Combining Multifocal Microscopy and Photometric Stereo with a Low Cost Motorized Positioner

Gustavo Lelis da Silva
Helena Cristina da Gama Leitão
Federal Fluminense University (UFF)
Niteroi, RJ, Brazil
Email: {g.lelis.silva@gmail.com
hcgl@ic.uff.br}

Jorge Stolfi
State University of Campinas (UNICAMP)
Campinas, SP, Brazil
Email: stolfi@ic.unicamp.br

Abstract—**Multifocal microscopy is an established technique to combine several reflected-light microscopy images of an object, each with a limited depth of focus, into a single image that shows the whole object in focus. Photometric stereo is an independent technique to recover the third dimension of an opaque object, given several images taken from the same viewpoint with different illumination conditions. We describe here the synergistic combination of both techniques to obtain accurate three-dimensional images of microscopic objects. In order to streamline the acquisition of the necessary images, we developed a low-cost positioner that automatically varies the microscope-to-object distance and the light direction, under control of an ARDUINO microcontroller. The main structural parts of the support were manufactured by a 3D printer. The technique and support are demonstrated with a low cost digital microscope.**

I. INTRODUCTION

An optical microscope usually can focus only on those points of the target object that lie in a narrow range of distances from the objective lens; that is, between two close planes perpendicular to the optical axis. The distance ε between these planes, called *depth of field*, is inversely proportional to the magnification of the microscope and to the effective diameter of the lens [1]. If the object of interest cannot fit between these two planes, it is necessary to acquire several images, varying the distance between the object and the microscope, so that every visible point of the object is in focus in at least one of the images. These images then can be composited to form a single *multifocus image* (MFI), showing the whole object in focus.

As a by-product, this technique also provides information about the shape of the object along the optical axis direction (Z-axis). This information has the form of an approximate *height map*, that, for each pixel of the combined image, specifies the Z coordinate of the point on the object's surface that is visible at that point. This method to recover the 3D shape of the object is known as *shape from focus* [2], [3], *depth from focus* [4] or *multifocus stereo* (MFS) for short.

Another way to acquire the three-dimensional shape of an object is to capture several images with the camera and object in the same relative position, varying the direction of the light that falls on the object [5], [6]. Analysis of how the apparent color of the object changes at each pixel yields the inclination of its visible surface in that pixel, as well as its "intrinsic" color. This slope information then can be integrated to recover the Z-coordinate at each pixel [7]. This technique, known as *photometric stereo* (PMS), can be used for scenes of arbitrary size, from planets (with the sun as light source) [8] to bacteria (by varying the position of the backscatter detector in a scanning electron microscope) [9].

In this paper, we describe how to synergistically combine these two techniques with low- to moderate-power reflected-light microscopy in order to obtain a height map of a small opaque object (up to a few mm in diameter) that is more accurate than what could be obtained with either technique alone. Namely, we obtain a collection of images $J_i^{(j)}$ for n microscope-to-object distances $i = 1, 2, \ldots, n$ and m illumination sources $j = 1, 2, \ldots, m$. The multifocal stereo technique is used to obtain a combined sharp image $\widetilde{J}^{(j)}$ for each light source j, and a *tentative height map* $\widetilde{Z}^{(j)}$. Then the photometric stereo technique is applied to these images, to obtain a single *final height map* $\widehat{Z}$ and an *intrinsic color map* $\widehat{A}$.

The photometric stereo technique is possible only if every visible point of the surface is illuminated by at least three of the light sources. Therefore, the number m of light sources must be at least 6 [10]. However, better results are obtained with a dozen or more light sources. As for MFS, the number n of images for each light source must be at least H/ε, where H is the total extent of the object in the Z-direction. For the objects and microscope used in this work, about two dozen images had to be acquired for each light source. Therefore, a minimum of a hundred images had to be acquired for each object.

In order to automate the acquisition of so many images, we built a simple motorized positioner that automatically varies the distance between the microscope and the object, and a set of 6 light LED light sources, all controlled by an ARDUINO microcontroller (see Figure 1). The frame is built from 3D-printed plastic elements and other easily-obtained parts. Our design is original, and may be an effective yet much cheaper alternative to commercial motorized microscope stands.
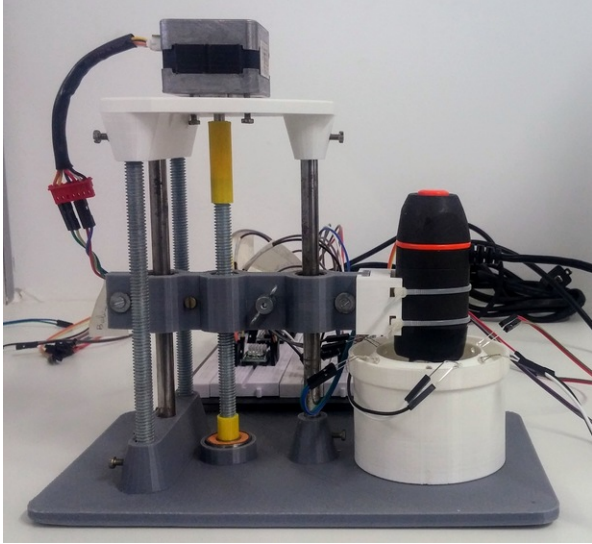
Fig. 1. Image of the motorized positioner with microscope.

