A Computationally Efficient Algorithm for Multi-Focus Image Reconstruction

Helmy A. Eltoukhy* and Sam Kavusi

Department of Electrical Engineering
Stanford University, 350 Serra Mall, Stanford, CA 94305

ABSTRACT

A method for synthesizing enhanced depth of field digital still camera pictures using multiple differently focused images is presented. This technique exploits only spatial image gradients in the initial decision process. The spatial gradient as a focus measure has been shown to be experimentally valid and theoretically sound under weak assumptions with respect to unimodality and monotonicity.[1] Subsequent majority filtering corroborates decisions with those of neighboring pixels, while the use of soft decisions enables smooth transitions across region boundaries. Furthermore, these last two steps add algorithmic robustness for coping with both sensor noise and optics-related effects, such as misregistration or optical flow, and minor intensity fluctuations. The dependence of these optical effects on several optical parameters is analyzed and potential remedies that can allay their impact with regard to the technique's limitations are discussed. Several examples of image synthesis using the algorithm are presented. Finally, leveraging the increasing functionality and emerging processing capabilities of digital still cameras, the method is shown to entail modest hardware requirements and is implementable using a parallel or general purpose processor.

Keywords: digital cameras, optics, multi-focus, depth of field

1. INTRODUCTION

A pinhole camera exhibits the remarkable property that all portions of the imaged scene are brought into near perfect focus since it possesses infinite depth of field. However, due to the pinhole's obvious sensitivity deficiencies, lenses remain the optics of choice. When acquiring images of certain types of scenes, we would like to have large depth of field despite low illumination conditions. An example is a scene containing both close objects and a distant background. In this case, usually it is possible to capture the near and far objects in good focus using only two or three different focus settings. This suggests that one can acquire a series of pictures with different focus settings and fuse them to produce an image with extended depth of field. Accordingly, we present a computationally efficient algorithm for enhancing depth of field using multiple differently focused images.

Previous work by others investigated fusion methods based on wavelet and discrete cosine transformations[2–4] or use of a known camera point spread function (PSF).[5] Some of these methods suffer from the fact that they require prior knowledge of the system, e.g., must first determine the camera PSF (which can be time consuming and is generally shift-variant) or are complicated as they apply to several types of image fusion. In this paper we present an efficient algorithm for combining the in-focus regions of multiple differently focused images. The resultant synthesized image possesses a depth of field that is greater than any of the constituent images in the set, while retaining a natural verisimilitude.

Before presenting the details of the algorithm, we review a model of a typical lens system in Section 2 and derive equations relating focus measures to several optical parameters. In Section 3 we present the proposed algorithm along with analysis of its sensitivity to noise and other non-idealities. Simulations and experimental results are presented in Section 4. Finally, in Section 5 we discuss the complexity of the algorithm and show that its modest computational requirements are well suited to general image processors.

E-mail: eltoukhy@stanford.edu, Phone: (650) 725-9696, Fax: (650) 724-3648

2. LENS SYSTEM

Figure 1 shows a paraxial geometric optics model of image formation. Although geometric optics is only useful for analysis of first order effects, diffraction-related effects as predicted by physical optics are not significant for this application, since the spatial resolution of most sensor arrays is quite below that of the diffraction limit. For example, assuming the Rayleigh criterion for the far-field Fraunhofer diffraction-limited case, $resolution = \frac{1.22\lambda f}{D} \approx 1.5\mu m$ for typical values. Indeed, most CCDs and CMOS imagers have pixel sizes greater than 3 μm and well above 5 μm. Continuing with the thin-lens geometrical model of Figure 1, the point P on the object plane at U in the scene is imaged and perfectly focused as point p' on the image plane at V. The well-known lens equation, $\frac{1}{f} = \frac{1}{U} + \frac{1}{V}$, relates the position of these two points, U and V, with that of the focal length, f, of the lens. Furthermore, each lens must have a finite aperture (assumed to be circular in this case), the diameter of which is denoted D. Finally, the actual distance of the sensor plane to the lens, S, will allow us to estimate R, the radius of the blur circle induced by a non-zero difference between S and V. Using similar triangles, we can solve for R in terms of S, V, and D obtaining

$$\frac{D}{V} = \frac{2R}{S-V} \tag{1}$$

$$R = (S-V)\frac{D}{2V}. \tag{2}$$

We can eliminate V and write R in terms of camera parameters using the lens equation yielding

$$R = \frac{1}{2}DS\left(\frac{1}{f} - \frac{1}{U} - \frac{1}{S}\right). \tag{3}$$

