Segmentation of three-dimensional objects from background in digital holograms

C. P. Mc Elhinney[1] and J. B. McDonald[1] and Albertina Castro[2] and Yann Frauel[3]
and Bahram Javidi[4] and T. J. Naughton[1,5]

[1]Dept. of Computer Science, National University of Ireland, Maynooth, Co. Kildare , Ireland
[2]Instituto Nacional de Astrofísica, Óptica y Electrónica, Apdo. Postal 51, Puebla, Puebla 72000, Mexico
[3]Departamento de Ciencias de la Computación, Instituto de Investigaciones en Matemáticas Aplicadas y
Sistemas, Universidad Nacional Autónoma de México, México D.F. 04510, Mexico
[4]Department of Electrical and Computer Engineering, University of Connecticut, U-157, Storrs, CT 06269,
USA
[5]University of Oulu, RFMedia Laboratory, Oulu Southern Institute, Vierimaantie 5, 84100 Ylivieska, Finland
conormce@cs.nuim.ie; johnmcd@cs.nuim.ie; tomn@cs.nuim.ie;

Abstract

We present a technique for performing segmentation of three-dimensional objects encoded using in-line digital holography from the scenes background. We create a volume of reconstructions through numerically reconstructing a digital hologram at a range of depths. For each reconstruction a variance map is created through calculating variance about a neighbourhood for each of the reconstructions pixels. We can then classify a pixel as object or background by thresholding the maximum variance of every pixel over all depths. We present segmentation results for objects of low and high contrast.

Keywords: digital holography, segmentation, focus detection, three-dimensional image processing

1. Introduction

Holography is an established technique for recording and reconstructing real-world three-dimensional (3D) objects. Digital holography [12, 18, 19, 20] and holographic image processing [17, 16] have recently become feasible due to advances in megapixel CCD sensors with high spatial resolution and high dynamic range. We use phase shift interferometry [20] to capture our in-line digital holograms (DHs), which are in an appropriate form for data transmission and digital image processing. We are interested in segmenting object regions from background regions in digital holographic reconstructions. Segmentation is generally the first step to object recognition, but also has applications in video compression and motion tracking [8]. Segmentation has also recently been applied to 3D data, or layers of two-dimensional reconstructions, in tomography and digital holographic microscopy [4].

In Sect. 2, we describe how 3D objects are recorded using phase-shift digital holography. We then describe approaches for the focus detection of digital holographic scenes and how they can be used for segmentation purposes in Sect. 3. Section 4 details our method for object segmentation using a depth-from-focus algorithm. Our results are presented in Sect. 5 where we demonstrate the segmentation of multiple digital holographic scenes. We conclude in Sect. 6.

2. Phase-Shift Digital Holography

We record Fresnel fields with an optical system based on a Mach-Zehnder interferometer (see Fig. 1). A linearly polarised Argon ion ($532.8 \, \text{nm}$) laser beam is expanded and collimated, and divided into object and reference beams. The object beam illuminates a reference object placed at a distance of approximately $d = 350 \, \text{mm}$ from a 10-bit 2048×2032 pixel Kodak Megaplus CCD camera. Let $U_0(x, y)$ be the complex amplitude distribution immediately in front of the 3D object. The linearly polarised reference beam passes through half-wave plate RP_1 and quarter-wave plate RP_2. By selectively removing the plates we can achieve four phase shift permutations of 0, $-\pi/2$, $-\pi$, and $-3\pi/2$. The reference beam combines with the light diffracted from the object and forms an interference pattern in the plane of the camera. At each of the four phase

shifts we record an interferogram. We use these four real-valued images to compute the camera-plane complex field
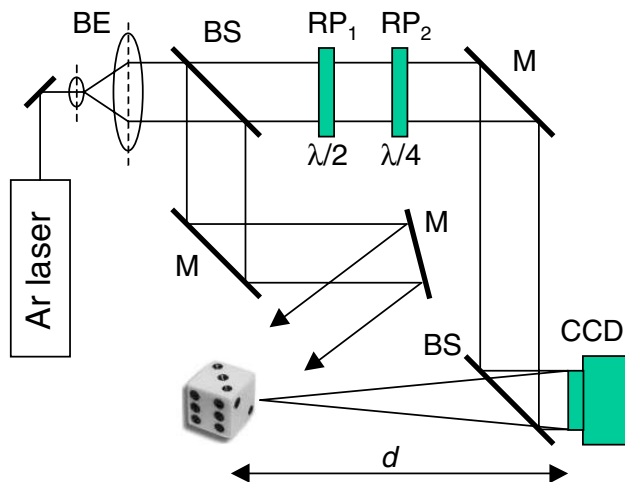

Figure 1. Experimental setup for PSI: BE, beam expander; BS, beam splitter; RP, retardation plate; M, mirror.

