

# Calculating depth maps from digital holograms using stereo disparity

Tomi Pitkäaho<sup>1</sup> and Thomas J. Naughton<sup>1,2,\*</sup>

<sup>1</sup>University of Oulu, RFMedia Laboratory, Oulu Southern Institute, Vierimaantie 5, 84100 Ylivieska, Finland

<sup>2</sup>Department of Computer Science, National University of Ireland Maynooth, Maynooth, County Kildare, Ireland

\*Corresponding author: tomn@cs.nuim.ie

Received March 30, 2011; revised April 28, 2011; accepted April 29, 2011;

posted May 2, 2011 (Doc. ID 145095); published May 26, 2011

Depth extraction is an important aspect of three-dimensional (3D) image processing with digital holograms and an essential step in extended focus imaging and metrology. All available depth extraction techniques with macroscopic objects are based on variance; however, the effectiveness of this is object dependent. We propose to use disparity between corresponding points in intensity reconstructions to determine depth. Our method requires a single hologram of a scene, from which we reconstruct two different perspectives. In the reconstruction the phase information is not needed, which makes this method useful for in-line digital holography. To our knowledge disparity based 3D image processing has never been proposed before for digital holography. © 2011 Optical Society of America

OCIS codes: 090.0090, 090.1995, 100.6890, 100.0100, 100.2000.

Digital holography makes it possible to record and reconstruct different perspectives of real-world three-dimensional (3D) objects [1,2]. Where the reconstructed phase is too noisy for phase unwrapping due to, for example, the presence of the twin as in some in-line single-shot architectures, or the long recording distance from a diffuse object causing too large phase jumps, or otherwise, the intensity can be used to extract depth maps. These techniques are important for extended focus imaging and metrology. In off-axis digital holographic microscopy, the reconstructed phase of the hologram can be used for accurate depth measurements [3–5]. However, reconstructed phase is not as useful from single digital holograms of macroscopic objects. Instead, different methods for depth map extraction have been introduced [6–8]. These methods are based on calculating variance, or otherwise estimating high spatial frequencies, at each region and at each depth. Besides being time consuming, the performance of each is highly object dependent. Variance needs a highly textured object to determine when it is in focus. Sharp edges are not enough unless between the edges there is also a texture. This is because regions of constant (either high or low) reflectivity will produce a low variance. We propose to use a different approach borrowed from the field of computer vision: stereo disparity. A single hologram can be used and no phase is needed in the reconstruction, which makes it useful for in-line single-shot digital holography.

A human's left and right eyes receive different images because of binocular parallax. The difference between locations of corresponding features in both images is called disparity [9,10], and it has been claimed that the visual system uses correlation to calculate disparity [11]. Disparity can correspond to a depth measurement [as illustrated in Fig. 1(a)], which gives information about three-dimensionality of a scene. Stereo disparity has been widely investigated in computer vision, but to our knowledge this is the first time it has been used in digital holography. Although we use correlation, any matching technique, such as one based on features, can be used.

Given a hologram  $f(x, y)$  of a 3D scene, centered on coordinates  $x = 0$ ,  $y = 0$ , a left and right perspective of the scene can be obtained from  $L(x, y) = \mathcal{F}_z[\Lambda(x, y)]$  and  $R(x, y) = \mathcal{F}_z[P(x, y)]$ , respectively, where

$$\Lambda(x, y) = \begin{cases} 0, & \text{if } x \geq 0 \\ f(x, y) & \text{otherwise} \end{cases}, \quad (1)$$

$$P(x, y) = \begin{cases} 0, & \text{if } x < 0 \\ f(x, y) & \text{otherwise} \end{cases}, \quad (2)$$

and where  $\mathcal{F}$  represents light propagation such as a Fresnel intensity approximation at distance  $z$  given by

$$F_z(x, y) = \left| \frac{-i}{\lambda z} H(x, y) * \exp \left[ i\pi \frac{(x^2 + y^2)}{\lambda z} \right] \right|^2, \quad (3)$$

where  $H$  can be either  $\Lambda$  or  $P$ ,  $\lambda$  is wavelength of the light, and  $*$  denotes a convolution operation.

By using these two perspectives  $L$  and  $R$ , we calculate the disparity space image as follows. For each pixel in the left image  $L(x, y)$ , we estimate its corresponding position  $(x', y)$  in the right image. In computer vision, the translation and rotation of one camera relative to the other must be found, and compensated for, in order to triangulate points in both captured images. In digital holography, this relation is known *a priori* since the two halves of the hologram lie in the same plane and an identical numerical reconstruction operation can be applied to each; therefore, we only have to search in the horizontal direction because we can ensure that there is a perfectly horizontal baseline between perspectives. The disparity for that pixel  $(x, y)$  is the difference in pixel locations  $x-x'$ , and the 2D matrix of disparity values, one for each pixel  $(x, y)$ , is the disparity space image (DSI).