Note that the significance of the negativity of R is only that the photodetector plane, S, is closer to the lens than the in-focus image plane, V. As is readily apparent, the larger the aperture the larger the blur circle engendered. In fact, using the above relation for the blur circle, we can derive the depth of field (DOF) for a lens system, where R now becomes the largest acceptable blur circle in the resultant image, which can be chosen based upon sensor resolution and human visual acuity limits. Combining the sensor plane displacement, $S - V = \frac{2RV}{D}$, with the lens equation yields

$$U_{\text{far}} = \frac{Uf(1 - 2\frac{R}{D})}{f - 2\frac{R}{D}U} \tag{4}$$

$$U_{\text{near}} = \frac{Uf(1 + 2\frac{R}{D})}{f + 2\frac{R}{D}U} \tag{5}$$

$$DOF = U_{\text{far}} - U_{\text{near}}, \tag{6}$$

where $U_{\text{near}}, U_{\text{far}}$ are the distances to the nearest and farthest object planes with blur circles less than or equal to the chosen R and U is the distance to the in-focus object plane as before. As $D \to \infty$, $U_{far} = U_{near} = U$ and DOF = 0. This result, of course, agrees with common knowledge that reducing aperture size increases the camera system depth of field. However, limiting aperture size is often not a luxury one can afford, since when imaging in indoor or overcast environments maximizing light collection becomes necessary in order to achieve acceptable SNR levels.

Now that the size of the blur circle has been calculated, it becomes necessary to evaluate the effect of its finite size on the resultant image. Although a real lens suffers from diffraction and aberration related limitations, which filter out higher spatial frequencies, we will continue to assume the first order geometrical model, which implies that the blur circle will be of uniform intensity and of radius R. The Fourier transform of such an intensity distribution with $f_r = \sqrt{f_x^2 + f_y^2}$ is

$$\mathcal{F}\{I(r)\} = 2\pi R^2 \left[\frac{J_1(Rf_r)}{Rf_r}\right], \tag{7}$$

where J_1 is a Bessel function of the first kind, order one.[6]   As R increases, i.e., the image is defocused, the width of the central maximum of the above transform decreases thereby implying increasingly attenuated higher frequencies. Unfortunately, the presence of the higher frequency side-lobes in the Bessel function-based OTF can complicate discrimination of focus quality, since some, albeit attenuated, high frequency components are passed through in even out-of-focus images. It should be noted, however, that in the limit of non-negligible further low pass filtering by diffraction and aberrations (chromatic or otherwise) of real lens systems and the non-ideal modulation transfer function (MTF) of the sensor, the above implied OTF transmutes to an increasingly more benign Gaussian distribution of the form, $e^{-\frac{1}{2}f_r^2 R^2}$.[7]

In the subsections that follow, various non-idealities, such as, brightness changes and magnification, as a result of the displacement of the sensor plane at different focus settings are analyzed.
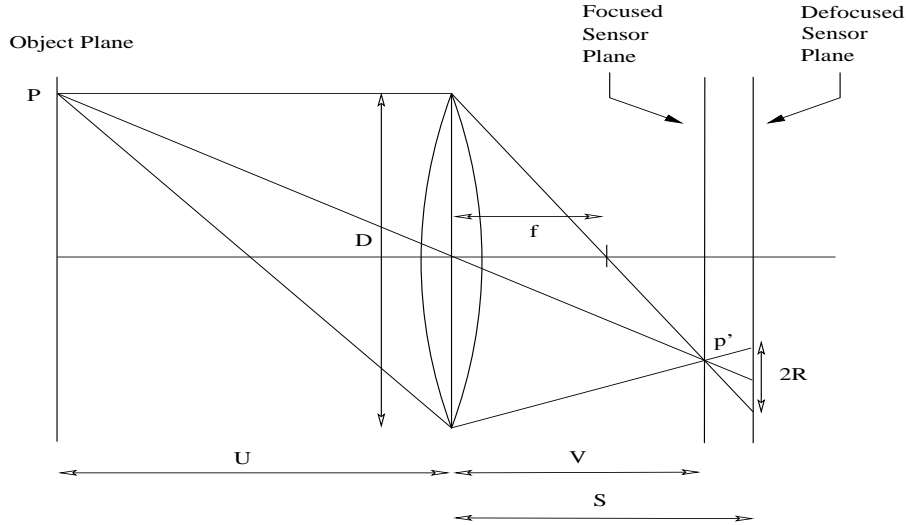

