# 3D single-pixel video

# 3D single-pixel video

**Yiwei Zhang, Matthew P Edgar, Baoqing Sun, Neal Radwell, Graham M Gibson and Miles J Padgett**

SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, UK

E-mail: ewayzhang@gmail.com

CrossMark

## Abstract

Photometric stereo is an established three-dimensional (3D) imaging technique for estimating surface shape and reflectivity using multiple images of a scene taken from the same viewpoint but subject to different illumination directions. Importantly, this technique requires the scene to remain static during image acquisition otherwise pixel-matching errors can introduce significant errors in the reconstructed image. In this work, we demonstrate a modified photometric stereo system with perfect pixel-registration, capable of reconstructing 3D images of scenes exhibiting dynamic behavior in real-time. Performing high-speed structured illumination of a scene and sensing the reflected light with four spatially-separated, single-pixel detectors, our system reconstructs continuous real-time 3D video at ~8 frames per second for image resolutions of $64 \times 64$ pixels. Moreover, since this approach does not use a pixelated camera sensor, it can be readily extended to other wavelengths, such as the infrared, where camera technology is expensive.

S Online supplementary data available from stacks.iop.org/JOPT/18/035203/mmedia

Keywords: real-time 3D imaging, single-pixel imaging, photometric stereo, Hadamard matrices, pixel registration

(Some figures may appear in colour only in the online journal)

## Introduction

Three-dimensional (3D) imaging is a heavily explored research field that supports a wide range of applications such as object and face recognition, robot navigation, surface mapping and medical operations. A variety of different techniques have been developed, each with different advantages and drawbacks dependent on the specific application [1–9]. Stereo imaging [10–14] is perhaps the most well-known technique, which uses multiple images obtained simultaneously from different viewpoints to reconstruct a 3D scene. However, the associated image processing, in particular the arduous step of performing pixel correspondence, can be problematic and computationally intensive. In contrast, photometric stereo, first introduced by Woodham [15], uses a single viewpoint and multiple lighting directions. This method reconstructs 3D images by defining the surface normals according to measured intensity differences between images taken with the different incident lighting directions [16–20]. Nevertheless, this photometric approach demands that the scene remains completely static whilst the lighting condition changes in order to prevent surface reconstruction errors, which limits its scope in real-time applications. Whilst various methods [21–24] have been proposed for improving the accuracy of different 3D shape recovery algorithms, it seems there has been relatively less work on eliminating the underlying problems associated with sequential acquisitions.