More formally, given left and right 2D perspectives of a 3D scene,  $L_{x,y}$  and  $R_{x,y}$ , respectively,

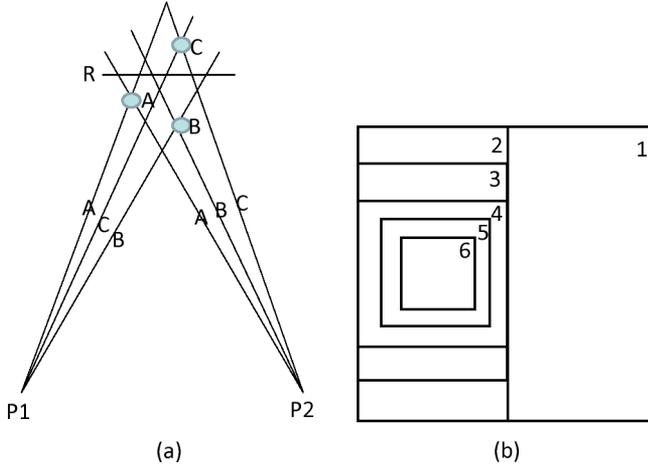


Fig. 1. (Color online) (a) Disparity with holograms: P1, P2, different perspectives; A, B, and C, points on the epipolar line in the reconstructed intensity image; R, reconstruction plane. (b) Illustration of size and position (perspective) of different window sizes, e.g. (1)  $2048 \times 2048$  pixels, (2)  $2048 \times 1024$ , and (6)  $512 \times 512$ .

$$DSI(x, y) = \begin{cases} \operatorname{argmax}_u(C_{x,y}(u)), & \text{if } \max_u(C_{x,y}(u)) > \tau \\ \text{undefined}, & \text{otherwise,} \end{cases} \quad (4)$$

where  $C_{x,y}$  is the normalized cross-correlation function used to estimate the disparity. Operator  $\operatorname{argmax}$  finds the disparity  $u$  that maximizes this correlation.  $\tau$  is an experimentally chosen threshold that is 0.3 in our case and is used to decide if each correlation peak is meaningful or noise. For each pixel the normalized cross correlation is calculated over a rectangular block of size  $(2t+1)^2$  pixels as

$$C_{x,y}(u) = \frac{\sum_{x,y} \tilde{L}_{x,y} \tilde{R}_{x,y}(u)}{[\sum_{x,y} (\tilde{L}_{x,y})^2 \sum_{x,y} (\tilde{R}_{x,y}(u))^2]^{1/2}}, \quad (5)$$

where  $\tilde{L}_{x,y} = L(x, y) - \bar{L}_{x,y}$  and  $\tilde{R}_{x,y}(u) = R(x-u, y) - \bar{R}_{x,y}(u)$ . In these,  $\bar{L}_{x,y} = \frac{1}{(2t+1)^2} \sum_{k=x-t}^{x+t} \sum_{l=y-t}^{y+t} L(k, l)$  is the mean value of the pixel values of the left image within the block of pixels around  $(x, y)$ , and  $\bar{R}_{x,y}(u) = \frac{1}{(2t+1)^2} \sum_{k=x-u-t}^{x-u+t} \sum_{l=y-t}^{y+t} R(k, l)$  is the mean value of the pixel values of the right image within the offset block around  $(x-u, y)$ ,  $u = 0, \dots, T$  is the offset, and  $T$  is the search length to be considered.

Our disparity algorithm is able to identify all of the objects at different depths in the Middlebury stereo data set tsukuba [10], as shown in Fig. 2. Next, we present the DSI (see Fig. 3) calculated using two different views from a hologram of an electronic chip component captured with in-line phase-shifting interferometry [12]. The hologram contained  $2048 \times 2032$  pixels with  $7.4 \mu\text{m}$  pixel pitch. Pin heights are 4.2 mm, the front pins were 163 mm from the sensor, and the distance between the front and back pins was 8 mm. This object was chosen because from the intensity image alone it is not obvious which pins are in front and which are behind. In the DSI the distances can