Figure 1. Geometric Optics Model of Lens System

2.1. Brightness Variations

The brightness of an image is dependent on the amount of light gathered by the lens system. The light gathering ability of a lens system is typically indicated by its numerical aperture (NA), which is defined as $n \sin \theta_{\max}$, where $\theta_{\max}$ is the angle between the central axis and the ray with the largest possible entrance angle. Since we are dealing with camera systems, n will be taken equal to one, which is the index of refraction of air. Once again using Figure 1, we obtain

$$NA = \frac{D}{\sqrt{D^2 + 4S^2}} \approx \frac{D}{2S}, \ 4S^2 \gg D^2. \tag{8}$$

Since objects at infinity focus at f, and this is when $\theta_{\max}$ is largest, NA is commonly denoted by $\frac{D}{2f}$. Interestingly enough, the brightness of the image is directly proportional to the square of the lens system NA.[8] This is, in turn, related to the speed or f-number of a camera system, which is commonly denoted by f/# and is equal to $\frac{f}{D} \approx \frac{1}{2NA}$. This brightness variation with S, can potentially cause a problem for any fusion algorithm, since uniform patches taken with different focus settings will have disparate brightness levels. It has been suggested that one can simply normalize the images with respect to the change in NA, or to the global mean of each image when NA is unknown. However, this works well only when saturated pixels represent a negligible portion of the scene, otherwise severe distortion will occur because of the non-linear clipping effect of saturation. Hence, when significant saturation does occur, for example when imaging bright scenes with the sky or white patches in their midst, a region-based rather than global mean normalization method is much better suited.

2.2. Magnification and Optical Flow

Another optical phenomenon which presents a possible obstacle to fusion of differently focused images is image misalignment due to magnification or misregistration. The former occurs when the sensor plane (or more accurately, the lens), is moved between frames, thereby changing the effective magnification of the imaged object. Once again using Figure 1, the change in magnification can be approximated by $\frac{S_2}{S_1}$, where S_1 and S_2 are the respective sensor to lens distances of each frame. This implies that we can normalize the images using the corresponding S_i's for each image to rescale them appropriately.[1] Even with this rescaling, there will be slight alignment problems due to blur and the change in lens position. Ostensibly, the proposed algorithm must be tolerant of such displacement (which can be on the order of several pixels). On a related note, misregistration can also occur as a result of slight camera movement between frames. An affine or perspective warping of the image using estimation of the global motion vectors is one possible remedy if misregistration is significant.

3. PROPOSED ALGORITHM

The basic crux of the problem is deciding which portions of each image are in better focus than their respective counterparts in the associated frames and combining these regions to form the synthesized extended depth of field image. In short, due to the low pass filtering nature of the modified Bessel function present in defocused images, the discrimination method of choice invariably involves quantification of high frequency content. Use of transformation methods, such as either discrete cosine or wavelet among others, has been reported extensively in the literature. Yet these techniques offer, in a sense, too much information for the required task, whereby frequency content (or other information) across the entire range is gleaned. Instead, "bulk" measures of high frequency content can be used. Various such measures, including image variance, image gradients, and image Laplacians have been employed and validated for related applications such as autofocusing.[1, 7, 9] Furthermore, modifications of these measures have been reported including the famous Tenengrad, which adds a thresholding operation to the accumulated image gradients in order to increase the sharpness of the measure. Although the particular sharpness of these measures is crucial for autofocusing applications, where precise discrimination between focus settings in a fairly large set must be made, our application is far more forgiving. Indeed, most scenes can be adequately imaged using at most two or three settings, which allows the use of coarser indicators for measuring focus quality. Out of the three listed focus measures above, it is clear that local image gradients offer the simplest means of focus discrimination in terms of implementational and computational complexity.