II. MULTIFOCAL STEREO

A. History and related work

Multifocal imaging and multifocus stereo are well established techniques of optical microscopy. A 1987 paper by Grossmann [4] described how to estimate the depth of different parts of an object from the degree of blurring in a single image of the same. In 1989, Shree Nayar [2] described how to combine multifocus image stacks to obtain a single sharp image and a depth map of the object, which is still the basis for present MFI and MFS.

There is an extensive literature on multifocus microscopy. A brief survey was provided in 2012 by Thiery and Green [11], and also in 2003 by Zamofing et al. [12]. Many papers describe its application to specific problems and disciplines, especially macro photography. Uses include the imaging of minerals [11], measurement of small metallic and wood parts [12], palynology [13], and archaeology and preservation of cultural artifacts [14].

An interesting variation on the technique is the 2006 proposal by Marc Levoy, Ren Ng et al. [15], which uses a special camera with multiple small lenses, each with a different depth of focus, in order to capture all the needed views in a single image shooting. This method could allow MFI and SF to be used on small moving objects, possibly in real time, at the cost of reduced lateral resolution.

B. Theory and notation

The input to the multifocal stereo technique is a series (*stack*) of n images $J_1, J_2, \ldots, J_n$ (*frames*), taken with object and microscope in the same relative position except for the microscope-to-object distance, and the same illumination conditions. The images are assumed to be perfectly aligned and scaled so that the pixel $J_i[x, y]$, for all j, shows the same point on the object's surface

The only pixels of each image J_i that are effectively in focus are those where the Z-coordinate of the visible surface lies in an interval $\overline{z}_i \pm \varepsilon/2$ [16], for some *focus midplane height* $\overline{z}_i$ specific to each image. See figure 2. Typically, the heights $\overline{z}_i$ are equally spaced with some step $\delta \leq \varepsilon$, and completely span the relevant range of Z coordinates – namely, with $\overline{z}_1$ below the lowest point visible in the image, and $\overline{z}_n$ above the highest point. We will define arbitrarily the first mid-plane position $\overline{z}_1$ as the origin of the Z-axis, so that $\overline{z}_i = (i - 1)\delta$ for all i.
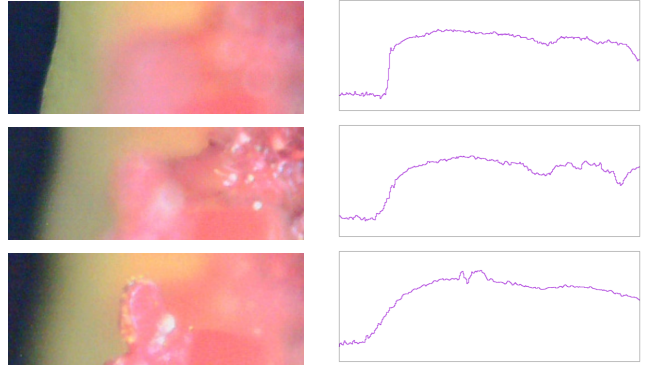

Fig. 2. Images of a small target object taken with different object-to-microscope distances. The plot at the right each image shows the brightness of pixels along a single scan line through the center of the image.

The first step in multifocal stereo is to identify the pixels of each image J_i where the object is in focus. This classification is based on the fact that the parts that are not in focus cannot have abrupt brightness and color variations, such as sharp edges and small spots. Thus, if the image gradients around a pixel $[x, y]$ are greater in an image J' than in another image J'', then we may conclude that that part of J' is better focus.

Using this criterion, a numerical *focus score* $\alpha_i[x, y]$ is given for each pixel $J_i[x, y]$ of each image in the stack. See figure 3. A simple choice could be the magnitude of the gradient in a small neighborhood of the pixel $[x, y]$, either absolute or relative to the average image intensity. Many other formulas have been proposed, using higher derivatives or the high-frequency components of wavelet-like operators [13]. A comparative survey of focus score operators was recently provided by Pertuz et al. [17].

In the simplest approach, for each pixel position $[x, y]$, the value $\widetilde{J}[x, y]$ of the combined image is copied from the corresponding pixel $J_i[x, y]$ with the largest score $\alpha_i[x, y]$. The approximate height $\widetilde{Z}[x, y]$ of that pixel is then set to the height $\overline{z}_i$, with implicit uncertainty $\pm\delta/2$.

Better results can be obtained by combining the pixels $J_i[x, y]$ in some sort of average, instead of simply selecting the one with largest score. Namely

$$\widetilde{J}[x, y] = \sum_{i=1}^{n} \lambda_i[x, y]\, J_i[x, y] \tag{1}$$

where the *blending coefficients* $\lambda_i[x, y]$, for each pixel $[x, y]$ are a partition of unity derived from the indicators $\alpha_i[x, y]$.

