. 2009.

Allene, C, Pons, J. Seamless image-based texture atlases using multi band blending. ICPR 2008, pp.1-4, Dec,2008

Han,B, Lin, X. A novel hybrid color registration algorithm for image stitching. Consumer Electronics, IEEE Transcations, 52(3):1129-1134, Aug. 2006

M.Brown, D.G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, Pp. 59-73, August. 2007

M.Uyttedaele, A.Eden. Eliminating ghosting and exposure artifacts in image mosaics. *CVPR2001*, pp.509-516, Dec.2001.

Tien-Der Yeh, Yon-Ping Che. An image stitching process using band-type optimal partition method. Asian Journal of Information Technology, 7(11):498-509, 2008.

Xiong, Yingen. Eliminating ghosting artifacts for panoramic images. ISM'09 pp: 14-16, Dec. 2009

Li Yao. Image mosaic on SIFT and deformation propagation. Knowledge Acquisition and Modeling Workshop, 2008, pp: 848-851, Dec. 2008.

Anat Levin, Assaf Zomet. Seamless image stitching in the gradient Domain. ECCV 2004, pp:377-389,May.2004

Baumberg, A. (2000). Reliable feature matching across widely separated views. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2000)*, pages 774–781, Hilton Head Island.

Brown, L. G. (1992). A survey of image registration techniques. Computing Surveys, 24(4), 325–376.

Brown, M., Szeliski, R., and Winder, S. (2004). Multi-Image Matching Using Multi-Scale Oriented Patches. Technical Report MSR-TR-2004-133, Microsoft Research.

Brown, M., Szeliski, R., and Winder, S. (2005). Multi-image matching using multi-scale oriented patches. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'2005),* pages 510–517, San Diego, CA

P. J. Burt, K.Hanna, and R.J. Kolczynski. Enhanced image capture through fusion. In Proceedings of the Workshop on Augmented Visual Display Research, pages 207-224. NASA-Ames Research Center.,Dec.1993.

J. DiCarlo and B. Wandell. Rendering high dynamic range images,. In Proceedings of SPIE, volume 3965, Jan. 2000.

F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. Computer Graphics Forum, 22:419–426, 2003.

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

E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21(3):267–276, July 2002.

J. Tumblin and H. E. Rushmeier. Tone reproduction for realistic images. IEEE Computer Graphics and Applications, 13(6):42–48, Nov. 1993.

Y. Li, L. Sharan, and E. H. Adelson. Compressing and companding high dynamic range images with subband architectures. ACM Transactions on Graphics, 24(3):836–844, Aug. 2005.

G. Ward. Fast, robust image registration for compositing high dynamic range photographcs from hand-held exposures. Journal of Graphics Tools: JGT, 8(2):17–30, 2003.

Yu Tang, Jungpil Shin. De-ghosting for Image Stitching with Automatic Content-Awareness. *Pattern Recognition (ICPR), 2010 20th International Conference on 23-26 Aug. 2010* 2210 – 2213.