ABSTRACT

Most of the focusing techniques need to estimate depth information for ensuring that the object of interest is at an appropriate distance for full frontal focus. Computational cameras which can variably focus different regions of the scene with large depth of field have been proposed. In this paper we propose a full auto-focusing algorithm using computational camera without involving any digital image restoration methods and just one input. The proposed computational camera uses multiple filter apertures corresponding to each color channel which can acquire three shifted views of a scene in the RGB color planes. We can make any region focused by appropriately shifting each color channel to be aligned. Depth map estimation is carried out to extract different regions from these channel shifted images which is later fused to produce the final image without any focal blur. Experimental results show performance and feasibility of the proposed algorithm for auto-focusing images with one or more differently out-of-focused objects.

Index Terms— Image restoration, image classification.

1. INTRODUCTION

Demand for digital auto-focusing techniques is rapidly increasing in many visual applications, such as camcorders, digital cameras, and video surveillance systems. Conventional cameras have come a long way in dealing with problems associated with focal settings and blur. Even though several steps have been taken, focal blur caused by varying distance of the object from the lens has been something that the conventional cameras still have to deal with. With focus set at either near, mid or far regions of the scene, the captured image tend to have only that particular region in focus whereas the remaining regions tend to be in out-of-focus. Post-processing steps in the form of blur restoration and multiple image fusion have been proposed to deal with the focusing problem.

Recently computational cameras have been developed that were capable of capturing additional information from the scene which when combined with post-processing can overcome several drawbacks of the imaging applications including: refocusing, increased dynamic range, depth-guided editing, variable lighting and reflectance, etc. The scope of this paper is to deal with the first factor that is the focal blur due to the mismatch in the distance of the object and focal length of the lens. In this paper we propose a combined hardware-software approach by which we tend to model the focal blur as channel dependent depth maps which are then used to remove the focal blur in images. The former (hardware) refers to the computational camera which employs a novel multiple filter aperture (FA) models for separating and distributing the blur into different color channels as shown in Fig. 1. The latter (software) refers to the computation or algorithm part which involves variable focusing, depth map estimation and fusion to generate a fully focused image from just single FA input. The block diagram of the proposed single view auto focusing algorithm is shown in Fig. 2.
2. RELATED WORKS

FA model: The idea of using multiple aperture lenses has been previously proposed using micro lens array and wave front coding [1] [2]. However, the imaging quality of these optical designs is fundamentally inferior to a camera system with a large single lens; the resolution of these small lens arrays is severely limited by diffraction. One can interpret the proposed optical configuration in this paper as an evolutionary step in which we deploy multiple detectors beneath each micro lens, instead of multiple array of sensors. More recent methods include single-lens multi-view image capture which proposed a post-exposure variable re-focusing [3]. The proposed FA model uses parallax cues instead of defocus cues and requires only color filters as additional optical elements to the lens without requiring multiple exposures.

Digital image processing methods for auto-focusing: Extensive work has been done using fusion and restoration-based methods for removal of out-of-focus blur in images [4-7]. Fusion algorithms using DCT, pyramids, and wavelets have been proposed to name a few where as restoration algorithms include blind deconvolution with no priori information as well as with PSF estimation. The auto-focusing method used in this paper needs only one input source image contrary to fusion-based methods and artifacts such as re-blurring and ringing due to restoration algorithm are also absent.

Depth from focus: Depth map algorithms have been extensively applied to stereo vision where the disparity estimate is computed as a correspondence measure through camera displacement [8]. Shape from focus can also estimate depth from a sequence of images taken by a single camera at different focus levels. Shape from focus methods employ spatial criteria including gray level variance (GLV), sum modified Laplacian (SML), Tanenbaum, mean method, curvature focal measure, etc [9, 10]. The proposed depth map estimation makes use of both the correspondence problem associated with stereo as well as spatial criteria [11] associated with focal blurs providing an accurate depth estimate.

3. MULTIPLE FILTER APERTURE (FA) MODEL

When the center of the aperture is not aligned on the optical axis, convergence is made off the optical axis, whose specific location depends on the distance between the lens and an object. The R, G, and B filters on FA are arranged so that their displacement with respect to center of the lens aligns with the row column displacement of the image sensor as shown in Fig. 3. By this arrangement, a scene point nearer or farther than the focused depth is captured as shift in R, G and B channels. The main advantage of the FA model is that it can provide an alternative method for the blur estimation in auto-focusing applications. Images acquired by using a conventional optical system have defocusing blur caused by a specific PSF. On the other hand the proposed multiple FA model the auto-focusing problem turns in to the alignment of R, G, and B channels with various depths of field. For shifting and aligning color channels we need to find the optimal pixel-of-interest at different positions in the image according to their focal measures. The pixel-of-interest can be referred to as a focal point pixel, around which channel shifting and alignment is carried out. For a given region, we then select the focal point pixel either from the center of the region or the pixel with the lowest focus measure. Similar operations repeat for differently selected focal point regions.

