Overview of three-dimensional integral imaging-based object recognition in low illumination conditions with visible range image sensors

Timothy O'Connor¹ · Adam Markman² · Bahram Javidi²

Received: 16 April 2020 / Accepted: 16 September 2020 / Published online: 23 September 2020
© Springer Nature Switzerland AG 2020

Abstract

We overview three-dimensional (3D) integral imaging-based object recognition in low illumination conditions. Imaging in very low illumination conditions, especially using passive, visible-range image sensors, remains a challenging endeavor that is particularly complicated by read-noise dominant images and provides unsatisfactory results when using conventional two-dimensional imaging strategies in photon-starved conditions. However, using passive three-dimensional integral imaging, which is optimal in a maximum likelihood sense, we are able to significantly improve imaging capabilities under low illumination conditions and allow for object detection and recognition. We overview reported work on 3D integral imaging object visualization, and recognition in low light including recent work utilizing convolutional neural networks for object recognition in very low illumination conditions.

Keywords Three-dimensional (3D) imaging · Integral imaging · Low light imaging · Object recognition

1 Introduction

Imaging in low illumination conditions is of interest for applications in remote sensing, underwater imaging, and night vision. Integral imaging is a three-dimensional (3D) imaging technique that incorporates the angular and intensity information from multiple viewing perspectives to reconstruct a 3D scene [1, 2], and has been shown to provide superior performance over 2D imaging strategies in low light environments due to being optimal in a maximum likelihood sense [3–8]. The pickup process which consists of recording images with different viewing perspectives, known as elemental images, can be accomplished using a lenslet or camera array, or by using a single camera on a moving translation stage. The 3D reconstruction of the scene can be performed either optically or computationally. Computational 3D integral imaging reconstruction is performed using the following equation:

$$I(x, y; z) = \frac{1}{O(x, y)} \sum_{a=0}^{A-1} \sum_{b=0}^{B-1} E_{a,b} \left(x - a \frac{L_x \times p_x}{c_x \times M}, y - b \frac{L_y \times p_y}{c_y \times M} \right) + \varepsilon,$$

where (x, y) is the pixel index, z is the reconstruction distance, $O(x, y)$ is the overlapping number on (x, y) , A and B are the total number of elemental images obtained in each column and row, respectively; $E_{a,b}$ is the elemental image in the a -th column and b -th row, and L_x and L_y are the total number of pixels in each column and row, respectively, for each $E_{a,b}$. M is the magnification factor and equals to z/f , f is the focal length, p_x and p_y represent the pitch between image sensors, c_x and c_y are the size of the image sensor, and ε is the additive read noise. Figure 1 shows the diagram for optical pickup and computational 3D reconstruction of integral imaging [3].

✉ Bahram Javidi, Bahram.Javidi@UConn.edu | ¹Biomedical Engineering Department, University of Connecticut, Storrs, CT 06269-3247, USA. ²Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT 06269-4157, USA.

Fig. 1 Diagram of integral imaging, **a** optical pickup of elemental images, and **b** 3D computational reconstruction. p is the pitch between sensors, g is the focal length, and c is the sensor size [3]

Fig. 2 Experimental results for two low light illumination conditions. **a, d** are the read noise limited 2D elemental images. **b, e** are the reconstructed 3D images with the faces detected. **c, f** are the detected faces from **(b, e)** respectively, after total variation denoising [3]

2 Results and discussion

In [3], 3D integral imaging-based low illumination object visualization and detection was presented using a conventional low-cost and compact CMOS sensor on a moving translation stage to capture the perspective elemental images of the scene. 36 elemental images were recorded of a person standing behind an occluding tree branch in low illumination conditions. Following image acquisition, computational 3D integral imaging reconstruction was performed, then the Total Variation (TV) denoising algorithm [9] was applied to the 3D reconstructed image. After denoising, the Viola-Jones object detection framework [10] was used on the reconstructed image for successful face detection. Experimental results under two low illumination conditions are provided by Fig. 2. Analysis of these results showed a reduction in entropy for the 3D reconstructed images, as well as an increase in the signal-to-noise ratio (SNR) in comparison to the traditional 2D imaging [3]. This overviewed work

[3] demonstrated 3D integral imaging for object visualization and detection in poor illumination conditions without the need for photon-counting or cooled CCD cameras and enabled detection of faces that was not possible in the conventional 2D images.