$H_0(x,y)$ and reconstructs the complex field $U_z(x,y)$ in a plane in the object beam at any distance z from the camera [18, 20]. Among several efficient options [11] is

$$U_z(x,y) = \frac{-\mathrm{i}}{\lambda z} \exp\left(\mathrm{i}\frac{2\pi}{\lambda}z\right) H_0(x,y) \star \exp\left[\mathrm{i}\pi\frac{(x^2+y^2)}{\lambda z}\right],$$
(1)

where λ denotes wavelength and $\star$ denotes convolution. As with conventional holography [10, 2], a windowed subset of the Fresnel field can be used to reconstruct a particular view of the object. As the window explores the field a different angle of view of the object can be reconstructed. The range of viewing angles is determined by the ratio of the window size to the full CCD sensor dimensions. Our CCD sensor has approximate dimensions of 18.5×18.5 mm and so a 1024×1024 pixel window has a maximum lateral shift of 9 mm across the face of the sensor. With an object positioned $d = 350$ mm from the camera, viewing angles in the range of $1.5°$ are permitted. Smaller windows will permit a larger range of viewing angles at the expense of image quality at eachviewpoint.

3. Focus detection in digital holograms

One method for reconstructing a DH at the most in-focus plane is to use a depth-from-focus (DFF) technique. Although there is no definitive criterion for finding the focal

plane of a scene or finding the focal distance for a region within a scene, a number of focus metrics have been proposed and demonstrated [9, 7, 13, 14, 5]. These employ self-entropy [9], phase changes [7], wavelet analysis [13], gray level variance [14], and integrated amplitude modulus [5] among others. Using these metrics, applications such as the detection of the focal plane [7, 13, 5] in digital holographic microscopy [3] and the measurement of 3D objects in digitised physical holograms [14] have been demonstrated. These techniques reconstruct over a range of depths and evaluate each 2D reconstruction using a focus metric, and when the focus metric returns a maximum the corresponding depth is returned. This relies on the assumption that a large majority of the scene is in focus at a single depth. This is not the case if there are multiple objects at different depths, or if the physical object itself extends in the z direction. This leads to a reconstruction containing some object regions in-focus with the rest out-of-focus. These blurred out-of-focus regions make it difficult to segment a single reconstruction of a DH based on texture or intensity. An approach for the recovery of 3D shape information from digitised physical holograms was proposed by Ma et al. [14]. By calculating variance on non-overlapping blocks from reconstructions of a DH at different depths they recovered depth information from a lower-resolution version of the sensed object. We choose to extend this variance-measurement approach in order to classify each 1D vector (x,y) in the reconstruction volume (each line of pixels parallel to the optical axis) as either belonging to the object or belonging to the background. The decision is taken as follows: if vector (x,y) contains an in-focus pixel from the object at any depth z then (x,y) is an object pixel, otherwise it is a background pixel.

4. Object Segmentation in digital holograms

We employ a DFF technique to create a maximum variance map. We then create our segmentation mask through thresholding the maximum variance map with a manually chosen threshold. We numerically reconstruct the digital hologram for each depth in the selected range of depths. Each reconstruction $I_z = |U_z|^2$ is of size $M \times N$ pixels. We first apply a Fourier filtering speckle reduction technique to each reconstruction [15]. We then calculate variance for each pixel by calculating variance on $n \times n$ pixel overlapping blocks approximately centred on each pixel, and address each block with (k,l) where $k \in [0, M-1], l \in [0, N-1]$. Variance of each overlapping block at each depth
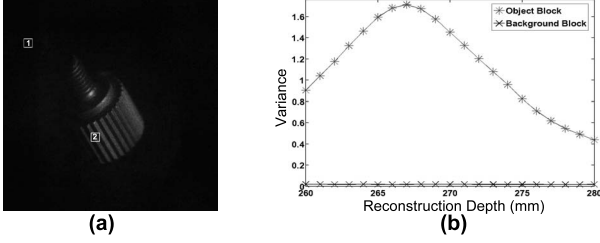
Figure 2. Screw object DH: (a) numerical reconstruction, (b) variance plot for background and object regions.

