Single-pixel 3D reconstruction via a high-speed LED array

Ming Wang\textsuperscript{1,2}, Ming-jie Sun\textsuperscript{2} and Chao Huang\textsuperscript{2}

\textsuperscript{1} Research Institute for Frontier Science, Beihang University, Beijing, People’s Republic of China
\textsuperscript{2} School of Instrumentation Science and Opto-Electronic Engineering, Beihang University, Beijing, People’s Republic of China

E-mail: mingjie.sun@buaa.edu.cn

Keywords: 3D reconstruction, single-pixel imaging, LED array

Abstract

Three-dimensional reconstruction can be performed in many ways, among which photometric stereo is an established and intensively investigated method. In photometric stereo, geometric alignment or pixel-matching between two-dimensional images under different illuminations is crucial to the accuracy of three-dimensional reconstruction, and the dynamic of the scene makes the task difficult. In this work, we propose a single-pixel three-dimensional reconstruction system utilizing structured illumination, which is implemented via a high-speed LED array. By performing 500 kHz structured illumination and capturing the reflected light intensity with detectors at different spatial locations, two-dimensional images of different shadows with 64 × 64 pixel resolution are reconstructed at 122 frame per second. Three-dimensional profiles of the scene are further reconstructed using the surface gradients derived by photometric stereo algorithm, achieving a minimum accuracy of 0.50 mm. Chromatic three-dimensional imaging via an RGB LED array is also performed at 40 frame per second. The demonstrated system significantly improves the dynamic performance of the single-pixel three-dimensional reconstruction system, and offers potential solutions to many applications, such as fast three-dimensional inspection.

1. Introduction

Three-dimensional (3D) reconstruction is an intensively explored technique, which is applied to areas such as public security, robotics, medical sciences, and military defense \cite{1}. Due to the high integration, charge coupled device and complementary metal-oxide semiconductor are usually used in 3D reconstruction systems \cite{2,3}. However, light sensors with a pixelated structure are unable to satisfy all application requirements, particularly in low light environment and in spectral ranges out of the visible region \cite{4-9}. Alternatively, images can be reconstructed by a single-pixel detector with computational algorithm \cite{10,11}. Therefore, single-pixel imaging has attracted much attention in 3D imaging in recent years \cite{12,13}. A variety of approaches have been proposed for different applications, among which time-of-flight \cite{12} and stereo vision \cite{13,14} are commonly used.

Time-of-flight measurement determines the distance to a scene by illuminating it with pulsed light and comparing the arrival time of the back-scattered light to the emitting time of the laser pulse. A major advantage of time-of-flight based 3D imaging is that its depth resolution is mainly determined by the pulse width of the pulsed light, and not dramatically affected by increases the distance between the system and the object, therefore a good candidate for long-distance 3D measurement \cite{15}. A narrower pulse width means a smaller uncertainty in time-of-flight measurement and less overlapping between back-scattered signals from objects of different depths, which in turn improves the system depth resolution. Single-pixel imaging based time-of-flight system uses pulsed laser for structured illumination and a time-resolving detector for detection \cite{16}. A series of two-dimensional (2D) images can be obtained at different depths, forming an image cube in 3D. However, the method is computational, because images at all different depths need to be reconstructed \cite{17}.

Stereo vision uses two or more images obtained simultaneously from different viewpoints to reconstruct a 3D image, while the geometric registration between several images may be problematic. Among stereo
vision methods, photometric stereo captures images with a fixed viewpoint but different illuminations [18]. The pixel correspondence is easier to perform than in traditional stereo vision, but images of different illumination have to be taken sequentially. If the images are captured at the same time, the imaging rate of photometric stereo will be improved to be much faster. In structured light illumination based single-pixel 3D imaging, several detectors measure the back-scattered light intensities form different locations simultaneously, solving the problem of shooting sequence [13, 19]. The 2D images reconstructed by different detectors appears to be illuminated from different directions. Thus, this architecture simplifies the 3D imaging system to one spatial light modulator, one lens, and several single-pixel detectors without compromising the quality of the reconstruction.