For example, one could simply take

$$\lambda_i = \frac{\alpha_i[x,y]}{\sum_{k=1}^{n} \alpha_k[x,y]} \qquad (2)$$

for all pixels $[x,y]$. See figure 3. The height $\widetilde{Z}[x,y]$ and its uncertainty can then be estimated by the weighted average and deviation of the focus plane heights $\overline{z}_i$, with weights $\lambda_i[x,y]$.
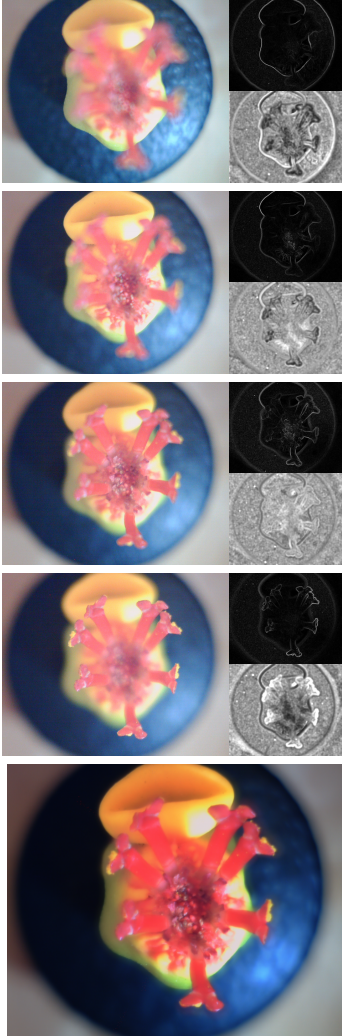

Fig. 3. Illustration of the multifocus imaging technique. The images at left are frames J_1, J_9, J_{17}, and J_{25} of a 25-frame stack. The image pairs at right are the focus indicator map α_i (top) and the blending coefficient map λ_i (bottom) for the frame at left. The larger image at bottom is the sharp composite image $\widetilde{J}$.

C. Software

There are several tools available to perform MFI and MFS from multifocus image stacks. Popular ones are COMBINEZP (free) [18], [19], HELICON FOCUS (commercial) [20], DEPTH FROM FOCUS (free) [21], and COSMOPLAYER (free) [21]. Also, many researchers have developed their own MFI/MFS tools, often with original algorithms [15], [21].

For the examples shown in this article, we aligned the frames $J_1^{(j)}, J_2^{(j)}, \ldots, J_n^{(j)}$ of each image stack j with the

`align_stack_image` tool from the HUGIN package [22], with parameters `−m −x −y −z −C −g 5`. The composite image $\widetilde{J}^{(j)}$ at the bottom of figure 3 was produced with our own experimental MFI composition program `multifok` [23]. For the examples in sections III and IV we used the HELICON FOCUS commercial MFI software.

III. PHOTOMETRIC STEREO

A. History and related work

The physical principles of photometric stereo have been known since the development of photometry in 1760 by Lambert [24], and their use by astronomers in the early 20th century [8].

The first proposals to exploit those principles for computer vision, such the pioneering 1977 article by B. K. P. Horn [25], assumed very simple settings: a single image of a smooth object with uniform color and Lambertian finish, illuminated by a single light source, distant and point-like (i.e. by a uniform and unidirectional light field). However, this basic single-image *shape from shading* problem was ill-posed, since the the shading of a surface element does not determine the direction of the normal.

The first description of the true photometric stereo method, using multiple images under different light fields, was provided by Woodham in 1980 [5]. This idea has been extensively researched since then, with significant improvements such as reduced constraints on the nature of the light sources and optical properties of the surface, and reduced need for calibration. For example, Hayakawa in 1994 [26] and Yuille et al. in 1997 [27] used SVD and PCA analysis to extract surface normals without a priori information about the light sources

The most liberal and reliable methods to determine the normals use an *example object* or *light gauge* — an object of known shape that is imaged under the same light conditions as the target object, and has similar surface characteristics except possibly for intrinsic color. This technique was used already by Woodham in 1980 [5], and was improved in speed, accuracy, and generality by Hertzmann [28], Saracchini et al. [29], and Quéau et al. [30].

B. Theory and notation

In photometric stereo, m images $J^{(1)}, J^{(2)}, \ldots, J^{(m)}$ are taken with the object and microscope in the same relative position, but using m different light sources $L^{(1)}, L^{(2)}, \ldots, L^{(m)}$ in different directions. As in multifocus imaging, we assume that the images are perfectly aligned and scaled, so that the pixel $J^{(j)}[x,y]$, for all j, shows the same point on the object's surface — which we denote by $p[x,y]$.

To simplify the description, let's assume that the light sources are monochromatic with the same spectral color, and the value $J^{(j)}[x,y]$ of each pixel is a real number proportional to the apparent brightness of the bit of surface around $p[x,y]$; that is, to the luminous power per unit of area scattered by that bit of surface towards the microscope's objective lens. (In other words, we assume a linear brightness encoding, rather than the "gamma" encoding used in most image files.)

Let $\vec{u}[x,y]$ be the local *normal*, that is, the mean unit direction vector perpendicular to the surface at $p[x,y]$ and pointing out of the object. According to physical optics, the apparent brightness of the image pixels can be expressed as

$$J^{(j)}[x,y] = A[x,y]\,\mathcal{L}^{(j)}[x,y](\vec{u}[x,y]) \qquad (3)$$

Here, $A[x,y]$ is a numerical coefficient between 0 and 1, the *albedo* or *total scattering coefficient* of the surface at that point, namely the percentage of incident light that is scattered back to the environment (rather than absorbed); and $\mathcal{L}^{(j)}[x,y]$ is the *shading function* for that part of the surface. The albedo is supposed to be an intrinsic property of the surface at that point, that describes its intrinsic "color", distinguishing a white surface ($A = 1$) from a gray one ($A < 1$) or a black one ($A = 0$).