Figure 3. Segmentation of screw object DH: (a) manual, (b) our DFF approach, (c) EM approach.

z is calculated with function $V_z : R^{n \times n} \to R^+$ defined by

$$V_z(k,l) = \frac{1}{n^2} \sum_{x=k-\lfloor \frac{n-1}{2} \rfloor}^{k+\lceil \frac{n-1}{2} \rceil} \sum_{y=l-\lfloor \frac{n-1}{2} \rfloor}^{l+\lceil \frac{n-1}{2} \rceil} \left[I_z(x,y) - \overline{I_z(k,l)} \right]^2 , \tag{2}$$

where $\overline{I_z(k,l)}$ is defined

$$\overline{I_z(k,l)} = \frac{1}{n^2} \sum_{x=k-\lfloor \frac{n-1}{2} \rfloor}^{k+\lceil \frac{n-1}{2} \rceil} \sum_{y=l-\lfloor \frac{n-1}{2} \rfloor}^{l+\lceil \frac{n-1}{2} \rceil} I_z(x,y), \tag{3}$$

and where any indexes (x,y) that go outside the extent of I_z evaluate to 0. V is therefore a volume storing a 2D variance image for each depth z. A location with high variance indicates the nearby presence of an object. We find the maximum value in each (k,l) vector with

$$V_{\max}(k,l) = \max_z \left[V_z(k,l) \right]. \tag{4}$$

Where the maximum variance is low, this indicates a background region. A threshold τ is chosen and $V_{\max}$ is transformed as

$$\mathrm{SMask}(k,l) = \{\, 1 \,, \text{if } V_{\max}(k,l) \geq \tau 0, \text{if } V_{\max}(k,l) < \tau, \tag{5}$$

where 0 denotes a background pixel and 1 denotes an object pixel. The binary image SMask is our segmentation mask. Finally, we apply a mathematical morphology erosion operation (with neighbourhood $\lceil n/2 \rceil \times \lceil n/2 \rceil$) to SMask to shrink the boundaries of the object; our use of overlapping blocks uniformly enlarges the mask. Erosion has the potential to change the shape of the mask and may remove fine features along the boundary of the mask.

5. Object Segmentation Experiments

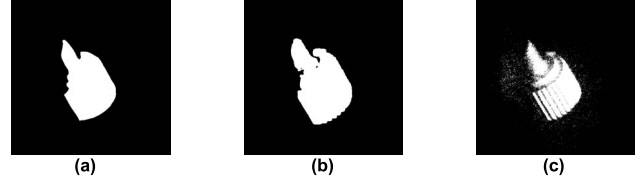We verify our DFF technique using DHs of real-world objects. The first object (screw) is $1\,\mathrm{cm}^3$ and was positioned $268\,\mathrm{mm}$ from the camera. Our second object (knight) is $2\,\mathrm{cm} \times 2\,\mathrm{cm} \times 0.6\,\mathrm{cm}$ and was positioned $371\,\mathrm{mm}$ from the camera. The third object (stormtrooper) is $1.7\,\mathrm{cm} \times 2\,\mathrm{cm} \times 1.1\,\mathrm{cm}$ and positioned $366\,\mathrm{mm}$ from the camera while the fourth object (two screws) is $1.8\,\mathrm{cm} \times 2.3\,\mathrm{cm} \times 1.6\,\mathrm{cm}$ and positioned $360\,\mathrm{mm}$ from the camera. Each DH has 2032×2048 pixels. We compute a sequence of reconstructions at different depths from a single perspective with a uniform interval of $1\,\mathrm{mm}$ between successive values of z. We apply our DFF technique to this sequence of reconstructions to obtain SMask. A reconstruction of the screw object is shown in Fig. 2(a), where two 81×81 pixel blocks labeled 1 and 2 have been manually selected to indicate example background and object regions, respectively. Plots of variance calculated on these blocks over a range of 21 mm are shown in Fig. 2(b). It can be seen that the variance of the background block is 10^{-2} lower than that for the object block for this hologram. We found this to be true in the general case and allowed us to choose the appropriate normalised values of $\tau = 2 \times 10^{-3}$, $\tau = 4 \times 10^{-5}$, $\tau = 1.4 \times 10^{-4}$ and $\tau = 2 \times 10^{-4}$ for the screw, knight, stormtrooper and two screws object respectively. In Fig. 3 we compare our results with the ground-truth case where we manually classified each pixel as either background or object, and with the well-known and robust intensity-based segmentation technique expectation-maximisation (EM) [8]. EM is a two-step iterative segmentation algorithm requiring two inputs: an image histogram h, and the desired number of segments k. This algorithm attempts to fit k Gaussians to the image histogram. The first step estimates the mean, variance and proportion of the k Gaussians for the input image histogram. The second step attempts to maximise these Gaussian parameters for the input data with a maximum likelihood estimate. This is repeated until the difference between successive maximum likelihood estimates is below a threshold. Our Gaussian mixture models are computed with

