MRT Letter: Recovering Weak-Textured Surfaces Using Image Focus

MANNAN SAEED MUHAMMAD AND TAE-SUN CHOI* 
Signal and Image Processing Laboratory, School of Information and Mechatronics, Gwangju Institute of Science and Technology (GIST), 261 Cheomdan-Gwagiro, Buk-Gu, Gwangju 500-712, South Korea

KEY WORDS shape from focus; texture; 3D shape recovery

ABSTRACT In nature, objects have partially weak texture and their shape reconstruction using focus based passive methods like shape from focus (SFF), is difficult. This article presents a new SFF algorithm which can compute precise depth of dense as well as weak textured objects. Segmentation is applied to discard wrong depth estimate and then later interpolating them from accurate depth values of their neighbors. The performance of the proposed method is tested, using different image sequences of synthetic and real objects, with varying textures. Microsc. Res. Tech. 72:703–706, 2009. © 2009 Wiley-Liss, Inc.

INTRODUCTION Shape from focus (SFF) is an optical passive method for 3D shape recovery of an object from its 2D images, using a sequence of images, taken by a single camera at different focus levels, to compute the depth of the object in the scene. At first, focus measure (FM) is applied to compute focus quality for each pixel in the image sequence. Sun et al. (2004) and Sundaram et al. (1997) have reported many famous FMs. An initial estimate is obtained by maximizing the FM along the optical axis, and then it is enhanced by using an approximation technique. In literature, a variety of approximation techniques for SFF, including SFF,TR (Subbarao and Choi, 1995), SFF,NN (Asif and Choi, 2001), and SFF,DP (Ahmad and Choi, 2005) have been reported.

SFF techniques are based on the assumption about the presence of prominent texture in the scene. Mostly existing methods perform well in rich textured scenes. However, their performance is deteriorated when the objects have fewer variations in the texture. In this article, we propose a new SFF method to approximate the 3D shape of the weak textured objects. To find the initial depth, for each object point, the intensities of pixels are modified and approximated by a Gaussian model, then, segmentation is used to discard the depth estimates of weak textured areas. Finally, 3D shape is recovered by interpolating the missing areas through accurate depth values of their neighbors. The performance of the proposed method is tested using different image sequences.

PROPOSED ALGORITHM

When an object point on the object is best focused, the corresponding pixel behavior is explained by Figure 1. For each object point the gray level values of the pixels are taken in a vector \( i^jP \left( i^jP \in \mathbb{R}^n \right) \), and are modified by Eq. (1).

\[
[i^j m_k]_{n1} = \left[ \sum_{\Omega} \left( i^j p_k - \max\{i^j p_1, i^j p_n\} \right)^2 \right]_{n1}
\]  

for all \( 1 \leq k \leq n \), where \( n \) is the total number of images in the image sequence \( I_i(X \times Y) \), \( i^jp_k \), and \( i^jm_k \) are the \( k \)th value in the original pixel intensity vector \( i^jP \) and modified pixel intensity vector \( i^jM \), respectively. \( i \) and \( j \) are the \( x \) and \( y \) positions of the pixel in spatial domain, and \( \Omega \) is the summing window \((3 \times 3)\). At the initial stage, the vector \( i^jM \) is approximated by the Gaussian model defined as \( G(k)_{i,j} = \Lambda \exp\{-{(k-B)^2/C^2}\} (i,j) \) where \( \Lambda \) is the peak, \( B \) is the mean, and \( C/2 \) is the standard deviation. The corresponding value of \( B \) is taken as the initial depth of the object point.

Presence of spikes is observed in the initial depth estimation, mainly due to the areas on the surface where the object has weak texture. One solution for this problem is to discard these wrong depth values and later estimating them from the accurate depth values of their neighbors. For this, segmentation based on threshold selection method is applied to separate the noise (spikes) from the initial depth map \((D)\). This depth map is then divided into the small window patches each of size \( \omega \times \omega \) (thus the total number of small window patches is \( N = (X \times Y)/\omega^2 \), and the histogram of each window is calculated. The histogram indicates the occurrence frequency of the object points with similar depth values. If we define probability \( p \) of occurrence of object points with similar depth values, then \( p = h/\omega^2 \); where \( \omega^2 \) is the total number of object points in the window patch, and \( h \) is the number of points of same depth. The occurrence probability of objects points belonging to the cluster \( C_k \) is given by Eq. (2), where \( K \) is total number of clusters.