The lighting function is supposed to be independent of the albedo, but may depend on the surface finish (matte, glossy, etc.). Let $\Phi^{(j)}[x,y](\vec{v})$ denote the incoming *light flux* reaching that spot from direction $\vec{v}$; that is, the apparent brightness of the environment as would be seen by a tiny observer standing on that surface element and looking towards the direction $\vec{v}$. Let also $\vec{w}[x,y]$ be the direction from $p[x,y]$ to the camera's nominal position. The value of $\mathcal{L}^{(j)}[x,y](\vec{u})$, for a surface with orientation $\vec{u}$, is determined by these quantities and by the intrinsic optical properties of the surface at that point, according to the equation

$$\mathcal{L}^{(j)}[x,y](\vec{u}) = \int_{\mathbb{S}^2} \Phi^{(j)}[x,y](\vec{v})\,\beta[x,y](\vec{u},\vec{w}[x,y],\vec{v})\,d\vec{v} \qquad (4)$$

Here $\beta[x,y]$ is the (*normalized*) *bidirectional radiance distribution function* (BRDF), that tells what fraction of the light that arrives from direction $\vec{v}$ and is not absorbed by the surface will be scattered in the direction $\vec{w}[x,y]$.

The primary goal of photometric stereo is to solve the m equations (3) for each pixel $[x,y]$, in order to recover the albedo $A[x,y]$ and the surface normal $\vec{u}[x,y]$. It should be noted that each equation (3) is useless if the point $p[x,y]$ is shadowed relative to light j, so that it receives no light (or only ambient light, which is often mostly isotropic and indeterminate). Even if the functions $\mathcal{L}^{(j)}[x,y]$ are fully known for all j, there are still three unknowns (A and the two horizontal components of $\vec{u}$), and therefore at least three of the equations must be independent and non-trivial. We conclude that, the problem can be solved only if $p[x,y]$ is sufficiently well illuminated by at least three of the light sources. Even for the simplest convex objects, six light sources is the minimum needed to satisfy this requirement.

Determining the functions $\mathcal{L}^{(j)}[x,y]$ is a hard problem by itself. Equation (4) requires knowledge of the BDRF $\beta[x,y]$ and of the incoming light field $\Phi^{(j)}[x,y]$. One way to attack this problem is to restrict the situation, e. g. by requiring $\Phi^{(j)}$ to be essentially unidirectional and uniform and/or requiring the surface to be Lambertian.

Alternatively, the functions $\mathcal{L}^{(j)}[x,y]$ may be estimated by imaging one or more example objects in the same lighting conditions. One such gauge is enough if the incoming light field Φ can be assumed to be uniform; that is, essentially the same for all fully illuminated points $p[x,y]$. That assumption can be made if both the object and the light source are small compared to the distance between the two — as is the case for the setup described in section V.

Once the normal map $\vec{u}[x,y]$ has been determined, the height map $\widehat{Z}$ can be obtained by one of several *surface gradient integration* method [7], [31], [32].
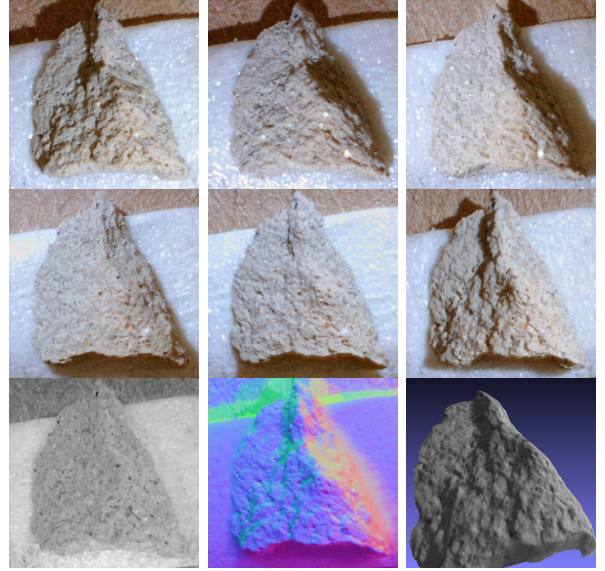

Fig. 4. Photometric Stereo input images J_i and output from the method: the albedo map A, normal map $\vec{u}$ and the derived height map from a surface integration method [7].

C. Software

Unlike multifocus imaging, photometric stereo is still a developing field. The quality of the results is very dependent on setup conditions such as stray light and uneven surface finish. Tools for PMS are relatively few and mostly created as research projects rather than software for general use.

For the examples in figures 4 and 5, the six multifocus images $\widetilde{Z}^{(1)}, \widetilde{Z}^{(2)}, \ldots \widetilde{Z}^{(6)}$ obtained as explained in section II-C were combined by the photometric stereo tools of Saracchini et al. [33] to obtain the estimated height map $\widehat{Z}$ and albedo map $\widehat{A}$.

Those tools use an example-based normal estimation algorithm [29] and a multi-scale gradient integrator [7]. The example object that we used for the former was a small glossy spherical object. A smoothing function based in radial basis [34] was applied over the measured samples to remove scattered "dust" noise.