$$(G)(k) = \frac{p_e(k)}{\sqrt{2\pi \times \sigma_e^2(k)}} \times \exp\left(-\frac{1}{2} \times \frac{(h - \mu_e(k))^2}{\sigma_e^2(k)} \right) \tag{6}$$

43

where $\mu_e(k)$,σ_e^2 and p_e are the estimated mean,variance and proportion respectively. The shallow focal range of the screw object DH allows for a comparison between our method and this 2D technique. We use receiver operating characteristic (ROC) analysis [6] to display the relative tradeoff between the true positive rate and the false negative rate of our classifiers. In this analysis the DFF block size is the variable in our choice of classifier. We estimate the true positive rate of a classifier as

$$\text{tp rate} \simeq \frac{\text{True Positives}}{\text{Total number of positives}} \qquad (7)$$

and the false positive rate as

$$\text{fp rate} \simeq \frac{\text{False Positives}}{\text{Total number of negatives}}. \qquad (8)$$

Classification results are plotted on orthogonal axes defined by tp rate and fp rate, allowing us to choose the block size n that best maximises the tp rate while simultaneously minimising the fp rate. For the screw object DH, we used a set of nine different block sizes, ranging from $n = 7$ to $n = 151$. The ROC curve this set of classifiers produced can be seen in Fig. 4(a). Since the data is clustered in the top left corner of the ROC graph a zoomed in version of the top left corner of the ROC graph is shown in Fig. 4(b). As desired, all points are located far from the random guess classifier performance. It is clear from these graphs that a small block size classifies background pixels perfectly at the expense of object pixels, and a large block size classifies object pixels perfectly at the expense of background pixels. Our compromise between perfect background/object segmentation is to minimise the distance between the points in ROC space and the point (0,1), since we regard false positives and false negatives as being equally undesirable. Through experiments with different objects we chose a block size of 81×81 pixels which has an average tp rate of 98.83% and an average fp rate of 1.04%. Using this block size, we created the segmentation mask shown in Fig. 3(b). By comparison, EM achieved a good fp rate of 1% but a relatively poor tp rate of 80%. The segmentation mask created by EM is shown in Fig. 3(c). As expected the segmentation mask obtained from a volume of variance maps outperforms one created using an individual reconstruction. We also present the results of applying our DFF technique to three other DHs: a low contrast knight object, a low contrast stormtrooper object and a high contrast two screws object. A reconstruction of the knight object, stormtrooper object and two screws object is shown in Fig. 5(a), 6(a) and 7(a) respectively. As part of the speckle reduction measures taken for these DHs, the reconstruction sizes have been reduced to 512×512 pixels. This is one quarter of the resolution of the previous experiments which leads us to choose a block size of 21×21 (one quarter the
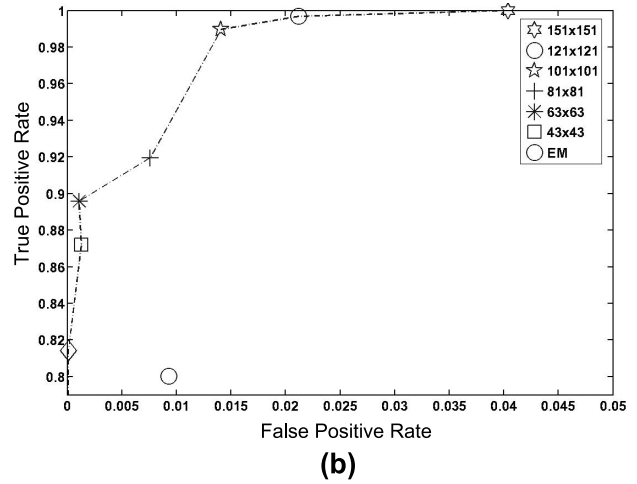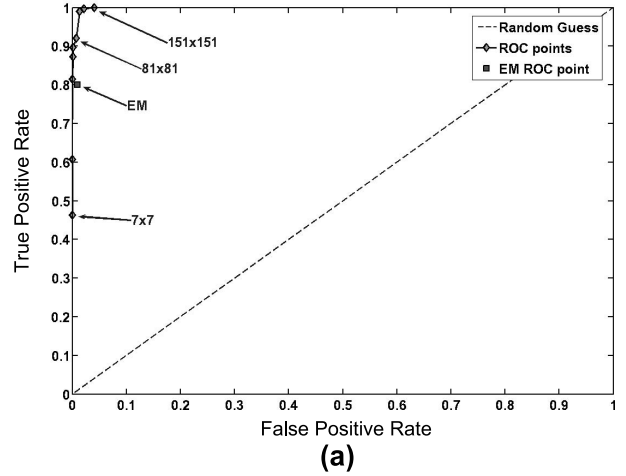