One state-of-the-art approach for solving pixel-matching errors is spectrally multiplexed photometric stereo, where a scene is photographed with a camera system configured to measure multiple spectral channels [25]. This approach uses two cameras aligned co-axially with a beam splitter and spectrally filtered using two different bespoke dichroic filters, in conjunction with three spatially-separated, white-light sources with unique spectral profiles. This approach captures per-pixel photometric normals and full color reflectance simultaneously, and requires no time-varying illumination.
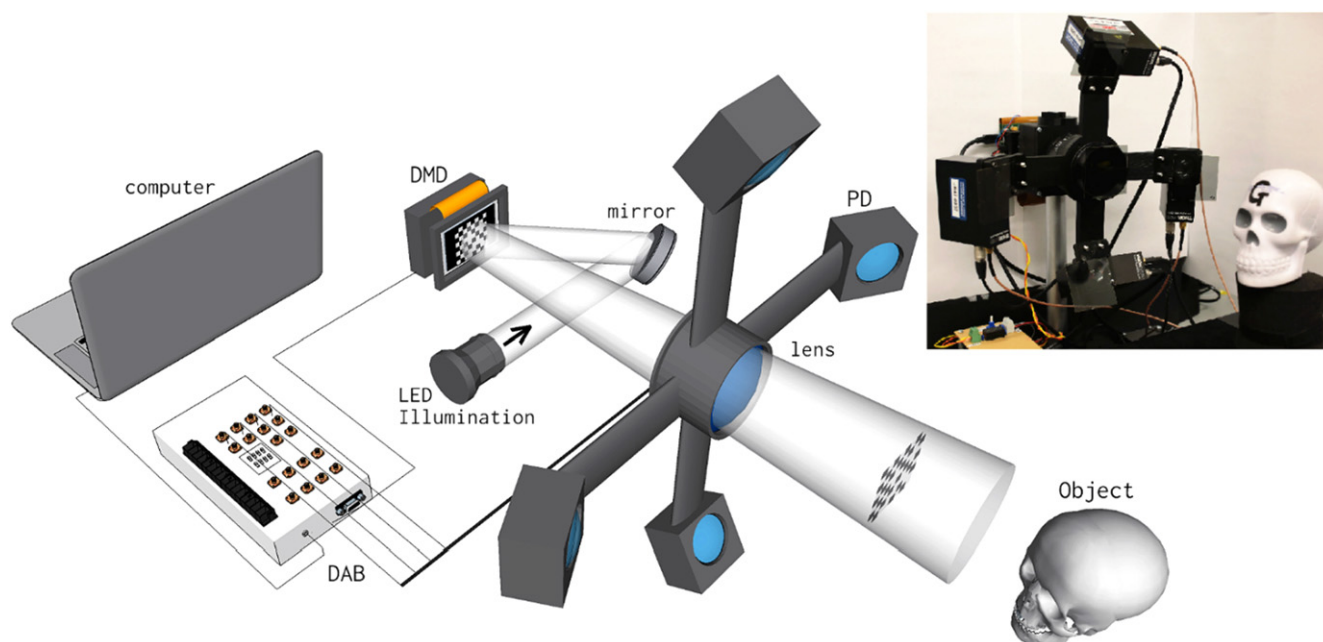
**Figure 1.** Video-rate 3D imaging system. The system contains a digital micro-mirror device (DMD), a camera lens and four spatially separated photodetectors (PD) fixed surrounding it. A white LED light source is used to illuminate the DMD chip and encoded into binary light fields (0 s and 1 s). A mirror is adjusted manually inside a central 3D-printed mount so that the LED light is reflected to the right direction on the DMD chip. The structured light patterns are projected through the lens onto the object. A plastic polarizer sheet is attached in front of each photodetector (horizontally) and the camera lens (vertically) to eliminate the specular reflection on the object. Both photodetectors and the LED light are controlled by a custom electric board (CEB). Each photodetector receives light scatted by the object to give a signal of intensity value which is then sent to the computer through a data acquisition board (DAB) to form four 2D images. These 2D images are analyzed using a photometric technique to give the 3D information of the scene.

However, reconstruction bias still occurs due to spectral variations for scenes with distinct materials, such as human faces.

Here we demonstrate an alternative approach, combining photometric stereo with single-pixel imaging, which utilizes an efficient real-time sampling scheme. Single-pixel imaging [26–28] is a computational imaging technique that allows a single-pixel detector to be used as an imaging device by using a spatial light modulator to provide either time-varying, structured detection of an image or by providing time-varying, structured illumination onto a scene. We have previously shown that when using structured illumination it is the position of the detector determines the apparent lighting condition for the reconstructed image [1]. By using a small number of single-pixel detectors in different spatial locations, multiple images of a scene with different shading profiles can be reconstructed with perfect pixel registration, even for moving objects. By utilizing crossed polarizers it is possible to observe the Lambertian surface reflectivity which allows the surface normals to be estimated via photometric stereo techniques and hence the recovery of 3D images. However, in our original work, many thousands of projected patterns were required leading to acquisition times in the order of several minutes per single image [29, 30].

More generally, within the field of single-pixel imaging, orthogonal and also pseudo-random bases have been employed which significantly reduces the acquisition time, however, the finite modulation rates of micro-electomechanical-systems

technology, typically ∼20 kHz, places restrictions on the achievable frame rates even for relatively low resolution images. A few studies [31–35] have aimed to improve the imaging speed by using compressive sensing which utilize 'a priori' knowledge of the scene, such as sparsity in the spatial frequency domain. Some of the most impressive results utilizing highly-compressed data demand intensive computational processing to recover an image, which does not lend itself well for applications that demand video-rate operation. A variety of alternative compressed sensing schemes have been developed in recent years to enable compressed sensing for large image resolutions. In this work we employ one of these compressive strategies, known as evolutionary compressed sensing [36], in order to demonstrate continuous real-time 3D video at ∼8 frames per second for image resolutions of $64 \times 64$ pixels, equivalent to a speed-up of 4 times compared to a conventional raster-scanning sampling strategy.