In single-pixel imaging, the number of mask patterns required for one image reconstruction is proportional to the pixel resolution of the image, therefore limiting the frame rate of the system. One approach is to decrease the required number via compressive sensing algorithms of different sophistications [20–26]. However, these algorithms either have a computational overhead or are incapable of reaching a significant low compressive ratio. An alternative and more fundamental approach will be to increase the rate of structured illumination, reaching 64 × 64 pixel resolution at 8 frame per second with a 22 KHz digital-micromirror-device [19], or 32 × 32 pixel resolution at 10 fps with a 10 KHz LED array [27]. A 32 × 32 resolution LED array with a 500 KHz illumination rate was proposed in our previous work [28], achieving 1000 fps single-pixel imaging at a 25% compressive rate. However, the poor spatial resolution of the LED array in previous works limited its potential in various applications.

In this paper, we demonstrate a single-pixel 3D reconstruction system with a high-speed LED array, which has a pixel resolution of 64 × 64 and a maximum illumination rate of 2.5 MHz. 3D profiling with 0.50 mm accuracy at 122 fps are performed, chromatic 3D reconstruction results at 40 fps are presented as well.

### 2. 3D reconstruction principle

The architecture of 3D single-pixel imaging in our experiment is shown in figure 1. The mask pattern \(P_i\) displayed by an LED array is projected onto the object \(O\) through a projection lens, and then a single-pixel detector measures the total intensity of the reflected light as the inner product of \(P_i\) and \(O\) as

\[
s_i = P_i \cdot O,
\]

where \(O = [o_1, o_2, \ldots, o_N]^T\) denotes the reflectivity of the object at corresponding spatial location, and \(P_i = [p_{i1}, p_{i2}, \ldots, p_{iN}]\) denotes the mask values, either binary or greyscales at each spatial location. With \(M\) measurements, a linear equation set can be formed as

\[
S = P \cdot O,
\]

where \(S = [s_1, s_2, \ldots, s_M]^T\) denotes \(M\) measured values, and \(P = [P_1, P_2, \ldots, P_M]^T\) denotes \(M \times N\) measuring mask matrix. If \(M = N\) and \(P\) is orthogonal, the image of the object can be reconstructed by

\[
O = P^{-1} \cdot S.
\]

The 3D profile of the object can be reconstructed using the 2D images obtained from single-pixel detectors of different spatial location, i.e. Detector 1, 2, 3, and 4 in figure 1. The intensity of a pixel location \((i, j)\) in the image obtained by the \(k\)th detector can be denoted as

\[
E_k = E_R \left( \hat{d}_k \cdot \hat{n}_{ij} \right),
\]

where \(E_R\) is the reflected intensity of the light source, \(\hat{d}_k\) is the unit detector vector from the object to the detector, and \(\hat{n}_{ij}\) is the surface normal unit vector of the object at a given pixel location \((i, j)\). In the case that four detectors at different spatial locations are used to yield images from different perspectives, equation (4) can be rewritten as

\[
E = E_R \left( D \cdot \hat{n}_{ij} \right),
\]

where \(D = [\hat{d}_1, \hat{d}_2, \hat{d}_3, \hat{d}_4]\) is the unit detector vectors matrix, and \(E = [E_1, E_2, E_3, E_4]\) is the pixel intensity array acquired by four detectors. The unit surface normal vector of the corresponding pixel is determined as

\[
\hat{n}_{ij} = \frac{1}{E_R} \cdot D^{-1} \cdot E.
\]
By normalizing the surface normal vector in z direction, the surface normal vector at the pixel location \((i, j)\) can be rewritten as \(n_{ij} = \left(-p_{ij}, -q_{ij}, 1\right)^T\), where \(p_{ij}\) and \(q_{ij}\) are the gradients of the surface in \(x\) and \(y\) directions at the corresponding pixel location \((i, j)\), respectively.