More recently, the use of convolutional neural networks (CNN) has been presented for 3D integral imaging-based object recognition in very low illumination conditions [5]. In this overviewed work, 3D integral imaging is used to improve the SNR followed by TV denoising. After TV denoising, regions of interest are extracted from the denoised reconstructed image using the Viola-Jones face detection framework [10] and input into a CNN. Moreover, the input to the CNN is a 2D slice of the reconstructed volume, with several different depths selected separately from each scene. The CNN is trained on different low illumination conditions, then performs object recognition on the 3D reconstructed images taken at an unknown illumination condition [5]. During the experiments, 72 elemental images were obtained using an Allied Vision Mako-192 camera on a translation stage in varying illumination conditions. 6 human subjects were used in the experiments and were located 4.5 m away from the camera array. The scene illumination was altered by adjusting the intensity of the light source. Data was collected at 17 different illumination conditions for each of the 6 subjects. The images were reconstructed at different depths between 4 and 5 m using a step size of 50 mm. From each of the reconstructed images, a region of interest was extracted using the Viola-Jones face detector to be input into the network. The data was then randomly split into testing and training data with 4 randomly chosen illumination conditions being held out of the training procedure for testing and the remaining 13 illumination conditions used for training the CNN. To increase the size of the training set, data augmentation was applied on the extracted regions of interest. After data augmentation, a total of 29,232 images were used for training the network. The overview of the classification scheme using CNN is depicted by Fig. 3. Using this scheme, 100% classification accuracy was achieved for object recognition among the 6 subjects in very low illumination conditions.

3 Conclusions

In summary, we have overviewed recent works [3, 5] for integral imaging-based 3D object detection and recognition in low illumination conditions. 3D integral imaging improves the SNR over the conventional 2D images in photon-starved conditions. Following 3D reconstruction, TV denoising further improves the image quality, then faces can be detected using the Viola-Jones face detector

Fig. 3 Overview for classification procedure using CNN for low light object recognition [5]

which fails on the read-noise dominant conventional 2D images. The detected faces can be recognized using a CNN for classification [5]. Continued research for integral imaging-based 3D object detection and recognition in low illumination conditions includes work with highly sensitive imaging sensors such as the scientific CMOS and electron multiplying CCD cameras, and work in low light polarimetric imaging [6–8].

Acknowledgement B. Javidi acknowledges support by Air Force Office of Scientific Research (FA9550-18-1-0338); and Office of Naval Research (N000141712405, N00014-17-1-2561); Night Vision and Electronic Sensors Directorate, Communications-Electronics Research, Development and Engineering Center, US Army (W909MY-12-D-0008). T. O'Connor acknowledges support by the Dept. of Education through the GAANN program.

Funding Air Force Office of Scientific Research (FA9550-18-1-0338); Office of Naval Research (ONR) (N000141712561, N000141712405, N000142012690); Night Vision and Electronic Sensors Directorate, Communications-Electronics Research, Development and Engineering Center, US Army (W909MY-12-D-0008).

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Lippmann G (1908) Épreuves réversibles donnant la sensation du relief. *J Phys Theor Appl* 7:821–825
2. Stern A, Javidi B (2006) Three-dimensional image sensing, visualization, and processing using integral imaging. *Proc IEEE* 94:591–607
3. Markman A, Shen X, Javidi B (2017) Three-dimensional object visualization and detection in low light illumination using integral imaging. *Opt Lett* 42:3068–3071
4. Stern A, Aloni D, Javidi B (2012) Experiments with three-dimensional integral imaging under low light levels. *IEEE Photon J* 4:1188–1195
5. Markman A, Javidi B (2018) Learning in the dark: 3D integral imaging object recognition in very low illumination conditions using convolutional neural networks. *OSA Continuum* 1:373–383
6. Shen X, Carnicer A, Javidi B (2019) Three-dimensional polarimetric integral imaging under low illumination conditions. *Opt Lett* 44:3230–3233
7. Markman A, O'Connor T, Hotaka H, Ohsuka S, Javidi B (2019) Three-dimensional integral imaging in photon-starved environments with high-sensitivity image sensors. *Opt Express* 27:26355–26368
8. Hotaka H, O'Connor T, Ohsuka S, Javidi B (2020) Photon-counting 3D integral imaging with less than a single photon per pixel on average using a statistical model of the EM-CCD camera. *Opt Lett* 1:1. <https://doi.org/10.1364/OL.389776>
9. Chan SH, Khoshabeh R, Gibson KB, Gill PE, Nguyen TQ (2011) An augmented Lagrangian method for total variation video restoration. *IEEE Trans Image Process* 20:3097–3111
10. Viola P, Jones M, Snow D (2005) Detecting pedestrians using patterns of motion and appearance. *Int J Comput Vis* 63:153–161

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.