## Methods

### Custom single-pixel system design

The application programming interface is written as a dynamic-link library file which provides a convenient interference between of control software and accessory light modulator package (ALP) driver. Patterns are first loaded from the controlling software to the ALP board RAM in
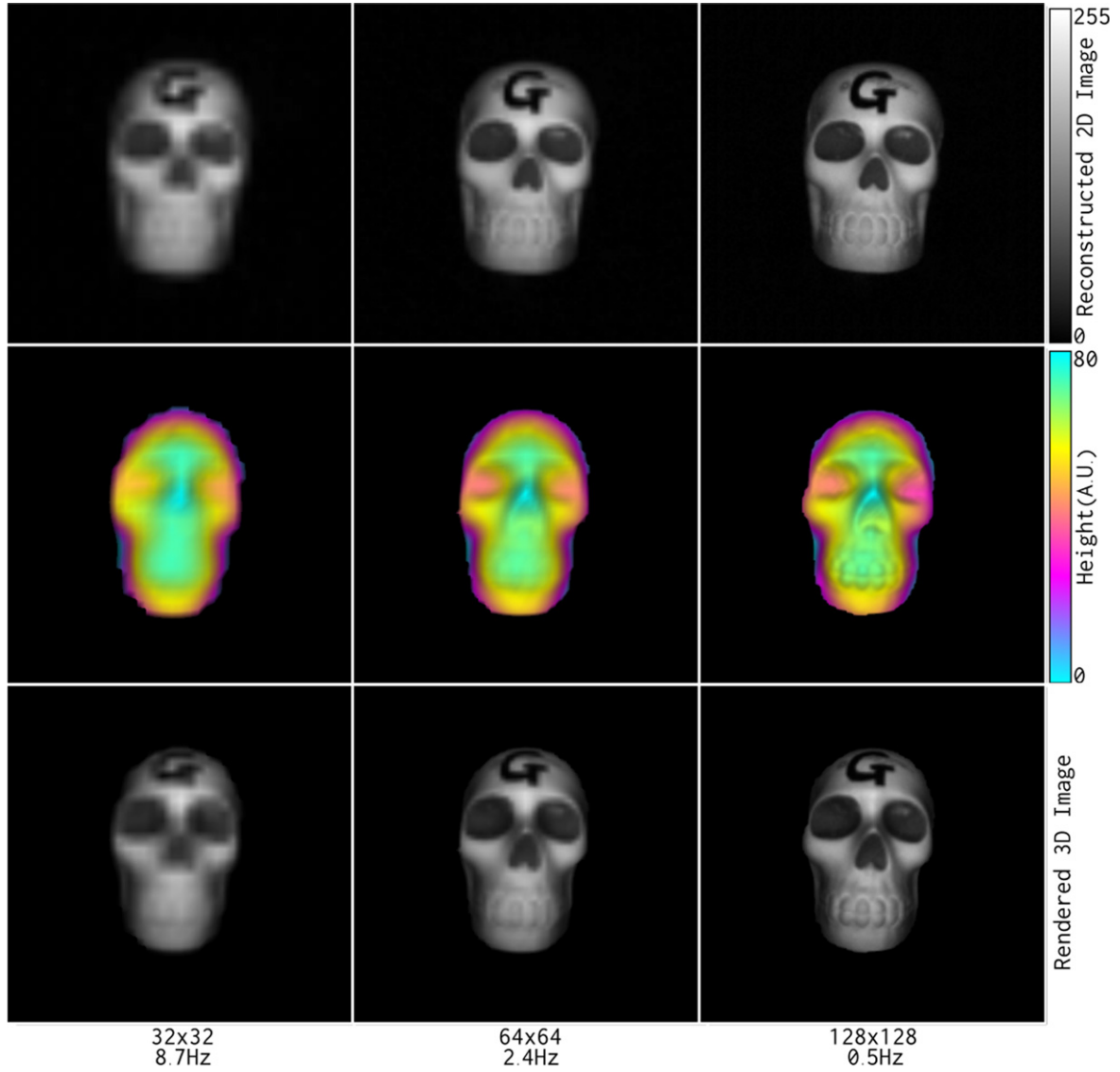
**Figure 2.** Comparison of 3D reconstructions of a static object using structured illumination with different resolutions. With illumination resolutions of $32 \times 32$, $64 \times 64$, and $128 \times 128$ Hadamard pattern pairs, the corresponding frame rate with our 3D imaging system are: 8.7 Hz, 2.4 Hz, and 0.5 Hz.

sequence. The display time can also be adjusted manually through the control software, which in this experiment is set to be $50\,\mu s$ as the minimum. Besides the high speed projection, another important feature of the ALP is that it provides synchronization trigger signals in reference to its display. That is, when one pattern gets displayed, there is a trigger signal released from digital micro-mirror device (DMD) control circuit to the DAQ input which then triggers a series of data acquisition. Signals are acquired after every trigger signal, and the sample numbers are determined by the display time and sampling rate. The DAQ used here is a National Instrument portable USB DAQ (NI USB-6221/16) with a maximum acquisition rate of 250 kHz for all channels. As there are four channels employed, sampling rate for each channel is set to 62.5 kHz. Given that each pattern is displayed for $50\,\mu s$, there are approximately three samples acquired for each pattern.