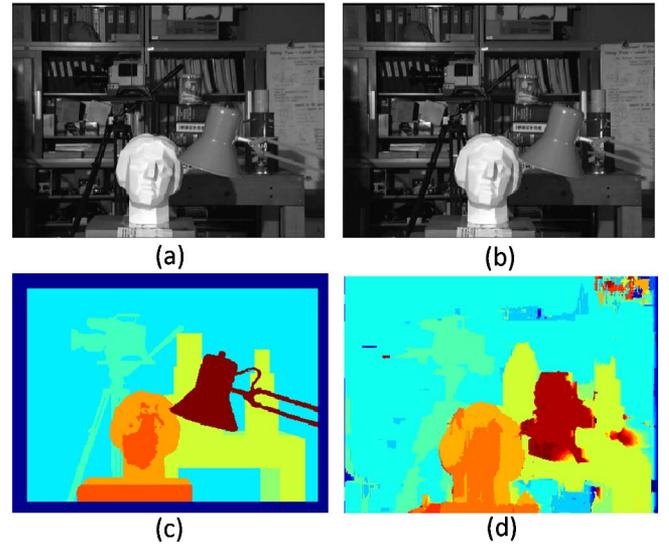


Fig. 2. (Color online) Middlebury's test image comparison: (a) left perspective, (b) right perspective, (c) ground truth, (d) output of our algorithm.

be readily identified. When calculating disparity with digital hologram reconstructions, the reconstruction depth affects the relative disparity values. Objects closer to the in-focus plane yield smaller absolute disparity values compared to objects farther from the in-focus plane. The sign on the disparity tells whether the object is in front of or behind the in-focus plane. Note that the objects in the DSI are enlarged due to the cross-correlation algorithm; they are appropriately reduced in subsequent figures using an image erosion operation. The fourth subobject from the right is one pin partially occluding another. Although this is not at all visible in the intensity reconstruction, the disparity algorithm responds appropriately. The reconstruction of a second object is shown in Fig. 4(a), and a relief plot of the DSI is shown in Fig. 5. This time, the front pins were 178 mm from the sensor. For reconstructions of size  $2048 \times 2032$ , block size of  $21 \times 21$ , and search length of 42 pixels, the technique takes 128 min using Matlab on a regular laptop computer.

While the window position in the hologram [see Fig. 1(b)] determines perspective, the window size controls resolution and depth of field in the reconstruction. A larger window means increased quality reconstructions and should allow one to more accurately find

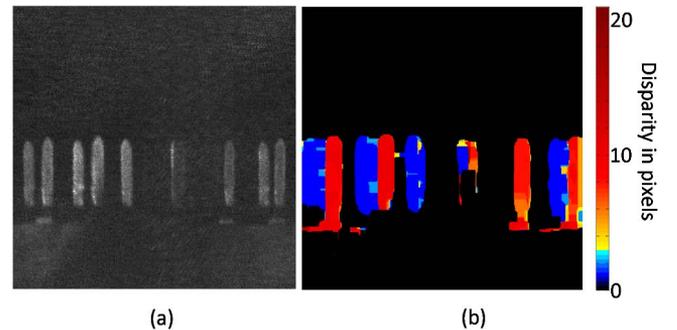


Fig. 3. (Color online) Input image (a) from left perspective, and (b) the calculated DSI. Increased absolute disparity means those parts of the object are farther away from the reconstruction distance.

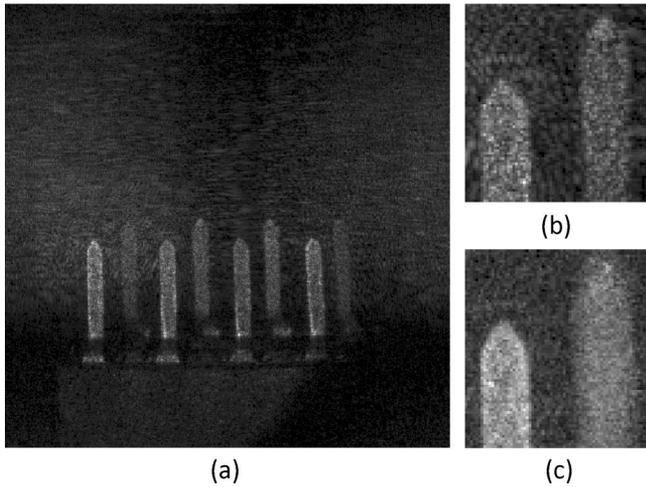


Fig. 4. Reconstruction using a different hologram. (a) Left perspective with window size  $512 \times 512$  pixels, (b) zoomed part of (a), (c) zoomed part with larger  $2048 \times 1024$  window showing increased quality for the left pin but lower depth of field.