IV. SYNERGY

Multifocus image composition is a necessary step before applying photometric stereo to small objects, because the normal estimation methods do not work on parts of the image that are out of focus. Except for fully illuminated Lambertian surfaces, the shading equations (3) are highly non-linear in

the normal $\vec{u}$, especially for rough surfaces. Therefore, the normal map is not only blurred in those parts, but in fact quite distorted.

But there is more: the two methods actually complement each other. Focus detection fails on regions where the object's surface is featureless [12]. Those parts will receive a low focus score $\alpha_i^{(j)}[x, y]$ even if the input image $J_i^{(j)}$ is perfectly in focus in that region. The composite image $\widetilde{J}^{(i)}[x, y]$ will be valid anyway, because any selection or blending criterion will give the same output pixel value. However, the estimated height $\widetilde{Z}[x, y]$ will be totally uncertain, and would have to be set to the entire interval $[\overline{z}_1 - \delta \_ \overline{z}_n + \delta]$. On the other hand, photometric stereo works well where the surface is smooth and has uniform BRDF. In those places the computed normal and the relative integrated heights are more likely to be correct.

In fact, the two methods can interact synergistically to yield a stereo method that is better than just using one or the other different parts of the image. Photometric stereo usually fails if the normal map is incorrect, even in a limited region of the images' domain, because the integration step can propagate the errors to the whole height map. In particular, the method fails if the height map has *cliffs* (discontinuities), such as along the silhouette edges of the target object. To obtain a meaningful height map, the user must provide a binary *validity mask* $W[x, y]$ that is zero where the normal $\vec{u}[x, y]$ is known or suspected to be incorrect. Moreover, if those marked-out pixels split the domain into two or more disconnected regions, the photometric stereo method cannot determine the relative heights of those regions.

The approximate height map $\widetilde{Z}$ produced by the multifocus stereo stage can be used to solve these problems. Significant cliffs and regions with ill-defined normal (such as distant background) can be usually be detected from $\widetilde{Z}$ By simple morphological operators such as thresholding, dilatation and erosion, the required validity mask W can be derived from $\widetilde{Z}$, even without manual user intervention. Even if this mask splits the domain into several disconnected regions, the map $\widetilde{Z}$ provides the approximate relative heights of those regions. See figure 5. Other examples of reconstructions using this approach are shown in figure 6

V. THE POSITIONER

A. Overview

The positioner consists of a stationary supporting *frame*, a moving *carriage* that holds the microscope, and a hollow cylindrical *lighting stage* that surrounds the target object. The carriage is displaced vertically along two fixed guide rails by a screw, which in turn is driven by a stepper motor. See figure 7.

The object is illuminated by a ring of six white light-emitting diodes (LEDs) around the top edge of the stage, so that the light makes an angle of about 45 degrees with the vertical.

Low-level control of the LEDs and the stepper motor is performed by an ARDUINO NANO [35] microprocessor, in response to higher-level commands received from a host computer through an USB serial interface. The ARDUINO
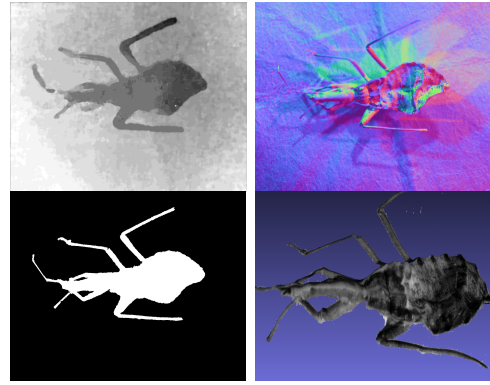

Fig. 5. Synergistic combination of multifocus stereo and photometric stereo to reconstruct complex objects. At the top, the approximate height map f$\widetilde{Z}$ from the multifocus stereo, and the normal map $\vec{u}$ from obtained by photometric stereo. At bottom, the validity mask W derived from $\widetilde{Z}$, and the height map obtained by integration of $\vec{u}$ with the validity mask W.

and other control circuitry was assembled on a protoboard. The microscope's camera is directly controlled by the host computer.

The microscope is initially positioned by the operator (by issuing commands directly to the ARDUINO, via its development environment and USB interface), so that its focus midplane lies below the lowest points visible in the camera. That action defines the first midplane position $\overline{z}_1 = 0$. Then a script is executed in the host computer, that repeatedly commands the ARDUINO to turn on and off the six LEDs in sequence, then displace the microscope up by δ, and capture the corresponding image $J_i^{(k)}$ after each setting.

The microscope we used to generate the examples in this article was the MINI HANDHELD DIGITAL MICROSCOPE manufactured by CELESTRON [36].

B. Mechanical

Most structural parts of the frame and carriage are made of polylactic acid (PLA) thermoplastic, and were manufactured with a desktop 3D printer (MAKERBOT REPLICATOR 2). The geometric models were created by us, using the CAD software supplied with the printer.

The frame consists of two plastic parts, a *base* plate, and a *head* plate connected by four steel *supporting rods* with 8 mm diameter. Two of the rods are smooth and also function as rails for the carriage. (The other two rods are threaded for incidental reasons, but their thread plays no role.) The support bars are inserted snugly into 20 mm deep holes in the plastic parts, and held in place by set screws self-threaded into the plastic.