*Photometric stereo*

The image appearance of an object varies based on the lighting illumination, object orientation, object shape and its reflectance. With a static object, the corresponding surface orientation can be determined by analyzing the object images under different illumination directions. Photometric stereo, which is ideally for Lambertian surfaces, allows depth and surface orientation to be estimated from multiple images of a static object taken from the same viewpoint, but under different illumination directions. The appearance of a diffuse object with a specular varying reflection may be modeled as [37]:

$$I_P = \sum_{i=1}^{k} \rho_{i,P}^t [f_i(n_P, L_P, v)],$$

Where $I_P$ is the pixel intensity at point $p$, $k$ is a fixed value of a linear combination of $k$ basis materials, $\rho^t$ is a reflection
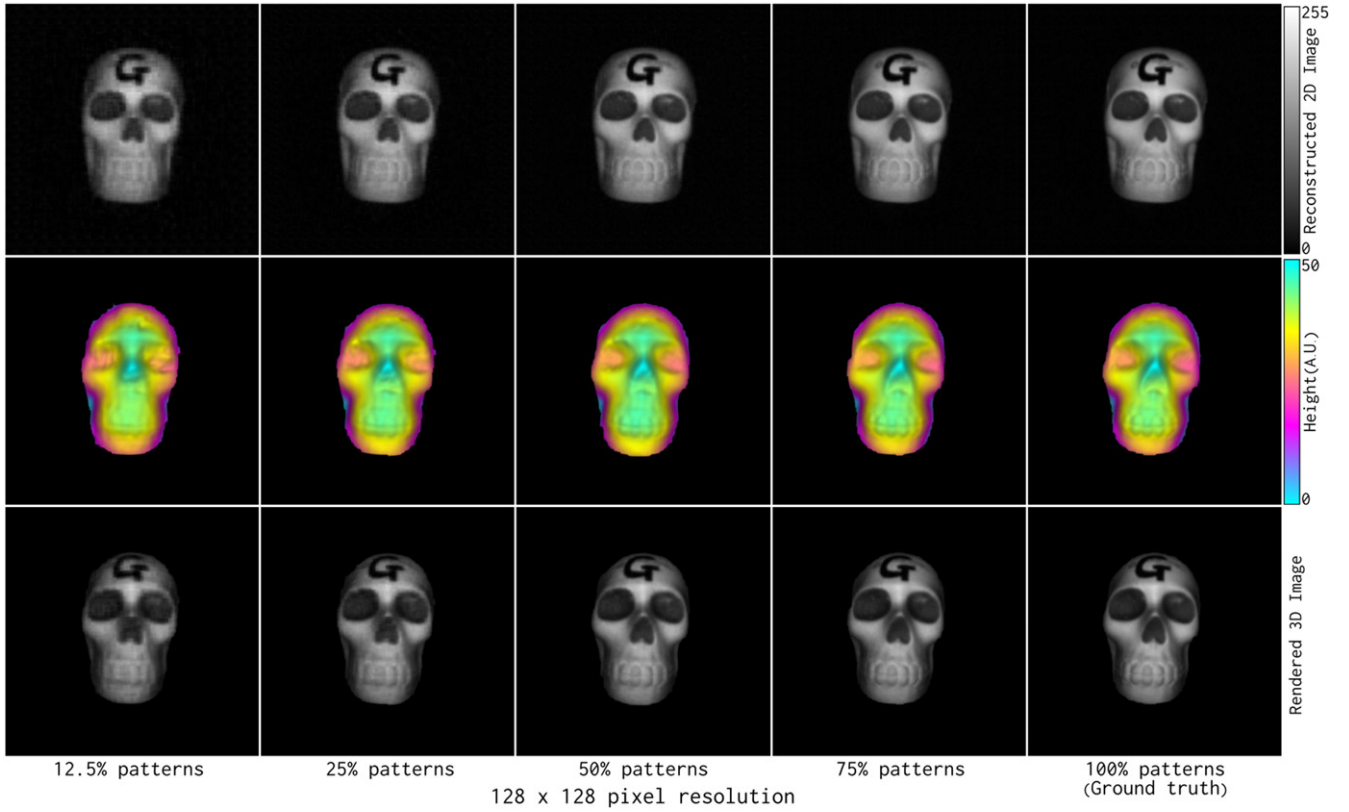
**Figure 3.** Comparison of 3D imaging using evolutionary compressed sensing. The object was reconstructed at 128 × 128 pixel resolution with five different compression ratios: 12.5%, 25%, 50%, 75% and 100%.