*Correspondence to: Tae-Sun Choi, Signal and Image Processing Laboratory, School of Information and Mechatronics, Gwangju Institute of Science and Technology (GIST), 261 Cheomdan-Gwagiro, Buk-Gu, Gwangju 500-712, South Korea. E-mail: tschoi@gist.ac.kr

Received 3 March 2009; accepted in revised form 28 July 2009

Contract grant sponsor: Bio Imaging Research Center at GIST.

DOI 10.1002/jemt.20768

Published online 1 September 2009 in Wiley InterScience (www.interscience.wiley.com).
Fig. 1. Pixel behavior in an image sequence. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Fig. 2. Experimental objects (left to right) two image sequence of simulated cone, LCD-TFT image sequence, original depth map of simulated cone.

Fig. 3. Depth estimation by proposed method from top left and clockwise. Normalized pixel and modified-pixel intensity vector, initial depth map of simulated cone; depth map with discarded depth, window patch example, window patch with discarded depth and approximated surface of the window patch.
The cluster with the highest probability is kept and the rest are discarded. The missing parts of the result-
ant depth map are then approximated by using inter-
polation, and the complete 3D shape is recovered.

RESULTS AND DISCUSSION

Three different image sequences have been used for the experiments. Two simulated cone images, with dif-
ferent textures, generated by computer simulations (Subbarao and Lu, 1994), and LCD images which are micro-
scopic images of LCD-TFT filter (Asif and Choi, 2001), as shown in Figure 2.

The dimensions of simulated cone images are $I_{100}(265 \times 265)$ and LCD images are $I_{90}(305 \times 305)$. Let us consider an object point (118,208) in the second image sequence of simulated cone. The pixel intensity vector and its modified intensity vector are shown in the Figure 3 (top left). The mean of the fitted Gaussian Model is found as 29. In this way the depth of the entire object is found. Figure 3 (top middle) shows the initial depth map of the simulated cone. The spikes in the figure are the noise (false depth estimation). The calculated depth map is then divided into small win-
dows, and the histogram is calculated for each window. Figure 3 (bottom left to right) shows the window patch taken from initial depth map, the empty patches with

\[
P(C_k) = \sum_{m=1}^{n^2} P_m, \quad \sum_{k=1}^{K} P(C_k) = 1 \quad (2)
\]

The cluster with the highest probability is kept and the rest are discarded. The missing parts of the result-
ant depth map are then approximated by using inter-
polation, and the complete 3D shape is recovered.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Method</th>
<th>Image seq. 1</th>
<th>Image seq. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>SFF TR</td>
<td>7.8160</td>
<td>6.9925</td>
</tr>
<tr>
<td></td>
<td>SFF DP</td>
<td>5.6456</td>
<td>4.8084</td>
</tr>
<tr>
<td></td>
<td>Proposed</td>
<td>5.0686</td>
<td>0.4840</td>
</tr>
<tr>
<td></td>
<td>SFF TR</td>
<td>0.9019</td>
<td>0.9125</td>
</tr>
<tr>
<td></td>
<td>SFF DP</td>
<td>0.9511</td>
<td>0.9675</td>
</tr>
<tr>
<td></td>
<td>Proposed</td>
<td>0.9532</td>
<td>0.9996</td>
</tr>
</tbody>
</table>

Fig. 4. Depth map by different methods. Top row simulated cone and LCD-TFT using proposed method. Bottom row simulated cone and LCD-TFT using SFF.DP.

Fig. 5. Comparison of different methods of SFF with proposed method (top: RMSE, bottom: correlation).
discarded depth and the interpolated surface. Figure 3 (top right) shows the depth map of simulated cone 2, with empty patches.

In Figure 4 the reconstruction of the simulated cone and LCD using proposed method (after interpolating the missing data) are shown along with the results of SFF.DP. Two statistical metrics, root mean squared error (RMSE) and correlation are applied to evaluate the performance. Table 1, Figures 4 and 5, clearly demonstrate the effectiveness of the proposed method.

**CONCLUSION**

In this article we have proposed a new SFF method to approximate the 3D shape of the object with less textured surfaces. Segmentation is used to discard wrong depth values which are caused mainly due to weak textured areas in the scene. Later, they are interpolated from the correct depth values of their neighbors. The results demonstrate the effectiveness of the proposed method.

**REFERENCES**