3.1. Image Gradient

To further ease implementation we confine all operations to separable FIR filters. In the x direction, $\frac{dI_i}{dx}$ can be implemented using a two-tap first-order differencing operation, such as $I_i(x, y) - I_i(x - 1, y)$, where $I_i(x, y)$ is the pixel intensity value at position (x, y) in image i. Similarly, $\frac{dI_i}{dy} = I_i(x, y) - I_i(x, y - 1)$. This achieves the required high-pass filtering operation in each spatial direction. A commonly employed focus measure is the energy of image gradients, which is defined as $\sum_x \sum_y (\frac{dI_i}{dx})^2 + (\frac{dI_i}{dy})^2$. This measure has been shown to be both experimentally and theoretically valid in discrimination of focus quality if a Gaussian or truncated Bessel blur function is assumed.[1, 7] Furthermore, under these conditions, it is monotonic and unimodal, properties that are important in ensuring reliable discrimination, since only one global maximum should exist if the optimally focused image is to be chosen with ease. In cases in which the defocus OTF exhibits significant side lobes, a Gaussian low pass filter can be applied to the image set a priori to validate use of the measure.[1] In order to somewhat simplify this measure, in terms of hardware complexity, we replace the squaring operation with that of absolute value. It is difficult to show theoretically that unimodality and monotonicity are maintained due to the non-linearity of the operator; however similar measures, such as Nayar's sum-modified Laplacian, SMLF,[10] in which the gradient is replaced with the Laplacian and the absolute values rather than squared magnitudes are calculated, have been used successfully in practice. As such, the first step of the proposed algorithm involves calculating the following gradient-based focus measure at each point for each image in the set:

$$G_i(x, y) = \left| \frac{dI_i(x, y)}{dx} \right| + \left| \frac{dI_i(x, y)}{dy} \right|. \tag{9}$$

Clearly, the larger the relative magnitude of the indicator, the higher the probability that its corresponding blur circle radius is smaller, since its respective image suffers less low-pass filtering. Hence, the pixel-level metric between a pair of images becomes, $M(x,y) = G_i(x,y) - G_j(x,y)$, where i, j correspond to two differently focused images. Thus, $M(x,y) > 0$ indicates the pixel value at location (x,y) in image i should be chosen otherwise we choose its counterpart from j. However, this measure alone in practice is not sufficient to pick out the better focused image on a pixel-by-pixel basis, since the above analysis of focus measure soundness assumed summation of these metrics over the entire image. Thus, aggregation of these measures is necessary.

3.2. Majority Filtering

Use of solely near pixel-level indicators, such as the image gradient, alone can make decisions vulnerable to wide fluctuations dependent on sensor (i.e. noise), optics (magnification and side lobes) and scene (local contrast) specific parameters. Furthermore, most methods employed by autofocusing techniques are global or semi-global in scale. Hence, corroboration from neighboring pixels of decision choices becomes necessary to maintain robustness of the algorithm in the face of the above adverse effects. Adding this corroboration while maintaining pixel-level decisions requires summing the $M(x,y)$'s over an $k \times k$ region surrounding each decision-point. This yields a new focus measure

$$\tilde{M}(x,y) = \sum_{i=-k/2}^{k/2} \sum_{j=-k/2}^{k/2} M(x+i, y+j). \tag{10}$$

The use of such aggregation consequently increases the accuracy of the decision by ensuring that pixels with large focus measures influence the decision of their neighbors. Implementation of such a summation can be accomplished through the convolution of the decision map consisting of the calculated M values with both a ones vector of length n and its transpose. Finally, since the decisions are now effectively blurred as a result of this aggregation, or in effect low pass filtering, a sigmoid function (or signum as an approximation) is applied to these filtered focus measures in order to transform them into a semi-hard decision map. This last step transforms the linear accumulation operation into a non-linear majority filter, while at the same time blending low contrast portions of the scene thereby providing partial immunity to brightness variations. Thus, the resultant measure becomes