Denote \(z_{ij}\) as the height of the surface at a pixel location \((i, j)\) and \(\{z\}\) as the set of the height for all pixel locations. The relationship between a pixel \((i, j)\) and its adjacent four pixels can be expressed by the Taylor series expansion as

\[
\begin{align*}
z_{i+1,j} &= z_{i,j} + \frac{\partial z}{\partial x} + \frac{1}{2!} \frac{\partial^2 z}{\partial x^2} + \frac{1}{3!} \frac{\partial^3 z}{\partial x^3} + \cdots \\
\frac{\partial z}{\partial x} &= z_{i-1,j} - \frac{\partial z}{\partial x} + \frac{1}{2!} \frac{\partial^2 z}{\partial x^2} - \frac{1}{3!} \frac{\partial^3 z}{\partial x^3} + \cdots \\
z_{i,j+1} &= z_{i,j} + \frac{\partial z}{\partial y} + \frac{1}{2!} \frac{\partial^2 z}{\partial y^2} + \frac{1}{3!} \frac{\partial^3 z}{\partial y^3} + \cdots \\
\frac{\partial z}{\partial y} &= z_{i,j-1} - \frac{\partial z}{\partial y} + \frac{1}{2!} \frac{\partial^2 z}{\partial y^2} - \frac{1}{3!} \frac{\partial^3 z}{\partial y^3} + \cdots
\end{align*}
\]

Adding the equations and ignoring any terms after the 3rd order, equation (7) can be rearranged as

\[
4z_{ij} = (z_{i+1,j} + z_{i-1,j} + z_{i,j+1} + z_{i,j-1}) = -\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right).
\]

In order to estimate the accurate surface height, a common reconstruction approach [29] is to minimize the least square error function of

\[
Erf(z) = \iint \left(\left(\frac{\partial z}{\partial x} - p_{ij}\right)^2 + \left(\frac{\partial z}{\partial y} - q_{ij}\right)^2\right) dxdy,
\]

where the minimum will be yield by Euler–Lagrange equation

\[
\nabla^2 z = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = \frac{dp_{ij}}{dx} + \frac{dq_{ij}}{dy}.
\]

Consequently, equation (8) becomes

\[
4z_{ij} = (z_{i+1,j} + z_{i-1,j} + z_{i,j+1} + z_{i,j-1}) = -\left(\frac{dp_{ij}}{dx} + \frac{dq_{ij}}{dy}\right),
\]

and by solving a series of equation (11) determined by all pixels in the obtained images, and using Robin boundary conditions [29] which consider the heights of adjacent pixels inside and outside the boundary are equal, each \(z_{ij}\) in the surface height \(\{z\}\) can be calculated.

3. Experiment results

In the experiment, the LED array has 64 × 64 pixels resolution with the size of 88 mm × 88 mm. The LED array can display binary patterns with red, green and blue at a maximum rate of 2.5 MHz, which is a significantly improved device of our previous work [28]. The improvements are achieved by choosing a chromatic LED chip with a small size (SMD18-038BT, chip size 1.3 mm). Hadamard matrix is used for the
mask patterns, which had a small computational overhead [26]. To display Hadamard patterns efficiently, the same two consecutive display strategy from [28] is employed here, with 128 I/O ports. A lens (f = 150 mm) is located at 515 mm far away from the LED array, and projected the light of patterns onto the object, which is 210 mm behind the lens. Photomultiplier tubes (PMTs, Thorlabs PMT2102) are placed up, bottom, left and right directions of the object, and 100 mm away. A high dynamic range digitizer (PicoScope 6404D) acquires and transfers the intensity data to the computer. Differential measurement is performed, i.e. 8192 patterns are displayed, including 4096 Hadamard patterns and their inverse. We used 50% compressive sensing, which has been proved to acquire images without significant signal-to-noise ratio (SNR) degrading [5, 16]. A ladder-shaped object with the size of 22 mm × 22 mm × 8 mm was used for 3D reconstruction as shown in figure 2(a). On the top of the surface, there was a small square with the size of 6 mm × 6 mm. It is worth mentioning that the PMTs used here have a much larger bandwidth (80 MHz) than the 2.5 MHz display rate of the LED array, which is mainly determined by the overall response time of the device, i.e. the switch time of LED chip combined the response time of the electronics on the driving circuit. PMT noise also plays an important role in the experiments. That is, when collecting reflected light intensity, due to the overall response time mentioned above, along with the rise/fall time of the PMTs, only a fractional portion of the acquired signals is valid for image reconstruction. The higher display rate applied on the LED array, the fewer valid samples will be acquired for each measurement, and PMT noise will dominate the measurement and degrade the image quality. Here the display rate of Hadamard patterns is set to 500 KHz, which is an optimal balance between system frame rate and reconstruction quality.