(a)

(b)

Figure 4. ROC graphs for object segmentation using different block sizes, (a) full ROC graph, (b) ROC graph magnified to show the top left corner.
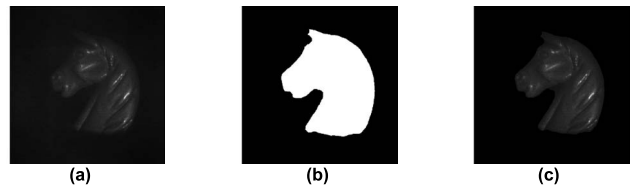

Figure 5. Segmentation of knight object DH: (a) numerical reconstruction, (b) segmentation mask obtained, (c) segmented reconstruction.
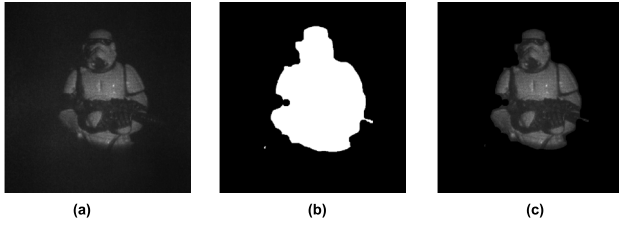
Figure 6. Segmentation of stormtrooper object DH: (a) numerical reconstruction, (b) segmentation mask obtained, (c) segmented reconstruction.
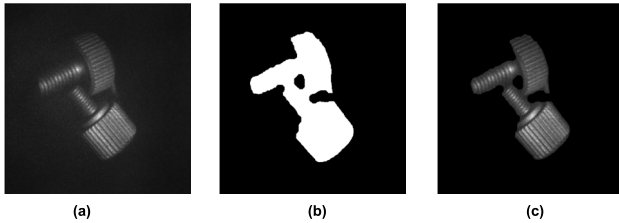

Figure 7. Segmentation of a two screws object DH: (a) numerical reconstruction, (b) segmentation mask obtained, (c) segmented reconstruction.

size of the block size chosen as a result of the ROC analysis). The erosion operator has a neighbourhood of 11×11 (half the block size used to create the segmentation mask). The resulting segmentation masks are shown in Fig. 5(b), 6(b) and 7(b). A segmented reconstruction at a in-focus plane for each of the DHs is shown in part (c) of Figures 5, 6 and 7. These segmented reconstructions are obtained from $I_z(x,y) \cdot \text{SMask}(x,y)$ where $\cdot$ means pointwise product. These results illustrate how objects can be successfully segmented from the background.