$$\hat{M}(x,y) = \frac{1}{1 + e^{-\beta \tilde{M}(x,y)}}, \tag{11}$$

where β is a constant and each pixel value of the synthesized image is

$$I_s(x,y) = \hat{M}(x,y) I_i(x,y) + (1 - \hat{M}(x,y)) I_j(x,y). \tag{12}$$

3.3. Hard vs. Soft Decision Regions

A typical problem that can occur with any type of image fusion is the appearance of unnatural borders between the decision regions resulting in a stitched together or "blue-screen" appearance. Furthermore, even if the discontinuity between image regions is negligible, the overlapping blur at focus boundaries makes all fusion algorithms (except for those that correct for camera-specific PSDs) susceptible to uncertainties in assigning the precise region boundary to the segmentation map. Indeed, objects in the foreground occlude background regions up to the extent of their blur circle radius. Unfortunately, neither image in the pair contains a focused representation of this occluded region. To combat this, soft decision boundaries can be employed using smoothing or low pass filtering of the decision map, $\hat{M}$. This creates weighted decision regions where a linear combination of the pixels in the two frames are used to generate corresponding pixels in the fused image. Accordingly, the above relation for I_s remains unchanged, however $\hat{M}$ is now a smoothed version of its former self. This technique does not globally reduce sharpness of the resultant combined image, since it has no effect on pixels located away from the decision region periphery. Moreover, adverse effects such as, optical flow due to minor optical magnification or camera movement, brightness variations, background occlusion, and algorithmic error can be tolerated to a larger extent without a noticeable degradation in the quality of the synthesized image.

3.4. Extension to Several Images

It is clear from the above that I_s is a linear combination of the pixel values in the two original images, i and j, where $\hat{M}$ represents the decision map indicating the weight applied to each. In order to extend the algorithm to more than two images, we propose an iterative solution whereby we first perform the method on the first two images in the set and subsequently between the synthesized image, I_s, and each remaining image. In other words, each pixel now becomes

$$I_s^{i+1}(x,y) = \hat{M}(x,y)I_s^i(x,y) + (1 - \hat{M}(x,y))I_{i+2}(x,y). \tag{13}$$

We assume in the following that only two images are required to synthesize the extended depth of field image.

4. SIMULATIONS AND EXPERIMENTAL RESULTS

Assessing the performance of such fusion algorithms using real images is difficult due to the uncertainty and subjectivity inherent in manually generating a "perfectly" fused image. Consequently, wishing to sidestep such issues, we use synthetically blurred pairs of images from an in-focus image in order to obtain a somewhat meaningful estimate of the performance of the algorithm amidst variation of significant parameters. Finally, an example using a real image pair is presented.

4.1. Simulations

In order to observe the dependence of the proposed algorithm on image noise, magnification, etc., we performed simulations using synthetically generated defocused regions using a Gaussian blurring function on sets of images. We assume that the original image is perfectly in focus everywhere and therefore use it as our comparison benchmark. As a first example, we tested the performance of the algorithm in the presence of very noisy images as shown in Figure 2. The SNR of these images is less than $20dB$. Figure 3 shows the fused image. Notice that the tower in the background and building in the foreground are both in focus. This bodes well for the performance of the algorithm in low light situations.
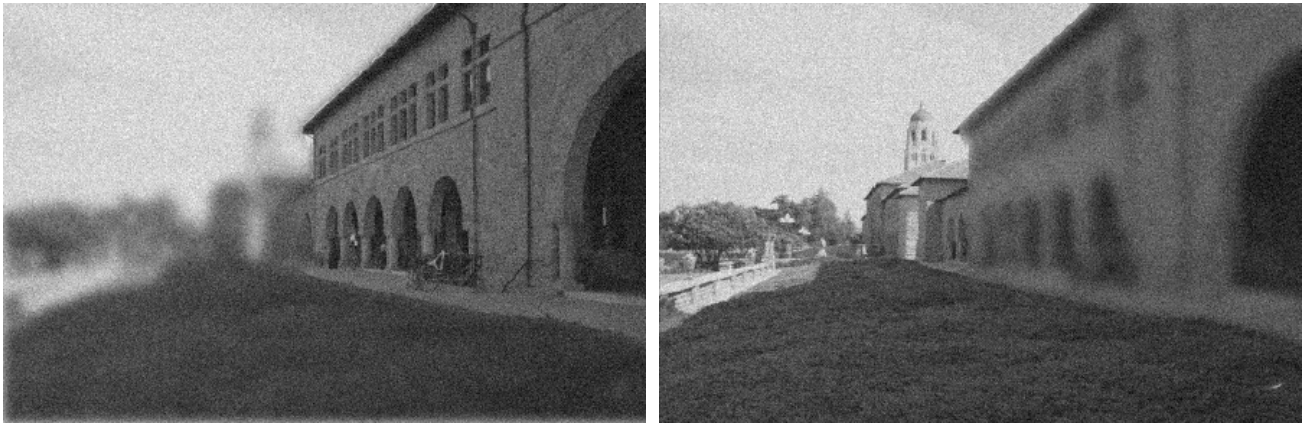
**Figure 2**. Two synthetic, noisy partially focused images.