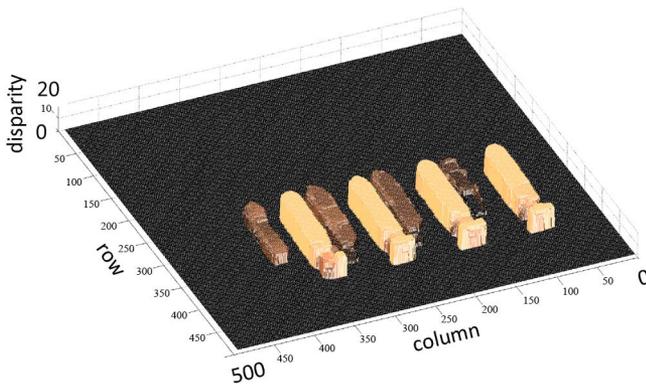


Fig. 5. (Color online) Relief plot of the DSI from object shown in Fig. 4(a).

correspondences between perspectives. However, for 3D scenes where the object extends along the optical axis, a smaller window means an increased depth of field [see Figs. 4(b) and 4(c)] and therefore can result in a more accurate disparity calculation because more objects are in focus. In addition, a smaller window allows a larger disparity range in the DSI and therefore more discrimination between depths. A smaller window also results in a quicker calculation time. We therefore verify how small the hologram window can be, with the trade-off that reconstruction noise will increase as fewer hologram pixels are used for the reconstruction. In Fig. 6, DSIs with different depths of field (different window sizes) are compared. Although we have not applied speckle reduction to our reconstructions, this is likely to enhance the accuracy of the depth map calculation.

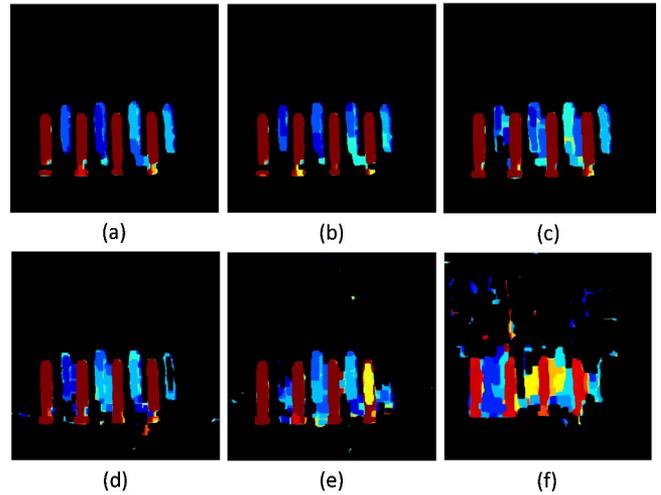


Fig. 6. (Color online) DSIs with different window sizes: (a)  $2048 \times 1024$ , (b)  $1536 \times 1024$ , (c)  $1024 \times 1024$ , (d)  $768 \times 768$ , (e)  $512 \times 512$ , and (f)  $128 \times 128$  pixels.

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 216105 ("Real 3D"), Academy of Finland, and Science Foundation Ireland under the National Development Plan. We thank Emmanouil Darakis for capturing the digital holograms and discussions.

## References

1. T. Kreis, *Handbook of Holographic Interferometry* (Wiley-VCH, 2005).
2. Y. Frauel, T. J. Naughton, O. Matoba, E. Tajahuerce, and B. Javidi, *Proc. IEEE* **94**, 636 (2006).
3. F. Charrière, A. Marian, F. Montfort, J. Kühn, T. Colomb, E. Cuche, P. Marquet, and C. Depeursinge, *Opt. Lett.* **31**, 178 (2006).
4. P. Ferraro, S. Grilli, D. Alfieri, S. De Nicola, A. Finizio, G. Pierattini, B. Javidi, G. Coppola, and V. Striano, *Opt. Express* **13**, 6738 (2005).
5. Y. Morimoto, T. Nomura, M. Fujigaki, S. Yoneyama, and I. Takahashi, *Exp. Mech.* **45**, 65 (2005).
6. L. Ma, H. Wang, Y. Li, and H. Jin, *J. Opt. A: Pure Appl. Opt.* **6**, 396 (2004).
7. S. Frey, A. Thelen, S. Hirsch, and P. Hering, *Appl. Opt.* **46**, 1986 (2007).
8. C. P. Mc Elhinney, J. B. McDonald, A. Castro, Y. Frauel, B. Javidi, and T. J. Naughton, *Opt. Lett.* **32**, 1229 (2007).
9. D. Marr and T. Poggio, *Science* **194**, 283 (1976).
10. D. Scharstein and R. Szelisk, *Intl. J. Comp. Vision* **47**, 7 (2002).
11. L. K. Cormack, S. B. Stevenson, and C. M. Schor, *Vision Res.* **31**, 2195 (1991).
12. E. Darakis, T. J. Naughton, and J. J. Soraghan, *Appl. Opt.* **46**, 4579 (2007).