The carriage consists mainly of two symmetric plastic *braces* and a plastic microscope *holder*. The two braces are held together by three screws with nuts and washers. The braces in turn hold a hex nut, 26 mm long, that rides on the driving threaded bar; and two linear bearings that slide along the rail bars. The holder block is attached to the front of the two braces by a set screw, and is supposed to be specific to
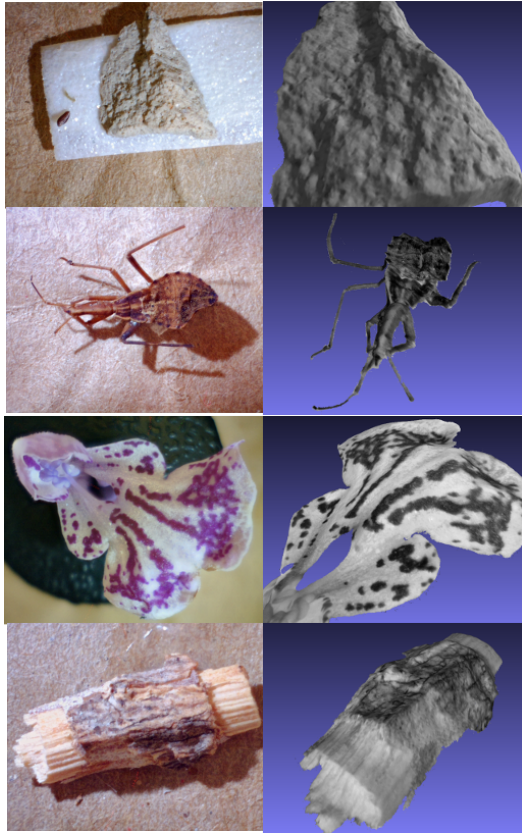
Fig. 6. Reconstruction of four additional test objects by the method and apparatus described here. At left, for each object, one of the six merged images $\widetilde{J}^{(j)}$ that were input to the photometric stereo method. At right, the reconstructed height map $\widehat{Z}$.
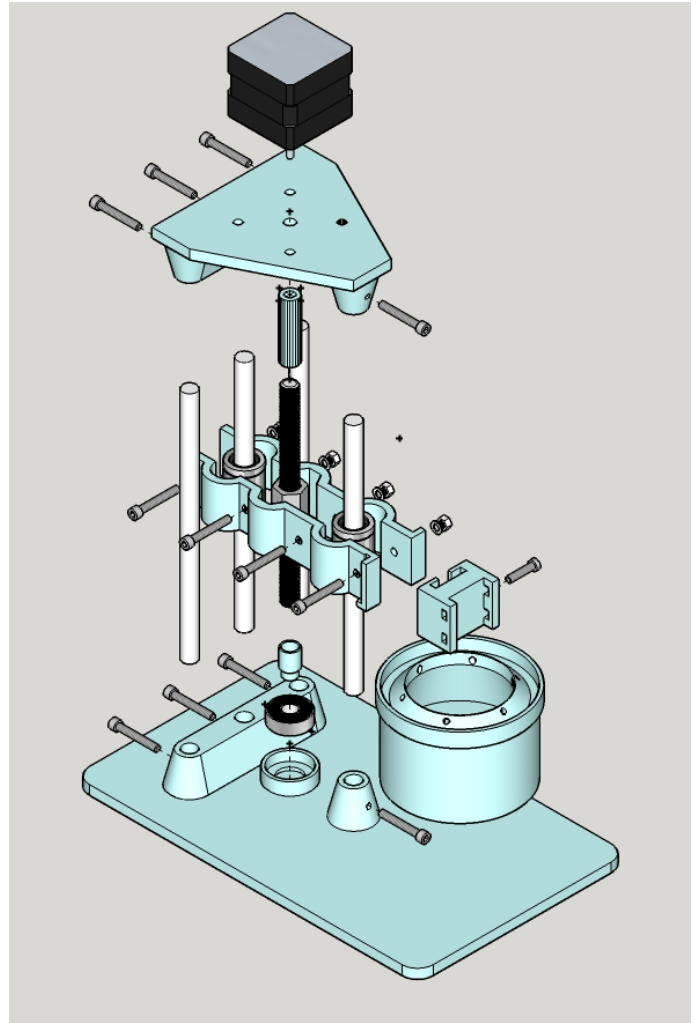

Fig. 7. Exploded view of the microscope positioner used in our experiments. The parts in light blue are made of PLA plastic. The steel support and rail rods are shown in white. The driving mechanism, comprising the stepper motor (at top), the threaded rod, and the roller bearing (at bottom), are shown in gray. The carriage consists of the two braces holding the nut and linear bearings, and the holder block attached to their front end.

the imaging device used. In our tests, it was a broad U-shaped bracket with four holes, and the microscope was fastened to it by two nylon cable ties.

The mechanism that drives the carriage consists of a stepper motor mounted on the head plate, connected to a threaded metal rod by a plastic coupler. The lower end of the threaded rod is inserted, via a second plastic coupler, into a skateboard roller bearing that is embedded in the base plate.

C. Stepper motor electronics

The stepper motor (NEMA 17 with 1.1 kgf of torque) is controlled by digital output pins 4 and 7 of the ARDUINO through a power drive module (brAllegro A4988). Output pin 2 is connected to the ENABLE port of the power module and is used to shut off the motor outputs.

The low current consumption of the power drive module allows the electrical power for its logic circuitry to be supplied by the ARDUINO's power supply ports (+5V and GND). The function of this module is to convert the low-power $0\,V/+5\,V$ binary signals from the ARDUINO to the 4-wire 16.8 V bipolar signals required by the motor. It connects to the motor via outputs 2B, 2A, 1B and 1A.