In order to quantify the error in the reconstructed image we calculate the S-CIELAB [11] ΔE image between the original image (all in focus) and the reconstructed one. ΔE is measured using a spatial extension of the CIELAB color metric, S-CIELAB. To simplify comparison using the S-CIELAB metric, the average spatial ΔE value is calculated to evaluate the quality of each synthetic image.

Figure 4 plots the average spatial ΔE versus added noise. As can be seen, the algorithm is fairly resistant to added random noise because of the aggregation operation. Note that for generating Figure 4, the decision map, which is generated using the noisy images, constructs a synthetic image from the noise-free images. This synthetic image is then compared with the original globally focused image in order to observe the spatial ΔE due only to the algorithm and not the added noise. As a comparison, Figure 5 illustrates the effect of aggregation
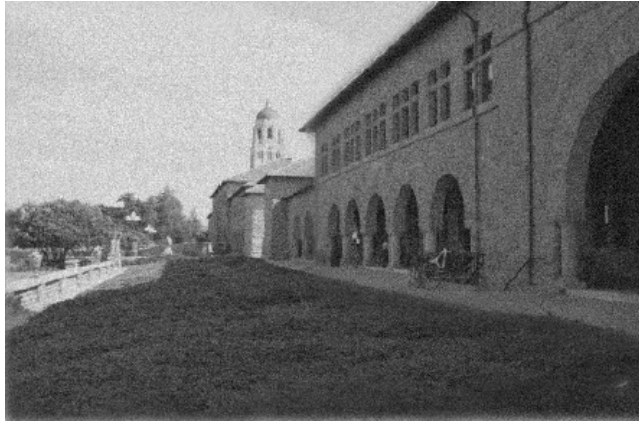
Figure 3. Fused image generated by the proposed algorithm.

size (k) on performance despite added noise for k=20,80,110. Clearly, the larger the value of k, the more resistant the algorithm is to noise, although at a cost of less specificity. Indeed, Figure 6 depicts the normalized (to the perimeter of the decision region) average spatial ΔE versus the width of one of the blur regions. For regions narrower than about 50 pixels (for k=60), the error rises sharply. Finally, Figure 7 shows the effect of misregistration due to magnification between the two images in terms of pixel shift. The soft decision regions allow for fairly significant displacement before image quality noticeably degrades.
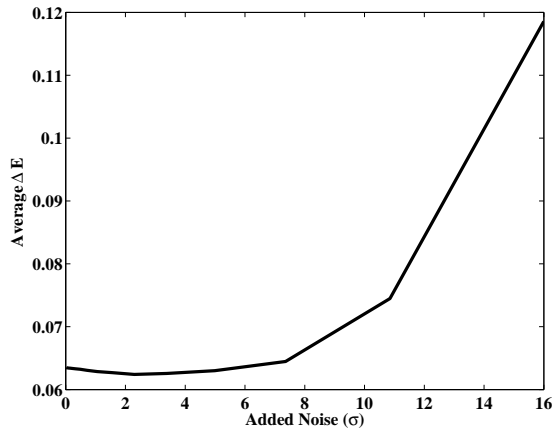

**Figure 4.** Plot of additive image noise versus average spatial ΔE.
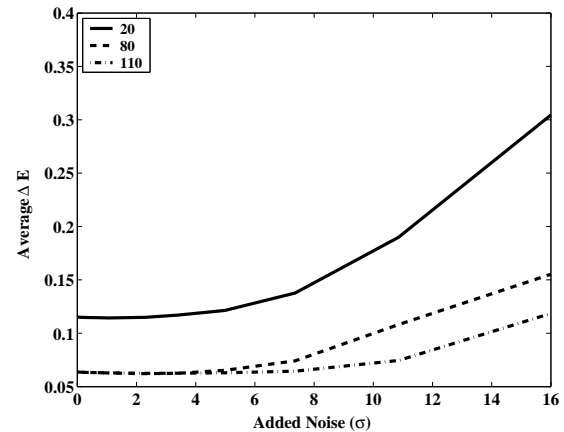
Figure 5. Effect of additive image noise for different sizes of aggregation, k.