**Table 1.** Relative rms error comparison at 128 × 128 pixel resolution. The corresponding relative rms error of 3D reconstructed height data with five different compression ratios.

| | 128 × 128 pixel resolution | | | | |
|---|---|---|---|---|---|
| Patterns used | 12.50% | 25% | 50% | 75% | 100% (Ground truth) |
| Relative rms error | 3.586 | 3.443 | 2.229 | 1.571 | 0 |

coefficient that varies on the surface, $f_i$ is any reflectance map as a function of the viewing direction $v$, $n_p$ is the surface normal at that point, and $L_P$ is the incident illumination field.

In general, this method requires those images taken successively following the change of illumination directions. In our system, we replace the lighting sources with single-pixel detectors and the camera with a patterned lighting, in which case those images are acquired simultaneously.

*Basis-scanning with Hadamard matrices*

To maximize sampling efficiency, an orthogonal series of 2D binary patterns, derived from a Hadamard matrix [38], are preloaded to onboard RAM on the DMD. For a Hadamard matrix of order $4^k$ ($k > 0$, integer), each row is reshaped into a 2D array with $2^k \times 2^k$ resolution and upscaled to fill the full height of the DMD, resulting in complete series of $4^k$ binary

patterns. For example, for complete sampling of a scene at a resolution of 32 × 32 would require a series of 1024 reshaped Hadamard patterns to be displayed and the corresponding intensities measured by a photodetector.