4. AUTO FOCUSING USING DEPTH MAPS

The color channel shifting and alignment process discussed in the previous section clearly states two distinctive advantages of the multiple FA model as; (i) each channel suffers distortion in proportion to the position of the aperture and (ii) the focal point change among various regions of the image to produce n-subimages with varying focal blur as

\[ I_{(R,G,B)} = \left[ I_{(R,G,B)}^{(1)}, ..., I_{(R,G,B)}^{(n)} \right], \]

where \( I_{(R,G,B)}^{(n)} \) represents n-subimages at varying focal points as shown in Fig. 4. A fully focused image can be generated from these n-subimages using estimated depth maps. In this paper we have adopted region-based depth map estimation in conjunction with segmentation. Segmentation allows us to decompose the reference image into regions of homogenous color, region perimeter based on the edge information which in turn supports the disparity estimate. We assume the pixel colors within a local window \( w(x,y) \) belong to one cluster and we use the magnitude of
the clusters elongation as the correspondence measure. More specifically, we consider a set \( P_t(x, y; d) \) of pixel colors with hypothesized disparity \( d \) as
\[
P_t(x, y; d) = \{(I_{Rt}(s + d, t), I_{Gr}(s, t - d), I_{Bt}(s - d, t)) | (s, t) \in \Omega(x, y)\}
\]
and search for \( d \) that minimizes the following color alignment measure.
\[
L(x, y; d) = \lambda_n \lambda_r \lambda_b / \sigma^r_n \sigma^g_n \sigma^b_n,
\]
where \( \lambda_n, \lambda_r, \) and \( \lambda_b \) respectively represent the eigenvalues of the covariance matrix \( \Sigma \) of the color distribution \( P_t(x, y; d) \) and \( \sigma_n^r, \sigma_n^g, \) and \( \sigma_n^b \) respectively the diagonal elements of \( \Sigma \) [12]. This measure gives an abstract disparity map in the disparity search range \([-10, 10]\) which can be used to develop the error energy matrix and calculate the depth map. The depth map extracted using this method alone is, however, insufficient for the proper extraction of focus maps needed to generate the fully focused image. Hence we have adopted segmentation as an additional criterion for more accurate extraction of the depth map \( \alpha(x, y) \) as,
\[
\alpha(x, y) = L(x, y) \cap M(x, y),
\]
where \( M(x, y) \) represents the mean-shift segmentation result. Using the proposed method focused regions \( I^{FR}_{(R,G,B)}, \ldots, I^{FR}_{(R,G,B)} \) from \( n \)-subimages can be extracted as:
\[
\]
The depth information gives us perimeter of different objects at varying distance from the camera. Using this depth map along with channel shifted images we can easily generate a fully focused image. The fusion process combines different regions from different channel shifted regions using the depth map information as follows:
\[
\]
where \( I^F_{(R,G,B)} \) represents the fully focused image.

5. EXPERIMENTS

For the experiments, we used a commercial gelatin filter (Kodak-Wratten Filter – G-58, B-47, and R-25) with sensors representing red, green, and blue spectral wavelengths. Fig. 4(a) represents a typical FA multiple object image captured with focus on ‘spring’ object on left. Figs. 4(b) and (c) represent channel shifting performed to shift the focus from ‘spring’ to ‘cowboy’ followed by ‘robot’ using Fig. 4(a). In order to combine Fig 4(a)-(c) to a single image with full focus the depth map estimation has been used as shown in Fig. 5. By using the generated depth map we can extract boundary of each object which in turn retrieve and combine pixels from various channel shifted images.

The auto-focusing algorithm was compared against standard restoration and fusion-based methods as shown in Table 1. Another comparison in the sense of pixel error count (PEC) and disparity error map (DEM) are shown in Tables 2 and 3. PEC is obtained by calculating the number of mis-classified pixels with ground truth segmentation map. Shape from focus measures including sum modified Laplacian (SML), gray level variance (GLV), Tanenbaum and Tanengrad were used in PEC comparison. It can be seen that the proposed depth map had comparable results with Tanenbaum for PEC but outperforming other measures significantly. The DEM was used to find the disparity error average for pre-defined range \([-10, 10]\) when compared to ground truth data. Stereo vision methods including sum and gradient absolute differences (SAD and GRAD), color and Bayes disparity were tested. In case of DEM the performance showed vast improvements over SAD and GRAD measures whereas comparatively competitive with color and Bayes disparity.