6. Conclusions

We have presented an approach for the segmentation of DHs into object and background. We have shown segmentation masks obtained from DHs containing high and low contrast objects. We have also demonstrated that our DFF approach produces a segmentation mask with a higher accuracy than one produced using a single reconstruction and a 2D segmentation algorithm. We expect our single object segmentation method will be successful for all macroscopic objects recorded by digital holography except pure phase objects. The performance of the algorithm for small objects has yet to be ascertained. For microscopic objects it is expected that a phase-unwrapping based approach would

be best. The accuracy of our approach is limited by an appropriate choice of block size and threshold value. Also, this method is currently limited to the segmentation of the scene from its background where that background is not itself composed of other objects. Extensions to our technique are currently planned for multi-object segmentation.

Acknowledgements

References

[1] J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio. Digital wavefront measuring interferometer for testing optical surfaces and lenses. *Applied Optics*, 13:2693–2703, 1974.

[2] H. J. Caulfield. *Handbook of Optical Holography*. Academic Press, New York, 1979.

[3] E. Cuche, F. Bevilacqua, and C. Depeursinge. Digital holography for quantitative phase-contrast imaging. *Optics Letters*, 24:291–293, 1999.

[4] M. DaneshPanah and B. Javidi. Segmentation of 3d holographic images using bivariate jointly distributed region snake. *Optics Express*, 14:5143–5153, 2006.

[5] F. Dubois, C. Schockaert, N. Callens, and C. Yourassowsky. Focus plane detection criteria in digital holography microscopy by amplitude analysis. *Optics Express*, 14:5895–5908, 2006.

[6] T. Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

[7] P. Ferraro, G. Coppola, S. D. Nicola, A. Finizio, and G. Peirattini. Digital holographic microscope with automatic focus tracking by detecting sample displacement in real time. *Optics Letters*, 28:1257–1259, 2003.

[8] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, third edition, 2003.

[9] J. Gillespie and R. King. The use of self-entropy as a focus measure in digital holography. *Pattern Recognition Letters*, 9:19–25, 1989.

[10] J. W. Goodman. *Introduction to Fourier Optics*. Roberts and Company, Englewood, Colorado, third edition, 2005.

[11] B. M. Hennelly and J. T. Sheridan. Fast numerical algorithm for the linear canonical transform. *J. Opt. Soc. Am. A*, 22:928–937, 2005.

[12] T. Kreis. *Handbook of Holographic Interferometry*. Optical and Digital Methods, Wiley-Vch, Weinheim, 2005.

[13] M. Liebling and M. Usner. Autofocus for digital fresnel holograms by use of a Fresnelet-sparsity criterion. *J. Opt. Soc Am A*, 21:2424–2430, 2004.

[14] L. Ma, H. Wang, Y. Li, and H. Jin. Numerical reconstruction of digital holograms for three-dimensional shape measurement. *Journal of Optics A*, 6:396–400, 2004.

[15] J. Maycock, B. M. Hennelly, J. B. McDonald, T. J. Naughton, Y. Frauel, A. Castro, and B. Javidi. Reduction of speckle in digital holography by discrete fourier filtering. *J. Opt. Soc. Am. A*, 24, 2007.

[16] J. Maycock, C. P. McElhinney, B. M. Hennelly, T. J. Naughton, J. B. McDonald, and B. Javidi. Reconstruction of partially occluded objects encoded in three-dimensional scenes by using digital holograms. *Applied Optics*, 45:2975–2985, 2006.

[17] T. J. Naughton, Y. Frauel, B. Javidi, and E. Tajahuerce. Compression of digital holograms for three-dimensional object reconstruction and recognition. *Applied Optics*, 41:4124–4132, 2002.

[18] L. Onural and P. D. Scott. Digital decoding of in-line holograms. *Opt. Eng.*, 26:1124–1132, 1987.

[19] U. Schnars and W. P. O. Jüptner. *Digital Holography: Digital Hologram Recording, Numerical Reconstruction*. Springer, 2004.

[20] I. Yamaguchi and T. Zhang. Phase-shifting digital holography. *Optics Letters*, 22:1268–1270, 1997.