## Results

Before assessing the 3D video quality with our 3D imaging system, we first considered fast 3D single-pixel imaging with a static object. In the system, as illustrated in figure 1, we combined a 3W white LED with a DMD of 1024 × 768 pixels, and a camera lens with a 24 mm focal length to deliver structured illumination at a rate of 22 kHz. A data acquisition board (DAB) was used to covert analog intensity measured by four photodetectors into digital signals at a rate of 250 kHz which subsequently were processed to reconstruct both 2D images via computational imaging and a 3D image based on photometric stereo. As used in other work with single-pixel cameras, the Hadamard basis was chosen for providing structured illumination, which yields better quality results compared to raster scanning techniques that suffer from poorer signal-to-noise [37]. In one investigation, we reconstructed a static object (a skull model) at three different resolutions: 32 × 32 pixels, 64 × 64 pixels and 128 × 128 pixels respectively, as shown in figure 2. As anticipated, increasing the image resolution provides improvement to image quality at the expense of the reconstruction frame-rate.
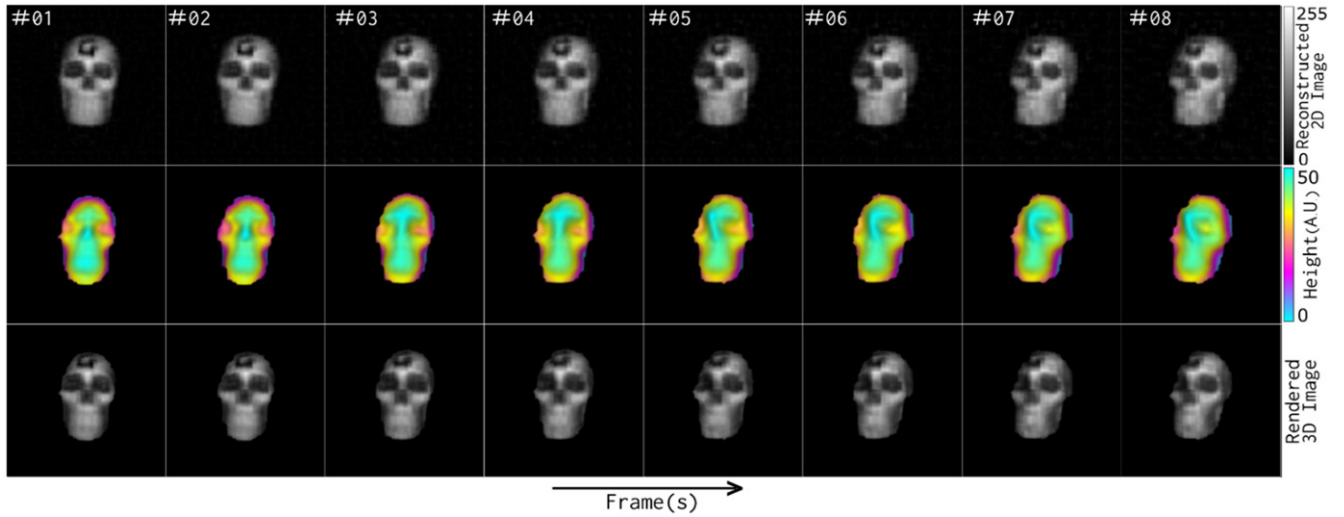
**Figure 4.** Sample of a compressed video-rate 64 × 64 pixel resolution 3D frames in 1 s. Each 3D frame is produced based on 1024 patterns (25% compression ratio) acquired at a frame rate of ∼8 Hz for an object physically rotated about the *z*-axis.
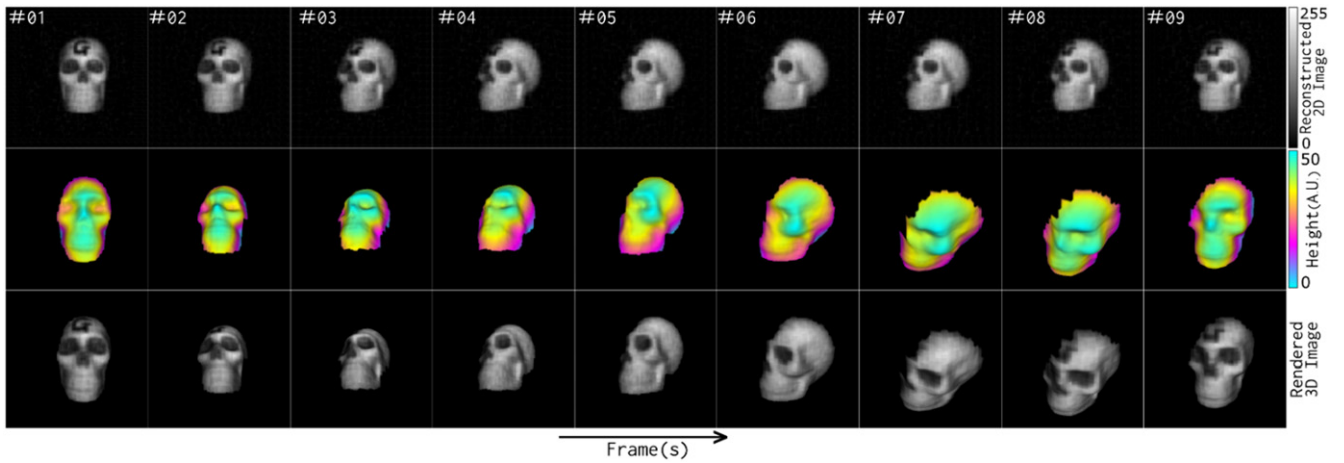


**Figure 5.** Sample of a 10 s 128 × 128 pixel resolution 3D video frames. Each 3D frame is produced based on 4096 patterns (25% compression ratio) acquired at approximately a frame rate of ∼1 Hz for an object physically rotated about the *z*-axis and the 3D reconstructed model rotated simultaneously about the *y*-axis on the front controller at an angle range of ±30°.

To improve the frame rate we applied compressive sensing algorithms. In choosing the optimal approach we noted that typical images can be represented by a subset of Hadamard patterns instead of a complete pattern set, and continuing adjacent frames are nearly coincident to each other with only slight variations. Hence we ordered the Hadamard patterns based on their corresponding mean signal intensities from the four photodetectors, and utilized the top-ranking patterns to form 2D images. With this compressive sensing algorithm, for each new frame we replaced a small percentage (we chose 10% patterns in this experiment) of low-ranking patterns among the top-ranking patterns with the ones that were randomly selected from the remainder of Hadamard patterns. For each frame, these four 2D images were combined using photometric stereo techniques to obtain 3D images. This evolutionary approach to the selection of a subset of Hadamard patterns is a compromise approach to maintain a high frame-rate without decreasing the spatial resolution.