6. CONCLUSION

In this paper we proposed an auto focusing algorithm for removing focal blur in images. The proposed algorithm is ideal for situations when the focal range of a scene is distributed over varying distance from the camera. Future works will be addressed to building a camera prototype that will be automatically able to re-focus the entire image without any user intervention and better color quality.

ACKNOWLEDGEMENT

This work was supported by National Research Foundation (NRF) grant funded by the Korea government (MEST) (2009-0081059), by the Industrial Technology Development program of the Ministry of Knowledge and Economy (MKE) of Korea, and by Ministry of Culture, Sports and Tourism (MCST) and Korea Culture Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2009.

REFERENCES

Fig. 4. Input images with different focal lengths and the result of proposed FA model; (a) input source image focused on the spring object (b) image with focus shifted from the spring to cowboy, (f) image with focus shifted from the cowboy to the robot.

Fig. 5. Experimental results: (a, d, g) input out-of-focus image captured using the multiple FA model. (b, e, h) depth map extracted using color channel dependency, and (c, f, i) focused images obtained by shifting focal point at different sections of the image followed by color channel alignment and fusion.

Table. 1. Comparisons for various auto-focusing methods

<table>
<thead>
<tr>
<th>AF Method</th>
<th>Pari</th>
<th>Mode</th>
<th>Input</th>
<th>Operation</th>
<th>RMSE</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiener Filter</td>
<td>PSF</td>
<td>Gray</td>
<td>1</td>
<td>Pixel</td>
<td>12.35</td>
<td>23.36</td>
</tr>
<tr>
<td>Iterative Filter</td>
<td>NIL</td>
<td>Gray</td>
<td>1</td>
<td>Pixel</td>
<td>8.56</td>
<td>26.32</td>
</tr>
<tr>
<td>Pyramid Fusion</td>
<td>NIL</td>
<td>Gray, Color</td>
<td>At least 2</td>
<td>Window and Pixel</td>
<td>5.68</td>
<td>28.42</td>
</tr>
<tr>
<td>Wavelet Fusion</td>
<td>NIL</td>
<td>Gray, Color</td>
<td>At least 2</td>
<td>Window and Pixel</td>
<td>5.02</td>
<td>29.95</td>
</tr>
<tr>
<td>Proposed</td>
<td>Nil</td>
<td>Color</td>
<td>1</td>
<td>Window and Pixel</td>
<td>8.06</td>
<td>26.41</td>
</tr>
</tbody>
</table>

Table. 2. Pixel error count (PEC) measure for focal operators

<table>
<thead>
<tr>
<th>Test</th>
<th>SML</th>
<th>Tanenbaum</th>
<th>GLV</th>
<th>Tenengrad</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deer</td>
<td>1.05</td>
<td>0.95</td>
<td>1.35</td>
<td>0.99</td>
<td>0.52</td>
</tr>
<tr>
<td>Toys</td>
<td>0.99</td>
<td>0.61</td>
<td>1.30</td>
<td>0.93</td>
<td>0.41</td>
</tr>
<tr>
<td>Sim</td>
<td>1.12</td>
<td>0.82</td>
<td>1.17</td>
<td>0.81</td>
<td>0.50</td>
</tr>
<tr>
<td>Doll</td>
<td>0.82</td>
<td>0.67</td>
<td>1.03</td>
<td>0.68</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Table. 3. Disparity error measure (DEM) comparisons for depth algorithms

<table>
<thead>
<tr>
<th>Test</th>
<th>SAD</th>
<th>GRAD</th>
<th>Color</th>
<th>Bayes</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deer</td>
<td>14.56</td>
<td>12.10</td>
<td>7.20</td>
<td>6.30</td>
<td>6.50</td>
</tr>
<tr>
<td>Toys</td>
<td>18.31</td>
<td>14.32</td>
<td>8.60</td>
<td>8.40</td>
<td>6.21</td>
</tr>
<tr>
<td>Sim</td>
<td>12.08</td>
<td>11.47</td>
<td>7.01</td>
<td>5.64</td>
<td>5.40</td>
</tr>
<tr>
<td>Doll</td>
<td>12.32</td>
<td>12.19</td>
<td>9.18</td>
<td>7.82</td>
<td>6.11</td>
</tr>
</tbody>
</table>