4.2. Experimental Results

Now we present several examples of the performance of the algorithm despite the challenges presented by real lens and their adverse effects. Figure 8 shows two partially focused images of a typical situation in which only either the subject or background is in focus. The fused image is shown in Figure 9. Figure 10 shows both the near and far focused images and their corresponding fused image. Notice that some saturation occurs in the white areas of the eye chart and that both brightness and magnification variation effects are present. Such a scene is challenging in that saturation can introduce spurious high frequency components in a defocused image. However, the proposed algorithm is still able to generate a reasonably well-fused image due to the sigmoid-based majority filtering and subsequent smoothing. Finally, Figure 11 shows an extreme example in which both a close subject (the flower) and a distant background are photographed in the same scene. Even with such disparity between the two images, the technique is able to adequately fuse the focused regions as illustrated.
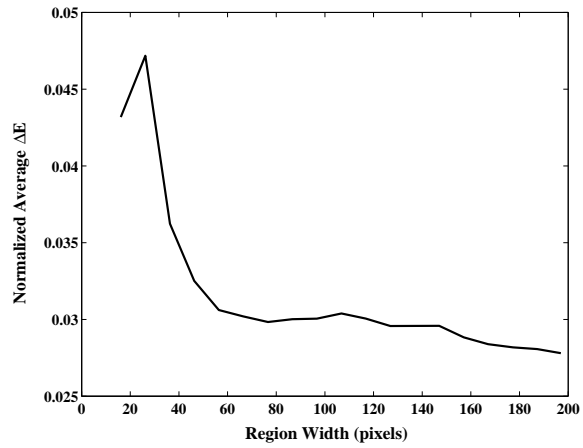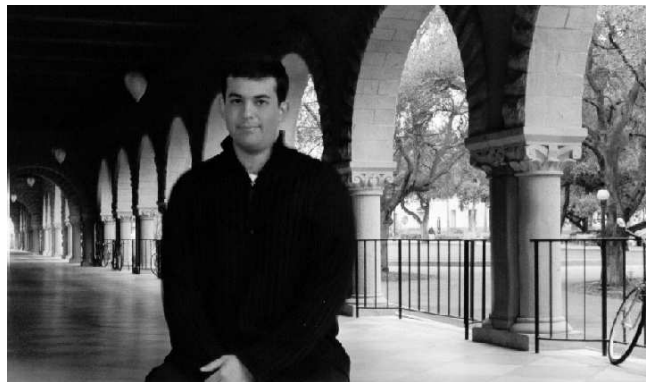
Figure 6. Plot of normalized average spatial ΔE versus region width.
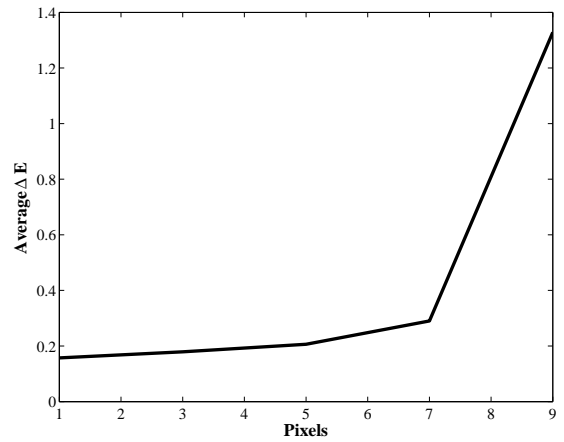
Figure 7. Effect of misregistration due to magnification in terms of pixel shift on average spatial ΔE.

Figure 8. Two partially focused images, near and far.

Figure 9. Reconstructed image using proposed algorithm.

Figure 10. The near focused, far focused and fused extended DOF image.

Figure 11. Example of extreme near focused, far focused and resultant fused image.