2D images obtained from four directions are shown in figure 2(b). With 50% compressive sensing, the time spent on once acquirement are 8.2 ms, corresponding to 122 fps for 2D imaging. Figures 2(c) and (d) show the 3D image reconstructed by 2D images, and the error map between the depth of 3D reconstruction and the ground truth. The accuracy of the 3D reconstruction, which is determined as the root mean squared error (RMSE) of the error map, is 0.88 mm. The error map shows that the inaccuracy of the reconstruction mainly happens near the edges of the object, which is caused by the fact that the photometric stereo algorithm used in the experiment is based on finite difference principle, and the edges would be reconstructed over smoothed.

To verify the dynamic performance of the proposed system, another experiment of imaging objects in motion is performed. Trapezoid, semicircle and triangle shaped prisms are evenly placed on a rotating disk with an interval of 60°, as shown in figure 3(a). The disk rotates at the speed of 4 rounds per second, and the objects, placed 33 mm to the center of the disk, move at 8.3 m s⁻¹ accordingly. The parameter configurations of the system are exact the same as the previous experiment, leading to 122 fps 3D reconstruction. Figure 3(b) shows the reconstructed 3D profiles of the trapezoid, semicircle and triangle objects. Figure 3(c) demonstrate the error maps between the reconstructed 3D profiles and their corresponding ground truth, and the RMSEs are 0.50 mm, 0.51 mm and 0.77 mm, respectively. The dynamic results demonstrate a profiling accuracy similar to that of the static one, though being slightly smaller due to the simpler structures.
Figure 3. 3D reconstruction of dynamic objects. (a) The diagram of trapezoid, semicircle and triangle objects, and their location on the rotating disk. (b) Reconstructed 3D profiles of the objects. (c) Error maps between the reconstructions and their corresponding ground truths in depth, the RMSEs are 0.50 mm, 0.51 mm and 0.77 mm, respectively.

Figure 4. Chromatic 3D imaging of a colored toy. (a) The colored toy object. (b) Reconstructed chromatic 3D image. (c) Chromatic 2D images obtained by PMTs of different locations.

of the objects. The maximum errors of the reconstructions happen near the edges of the object, where the gradient of the surface changes abruptly.

The LED chips on the array are chromatic, that is, each LED chip contains one red, one green and one blue light emitting diode. Therefore, by displaying the set of Hadamard mask patterns in red, green and blue sequentially, the proposed system is capable of chromatic imaging. An experiment is performed using a colored toy, shown in figure 4(a), as an object. Same acquisition configuration as previous experiments is implemented. Figures 4(b) and (c) show the reconstructed 3D image of the toy, and the 2D images used for the reconstruction. It is worth mentioning that the 3D profile of the toy is reconstructed at the frame rate of 122 fps, but the chromatic 2D images in figure 4(c) are obtained at 40 fps because they require three consecutive images, one red, one green and one blue, to recover chromatic information.

4. Conclusion

In this work, a 3D single-pixel imaging is presented with a high-speed LED array with a maximum 2.5 MHz refreshing rate. Dynamic experiments demonstrate that the proposed system is capable of 3D reconstruction at the frame rate of 122 fps with an accuracy of 0.50 mm, which is a significant improvement over the existing single-pixel imaging based works. The using of RGB LED chips on the array also enables chromatic 3D imaging at 40 fps. Nevertheless, the accuracy of the system could be further improved by an optimized photometric stereo algorithm, an LED array with more pixels or a higher pixel density, and higher performance detectors. The demonstrated system significantly improves the dynamic performance of the single-pixel 3D reconstruction system, and offers potential solutions to many applications, such as fast 3D inspection.
Funding

National Natural Science Foundation of China (61922011, 61675016).

Conflict of interest

The authors declare no conflicts of interest.