With implementation of this photometric algorithm, we reconstructed the object at 128 × 128 pixel resolution by using five different numbers of pattern pairs: 16 384 pattern pairs, 12 288 pattern pairs, 8192 pattern pairs, 4096 pattern pairs, and 2048 pattern pairs, equivalent to 100% (zero-compression), 75%, 50% 25% and 12.5% compression ratio (see figure 3), and compared the relative root-mean square (rms) errors of the height value in those 3D images to the zero compression result (see table 1). The 3D reconstruction with zero-compression in this figure is the same as the 3D reconstruction of 128 × 128 pixel resolution in figure 2. The result shows that, as expected, the rms error of the object's height value increases when using less pattern pairs (higher compression).

In figure 4, we present a sample of real-time compressed 64 × 64 pixel resolution 3D video frames in 1 s. Each 3D frame is produced based on 1024 patterns (25% compression ratio), which equals to the same amount of patterns for a zero-compression 32 × 32 pixel resolution 3D image

reconstruction. The frame rate of this 3D video is 7.6 Hz, approximately 4 times faster than the zero-compression 64 × 64 pixel resolution one. In figure 5, we demonstrate a 10 s real-time 3D video frames reconstructed at 128 × 128 pixel resolution at a frame rate of 0.9 Hz, using 4096 Hadamard patterns (25% compression ratio). The object is physically rotated about the *z*-axis, whilst the 3D reconstructed model is rotated about the *y*-axis at an angel range of [−30, 30] controlled by the system since we are fully aware of the height map of the object. We notice that the frame rate in this video is restrained due to the fact that the 3D reconstruction process at 128 × 128 pixel resolution starts to play an important role in overall time performance when the 2D reconstruction time decreases with compressive sensing algorithm.

## Discussion and conclusion

We have experimentally demonstrated a video-rate 3D imaging system based on photometric stereo and exhibiting perfect pixel registration by utilizing a high-speed structured-illumination and single-pixel detectors. As an extension of our previous related work, we have now shown continuous operation for 3D image reconstruction at image resolutions of 32 × 32, 64 × 64 and 128 × 128 pixels, with Nyquist-sampling frame rates of 8.7 Hz, 2.4 Hz, and 0.5 Hz, respectively. Additionally we have made use of sub-Nyquist sampling (compressive sensing) to speed-up the frame rates for 64 × 64 and 128 × 128 pixel resolution images at the expense of only a modest reduction in image quality, as evidenced by a quantitative analysis. For 64 × 64 and 128 × 128 pixel 3D images, 25% compressive sensing provides increased frame rates of ~8 Hz and ~1 Hz respectively, compared to Nyquist sampling. We note that at 128 × 128 pixel resolution, the total computational time for 2D and 3D image reconstruction, as performed on an octa-core processor, placed a limit on the achievable frame rate, however we anticipate that this may be improved with other processing hardware, such as dedicated high-performance GPU's. Since this 3D imaging approach does not rely on a pixelated camera sensor and the operational bandwidth of DMD extends beyond this visible, this technique can be applied at wavelengths where 3D imaging is prohibitively expensive.

## Acknowledgments

## References

[1] Sun B, Edgar M P, Bowman R, Vittert L E, Welsh S, Bowman A and Padgett M J 2013 3D computational imaging with single-pixel detectors *Science* **340** 844–7

[2] Horn B K and Brooks M J 1989 *Shape from Shading* (Cambridge, MA: MIT Press)

[3] Pastoor S 1995 Human factors of 3D imaging: results of recent research at Heinrich-Hertz-Institut Berlin *Proc. IDW* **95** 69–72

[4] Dickson R M, Norris D J, Tzeng Y L and Moerner W E 1996 Three-dimensional imaging of single molecules solvated in pores of poly (acrylamide) gels *Science* **274** 966

[5] Hill M, O'Mahony C, Berney H, Hughes P J, Hynes E and Lane W A 2001 Verification of 2D MEMS model using optical profiling techniques *Opt. Laser Eng.* **36** 169–83

[6] Scharstein D and Szeliski R 2003 High-accuracy stereo depth maps using structured light *IEEE Comput. Vis. Pattern Recogn.* **1** 195–202

[7] Hariharan P 2010 *Basics of Interferometry* (Boston: Academic)

[8] Geng J 2011 Structured-light 3D surface imaging: a tutorial *Adv. Opt. Photon.* **3** 128–60

[9] Jung C and Kim C 2012 Real-time estimation of 3D scene geometry from a single image *Pattern Recogn.* **45** 3256–69

[10] Keppel E 1975 Approximating complex surfaces by triangulation of contour lines *IBM J. Res. and Dev.* **19** 2–11