5. COMPUTATIONAL COMPLEXITY

The proposed algorithm has two basic parts. The first part is the computation of the spatial gradients, which can be as simple as the sum of two differencing operations. For a SIMD processor (or even desktop processors) only three to five unsigned addition per pixel would be required. The complexity for a $N \times N$ image would be of $O = \alpha N^2$ where α is a small number (for our example $\alpha = 2$). On the other hand, techniques that employ wavelet transforms would require convolution of the rows and columns and undersampling of them with at least one high pass and one low pass filter (e.g. Mallat filter banks[12]). Therefore, having $4L$ multiplications and additions per pixel is unavoidable where L is the length of the filter[13] (a reasonable choice is $L = \frac{N}{2}$). This means the computational complexity of the wavelet transform for even one stage is $O(N^3)$ and the multiplicative constant depends on the number of stages. For aggressive implementations of the transform the complexity can be reduced to $O(N^3 \log(N))$. Moreover, all other methods require multiplication in the convolution process and generally exhibit much larger multiplicative constants for their computational complexity order even when $O(N^2)$. In contrast, no multiplications are required for the convolution operations of the proposed method, since even the subsequent aggregation is performed solely via addition. Multiplication is required only at the final step if soft decision regions are used, since a linear combination of two pixel values at each point is then calculated.

The second major part of the algorithm involves performing the aggregation (majority filtering) which is basically a convolution operation. The optimal number of taps, k, in the majority filter is loosely image dependent, however 60 was found to work well in most cases. We perform the 2D convolution of the ones matrix with the image as two separable 1D convolutions of ones vectors. Accordingly, for an $N \times N$ image and a k-tap filter the complexity would be $O = \alpha k N^2$ where α is a small number (for our example $\alpha = \frac{4}{4} = 1$); the assumption is SIMD addition accepting four inputs. Moreover, this portion of the method can significantly gain from increased parallelism in the processor performing the calculation. As a matter of fact, for a width w of data in the SIMD add instruction and for an $m \times m$ processor array, the complexity would be of $O = 2\frac{k}{w}\frac{n^2}{m^2}$. This corroborates the idea that parallelism enhances convolution computations very efficiently.[14, 15] Indeed, the proposed algorithm leverages the functionality of parallel image processors quite efficiently.

6. CONCLUSION

In this paper, we proposed a method for the synthesis of extended depth of field images through the fusion of differently focused images. The method is predicated upon discrimination of focus quality using spatial image gradients, which are then aggregated using a majority filter. We analyzed the dependence of various optical effects on camera parameters and presented experimental results, along with simulations of the algorithm's performance despite such effects. Finally, the computational complexity of the algorithm was shown to be modest and the method compatible with the increasing parallelism of modern processors.

ACKNOWLEDGMENTS

REFERENCES

1. M. Subbarao, T. S. Choi, and A. Nikzad, "Focusing techniques," *SPIE Conference OE/Technology* **1823**, pp. 163–174, 1992.
2. X. Yang, W. Yang, and J. Pei, "Different focus points images fusion based on wavelet decomposition," *Proceedings of Third International Conference on Information Fusion* **1**, pp. 3–8, 2000.
3. Z. Zhang and R. S. Blum, "Image fusion for a digital camera application," *Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems and Computers* **1**, pp. 603–7, 1998.
4. W. Seales and S. Dutta, "Everywhere-in-focus image fusion using controllable cameras," *Proceedings of SPIE* **2905**, pp. 227–34, 1996.
5. K. Aizawa, K. Kodama, and A. Kubota, "Producing object-based special effects by fusing multiple differently focused images," *IEEE Transactions on Circuits and Systems for Video Technology* **10**(2), pp. 323–30, 2000.
6. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill Book Co., Singapore, 1996.
7. E. Krotkov, "Focusing (video camera, automatic)," *International Journal of Computer Vision* **1**, pp. 223–37, 1987.
8. E. Hecht, *Optics*, Addison Wesley, San Francisco, 2002.
9. G.Ligthart and F. Groen, "A comparison of different autofocus algorithms," *IEEE 6th International Conference on Pattern Recognition* **2**, pp. 597–600, 1982.
10. S. Nayar, "Shape from focus system for rough surfaces," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* , pp. 302–8, 1992.
11. X. Zhang and B. A. Wandell, "A Spatial Extension of CIELAB for Digital Color Image Reproduction," *Society for Information Display Symposium Technical Digest* **27**, pp. 731–734, 1996.
12. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Pattern Anal. and Machine Intell.* **11**, pp. 674–693, 1989.
13. P. Duhamel and O. Rioul, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Transactions on Information Theory* **38**, pp. 569 –586, 1992.
14. K. Batcher, "Design of a massively parallel processor," *IEEE Transactions on Computers* , pp. 810–836, 1980.
15. D. Helman and J. JaJa, "Efficient image processing algorithms on the scan line array processor," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**, pp. 47 –56, 1995.