ORCID iDs

Ming Wang https://orcid.org/0000-0002-9532-3616
Ming-jie Sun https://orcid.org/0000-0003-0649-522X

References

[1] Bosch T, Lesoure M, Myllyla R, Rioux M and Amann M C 2001 Laser ranging: a critical review of usual techniques for distance measurement Opt. Eng. 40 10–19
[2] Janesick J and Putnam G 2003 Developments and applications of high-performance CCD and CMOS imaging arrays Annu. Rev. Nucl. Part. Sci. 53 263–300
[3] Fossum E R and Hondongwa D B 2014 A review of the pinned photodiode for CCD and CMOS image sensors IEEE J. Electron Dev. Soc. 2 33–43
[4] Edgar M P, Gibson G M, Bowman R W, Sun B, Radwell N, Mitchell K J, Welsh S S and Padgett M J 2015 Simultaneous real-time visible and infrared video with single-pixel detectors Sci. Rep. 5 10669
[5] Radwell N, Mitchell K J, Gibson G M, Edgar M P, Bowman R and Padgett M J 2014 Single-pixel infrared and visible microscope Optica 1 285–9
[6] Thibault P, Dierolf M, Menzel A, Bunk O, David C and Pfeiffer F 2008 High-resolution scanning x-ray diffraction microscopy Science 321 379–82
[7] Padilla J 2014 Terahertz compressive imaging with metamaterial spatial light modulators Nat. Photon. 8 605–9
[8] Stantchev R I, Sun B Q, Hornett S M, Hobson P A, Gibson G M, Padgett M J and Hendry E 2016 Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector Sci. Adv. 2 e1600190
[9] Withawat W and Derek A 2014 Terahertz imaging: compressing onto a single pixel Nat. Photon. 8 593–4
[10] Shao J H 2008 Computational ghost imaging Phys. Rev. A 78 061802
[11] Sun M J and Zhang J M 2019 Single-pixel imaging and its application in three-dimensional reconstruction: a brief review Sensors 19 732
[12] McCarthy A, Collins R J, Krichel N J, Fernandez V, Wallace A M and Buller G S 2009 Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting Appl. Opt. 48 6241–51
[13] Sun B, Edgar M P and Bowman R 2013 3D computational imaging with single-pixel detectors Science 340 844–7
[14] Boyd A 1985 Stereoscopic images in confocal (tandem scanning) microscopy Science 230 1270–2
[15] Schwarz B 2010 Lidar: mapping the world in 3D Nat. Photon. 4 429–30
[16] Sun M J, Edgar M P, Gibson G M, Sun B, Radwell N and Lamb R 2016 Single-pixel three-dimensional imaging with time-based depth resolution Nat. Commun. 7 12010
[17] McCarthy A, Krichel N J, Gemmell N R, Ren X, Tanner M G and Dorenbos S N 2013 Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detector Opt. Express 21 8904–15
[18] Woodham R J 1980 Photometric method for determining surface orientation from multiple images Opt. Eng. 19 139–44
[19] Zhang Y W, Edgar M P, Sun B Q, Radwell N, Gibson G M and Padgett M J 2016 3D single-pixel video J. Opt. 18 035203
[20] Magalhaes F, Araujo F M, Correia M W, Abolbashari M and Farahi F 2011 Active illumination single-pixel camera based on compressive sensing Appl. Opt. 50 405–14
[21] Baraniuk R G 2007 Compressive sensing IEEE Signal Proc. Mag. 24 118–21
[22] Assmann M and Bayer M 2013 Compressive adaptive computational ghost imaging Sci. Rep. 7 16545
[23] Zhang Z B, Liu S J and Peng J Z 2018 Simultaneous spatial, spectral, and 3D compressive imaging via efficient Fourier single-pixel measurements Optica 5 315–19
[24] Jia T, Chen D Y and Wang J 2018 Single-pixel color imaging method with a compressive sensing measurement matrix Appl. Sci. Based. 8 1293
[25] Qian Y, He R Q and Chen Q 2019 Adaptive compressed ghost imaging based on the variation of surface normal Opt. Express 27 27862–72
[26] Sun M J, Meng L T and Edgar M P 2017 A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging Sci. Rep. 7 3464
[27] Eva S B, Pedro L C and Carlos C 2018 Low-cost single-pixel 3D imaging by using an LED array Opt. Express 26 15623–31
[28] Xu Z H, Chen W, Penuelas J, Padgett M and Sun M J 2018 1000 fps computational ghost imaging using LED-based structured illumination Opt. Express 26 2427–34
[29] Horn B K P 1986 Robot Vision (Cambridge, MA: MIT Press) pp 259–67