Power for the motor itself is provided by a standard industrial power source that converts $\approx 120\,V$ AC to 20 V DC.

The output of this power source is connected to the power driver module through a voltage regulator circuit (TEXAS INSTRUMENTS LM2596) that lowers the voltage to 16.8 V. In order to to attenuate possible voltage spikes and prevent a possible power-drive burnout, a $10\,\mu F$ capacitor was connected in parallel at the output of the regulator circuit.

D. Lighting electronics

The light sources used for this article were 6 white-light LEDs, rated 3.6 V and 30 mA, controlled by outputs 3, 5, 6, 9, 10, and 11 of the ARDUINO. These output were operated in pulse 500 Hz width modulation (PWM) mode, to allow their brightness to be adjusted.

We found that the 500 Hz flicker interfered badly with the image capture. Therefore we found it necessary to use low pass filters to turn the $0\text{–}5\,uV$ PWM signal into a continuous voltage and current drive, to keep the LEDs at constant

brightness. Each filter consists of a 100μF capacitor in parallel with the LED, and a $220\,\Omega$ resistor in series with both.
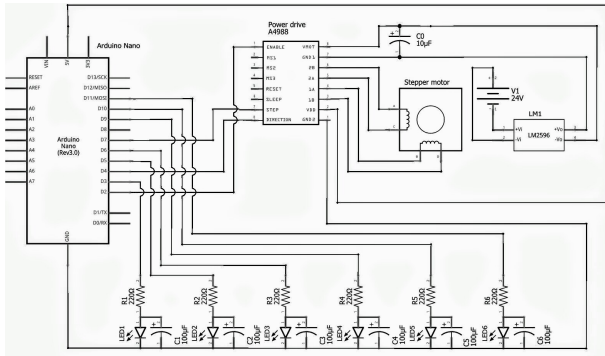

Fig. 8. Motorized Positioner Circuit Diagram.

E. Embedded ARDUINO program

We developed the embedded program that runs on the microcontroller using the standard ARDUINO C-like programming language and various free libraries for stepper motor control and other tasks [37].

The embedded program receives higher-level commands and parameters through the ARDUINO's serial communication (USB) port. The commands can be typed directly by the user, through the serial communication window of the ARDUINO development environment. Alternatively, they can be issued automatically by a program running on the host computer.

The commands understood by the embedded software include "displace the microscope" by a previously specified step δ (in μm) and "turn LED j on/off". The complete list is provided in the software's documentation [38].

F. Image acquisition script

Apart from initial positioning, the image acquisition process is automated by a simple Python program (**Focus**) that runs in the host computer and sends high-level commands to the ARDUINO. It receives as parameters the step δ between successive focus midplanes, the number n of images in each image stack, and the intensity of the light sources.

The open source software **uvccapture** was used to download the images from the microscope. Images were stored at 2560×2048 resolution in **jpg** format.

VI. CONCLUSIONS AND FUTURE WORK

We demonstrated the effective synergistic combination of multifocal stereo and photometric stereo to obtain fully in-focus images and height maps for microscopic objects, using simple and extremely low cost hardware.

Many improvement suggest themselves, to the positioner and the image capture process, as well as to the processing of the captured images. In particular, the depth information resulting the photometric stereo step could be used to improve the accuracy of a second pass of the multifocal stereo procedure.

With proper light sources and filters, the photometric stereo procedure (and hence the synergistic combination of MFS

and PMS) can be used to obtain albedo maps in arbitrary color bands, which could allow identification of materials indistinguishable by eye.

As for the motorized stand, it would be desirable to increase the number m of distinct light sources for each frame (to, say, 20 or more), to improve the accuracy and reliability of the PMS step. However, the number of LEDs is limited by the number of PWM output pins available from the ARDUINO. One way to increase m, without using more LEDs and pins, would be to turn the lighting stage around the optical axis with a second stepper motor. Then one could place the six LEDS at various inclination angles from the axis, to obtain better coverage of the direction sphere.

REFERENCES

[1] V. Ciciliato and J. G. Chiquito, "Modelamento do microscópio óptico como um sistema linear variante com o deslocamento," in *Anais do 13o Simpòsio Brasileiro de Telecomunicações (SBrT 1995)*, vol. I, 1995, pp. 258–263.

[2] S. K. Nayar, "Shape from focus," Robotics Institute, Carnegie Mellon University, Tech. Rep. CMU-RI-TR-89-27, Nov. 1989.

[3] J. L. Fernandes and J. R. A. Torreão, "Shape from shading through photometric motion," *Pattern Anal. Appl.*, vol. 13, no. 1, pp. 35–58, 2010.

[4] P. Grossmann, "Depth from focus," *Pattern Recognition Letters*, vol. 5, no. 1, pp. 63–69, 1987.

[5] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, pp. 139–144, 1980.

[6] E. Coleman and R. Jain, "Obtaining 3-dimensional shape of textured and specular surface using four-source photometry," *Computer Graphics and Image Processing*, vol. 18, pp. 308–328, 1982.

[7] R. F. V. Saracchini, J. Stolfi, H. C. da Gama Leitão, G. A. Atkinson, and M. L. Smith, "A robust multi-scale integration method to obtain the depth from gradient maps," *Computer Vision and Image Understanding*, vol. 116, no. 8, pp. 882–895, 2012.