[11] Boyde A 1985 Stereoscopic images in confocal (tandem scanning) microscopy *Science* **230** 1270–2

[12] Greenberg G and Boyde A 1993 Novel method for stereo imaging in light microscopy at high magnifications *Neuroimage* **1** 121–8

[13] Biskup B, Scharr H, Schurr U and Rascher U W E 2007 A stereo imaging system for measuring structural parameters of plant canopies *Plant Cell Environ.* **30** 1299–308

[14] Negahdaripour S 2007 Epipolar geometry of opti-acoustic stereo imaging *IEEE Trans. Pattern Anal. Mach. Intell.* **29** 1776–88

[15] Woodham R J 1980 Photometric method for determining surface orientation from multiple images *Opt. Eng.* **19** 139–44

[16] Okatani T and Deguchi K 2001 On uniqueness of solutions of the three-light-source photometric stereo: conditions on illumination configuration and surface reflectance *Comput. Vis. Image Und.* **81** 211–26

[17] Hertzmann A and Seitz S M 2003 Shape and materials by example: a photometric stereo approach *Proc. CVPR IEEE* vol 1 pp 533–40

[18] Barsky S and Petrou M 2003 The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows *IEEE Trans. Pattern Anal. Mach. Intell.* **25** 1239–52

[19] Basri R, Jacobs D and Kemelmacher I 2007 Photometric stereo with general, unknown lighting *Int. J. Comput. Vision* **72** 239–57

[20] Zhang Y, Graham M G, Hay R, Bowman R W, Padgett M J and Edgar M P 2015 A fast 3D reconstruction system with a low-cost camera accessory *Sci. Rep.* **5** e10909

[21] Horovitz I and Kiryati N 2004 Depth from gradient fields and control points: bias correction in photometric stereo *Image Vision Comput.* **22** 681–94

[22] Hernández C, Vogiatzis G and Cipolla R 2008 Multiview photometric stereo *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 548–54

[23] Kuparinen T and Kyrki V 2009 Optimal reconstruction of approximate planar surfaces using photometric stereo *IEEE Trans. Pattern Anal. Mach. Intell.* **31** 2282–9

[24] Chandraker M, Bai J and Ramamoorthi R 2013 On differential photometric reconstruction for unknown, isotropic BRDFs *IEEE Trans. Pattern Anal. Mach. Intell.* **35** 2941–55

[25] Fyffe G, Yu X and Debevec P 2011 Single-shot photometric stereo by spectral multiplexing *Proc. IEEE Int. Conf. Computational Photography* pp 1–6 DOI: 10.1109/ICCPHOT.2011.5753116

[26] Shapiro J H 2008 Computational ghost imaging *Phys. Rev.* A **78** 061802

[27] Duarte M F, Davenport M A, Takhar D, Laska J N, Sun T, Kelly K E and Baraniuk R G 2008 Single-pixel imaging via compressive sampling *IEEE Signal Process. Mag.* **25** 83

[28] Bromberg Y, Katz O and Silberberg Y 2009 Ghost imaging with a single detector *Phys. Rev.* A **79** 053840

[29] Lee K C, Ho J and Kriegman D 2001 Nine points of light: acquiring subspaces for face recognition under variable lighting *Proc. CVPR IEEE* **1** 519–26

[30] Schechner Y Y, Nayar S K and Belhumeur P N 2003 A theory of multiplexed illumination *Proc. CVPR IEEE* **2** 808–15

[31] Donoho D L 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289–306

[32] Candès E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489–509

[33] Chan W L, Charan K, Takhar D, Kelly K F, Baraniuk R G and Mittleman D M 2008 A single-pixel terahertz imaging system based on compressed sensing *Appl. Phys. Lett.* **93** 121105

[34] Beck A and Teboulle M 2009 A fast iterative shrinkage-thresholding algorithm for linear inverse problems *SIAM J. Imaging Sci.* **2** 183–202

[35] Baraniuk R G, Cevher V, Duarte M F and Hegde C 2010 Model-based compressive sensing *IEEE Trans. Inf. Theory* **56** 1982–2001

[36] 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** e10669

[37] Hertzmann A and Seitz S M 2005 Example-based photometric stereo: shape reconstruction with general, varying BRDFs *IEEE Trans. Pattern Anal. Mach. Intell.* **27** 1254–64

[38] Pratt W K, Kane J and Andrews H C 1969 Hadamard transform image coding *Proc. IEEE* **57** 58–68