[8] T. Rindfleisch, "Photometric method for Lunar topography," *Photometric Engineering*, vol. 32, no. 3, pp. 262–276, Mar. 1966.

[9] A. Miyamoto, H. Matsuse, and G. Koutaki, "Robust surface reconstruction by design-guided SEM photometric stereo," *Measurement Science and Technology*, vol. 28, no. 4, Feb. 2017.

[10] H. C. da Gama Leitão, R. F. V. Saracchini, and J. Stolfi, "Matching photometric observation vectors with shadows and variable albedo," in *Proceedings of the 21th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2008)*. IEEE Computer Society Press, Oct. 2008, pp. 179–186.

[11] V. Thiery and D. I. Green, "The multifocus imaging technique in petrology," *Computers & Geosciences*, vol. 45, pp. 131 – 138, 2012.

[12] T. Zamofing and H. Hugli, "Applied multifocus 3D microscopy," in *Proc. SPIE*, vol. 5265, 2004, pp. 134–144.

[13] A. Bercovici, A. Hadley, and U. Villanueva-Amadoz, "Improving depth of field resolution for palynological photomicrography," *Palaeontologia Electronica*, vol. 12, no. 2, p. 12, 2009.

[14] H. Plisson and L. V. Zotkina, "From 2D to 3D at macro- and microscopic scale in rock art studies," *Digital Applications in Archaeology and Cultural Heritage*, vol. 2, no. 2–3, pp. 102–119, 2015.

[15] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," in *ACM SIGGRAPH 2006 Papers*, ser. SIGGRAPH 2006. New York, NY, USA: ACM, 2006, pp. 924–934.

[16] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image and Vision Computing*, vol. 26, no. 7, pp. 971–979, 2008.

[17] S. Pertuz, D. Puig, and M. Ángel García, "Analysis of focus measure operators in shape-from-focus," *Pattern Recognition*, vol. 46, no. 5, Nov. 2012.

[18] A. Hadley, "COMBINEZP," Free software, available online at http://hadleyweb.pwp.blueyonder.co.uk/, 2017.

[19] J. Russell and R. Cohn, *CombineZP*. Book on Demand, 2012.

[20] H. Ltd., "HELICON FOCUS 6," Commercial software available at http://www.heliconsoft.com/heliconsoft-products/helicon-focus/, 2017.

[21] M. Niederösta, J. Niederösta, and J. Ščučkab, "Automatic 3D reconstruction and visualization of microscopic objects from a monoscopic multi-focus image sequence," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIV-5/W10, pp. 1–8, 2003.

[22] P. D'Angelo, "HUGIN panorama photo stitcher," Free software available at http://hugin.sourceforge.net/, 2017.

[23] G. L. da Silva, J. Stolfi, and H. C. da Gama Leitao, "MULTIFOK: A tool for multifocus stereo," Free software, available from authors, 2017.

[24] J. H. Lambert, *Photometria, sive de Mensura et Gradibus Luminis, Colorum et Umbrae*, Augsburg, 1760.

[25] B. K. P. Horn, "Understanding image intensities," pp. 201–231, 1977.

[26] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Journal of the Optical Society of America, Series A*, vol. 11, no. 11, pp. 3079–3089, 1994.

[27] A. Yuille and D. Snow, "Shape and albedo from multiple images using integrability," in *Proc. 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 1997, pp. 158–164.

[28] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying brdfs from photometric stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1060–1071, 2010.

[29] R. F. V. Saracchini, J. Stolfi, and H. C. da Gama Leitão, "A uniform grid structure to speed up example-based photometric stereo," *IEEE Trans. Image Processing*, vol. 20, no. 12, pp. 3495–3507, 2011.

[30] Y. Quéau, F. Lauze, and J.-D. Durou, "A L^1-TV algorithm for robust perspective photometric stereo with spatially-varying lightings," in *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer, 2015, pp. 498–510.

[31] R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 10, no. 4, pp. 439–451, 1988.

[32] A. Agrawal, R. Raskar, and R. Chellappa, "What is the range of surface reconstructions from a gradient field?" *Computer Vision–ECCV 2006*, pp. 578–591, 2006.

[33] R. Saracchini and J. Stolfi, "PHOTOMETRIC STEREO TOOLS," Free software, available online at http://www.ic.unicamp.br/~stolfi/EXPORT/projects/photo-stereo/.

[34] M. J. D. Powell, "Radial basis functions for multivariable interpolation: A review," in *Algorithms for approximation*. Clarendon Press, 1987, pp. 143–167.

[35] A. AG, "Getting started with the ARDUINO NANO," Online document at https://www.arduino.cc/en/Guide/ArduinoNano, 2017.

[36] Celestron Inc., "Mini handheld digital microscope," Online document at http://www.celestron.com/browse-shop/uncategorized/mini-handheld-digital-microscope, 2017.

[37] M. McCauley, "Accelstepper library for ARDUINO," Online document at http://www.airspayce.com/mikem/arduino/AccelStepper/index.html, 2017.

[38] G. L. da Silva, J. Stolfi, and H. C. da Gama Leitao, "Low-cost positioner control package," Free software available at http://www.ic.uff.br/~hcgl/LCPCPakage.tgz, 2017.