High-resolution, Real-time 3D Shape Acquisition

Song Zhang and Peisen Huang
Department of Mechanical Engineering, State University of New York at Stony Brook
Stony Brook, New York, 11794-2300, USA
Email: {song.zhang, peisen.huang}sunysb.edu

Abstract

In this paper we describe a high-resolution, real-time 3D shape acquisition system based on structured light techniques. This system uses a color pattern whose RGB channels are coded with either sinusoidal or trapezoidal fringe patterns. When projected by a modified DLP projector (color filters removed), this color pattern results in three grayscale patterns projected sequentially at a frequency of 240 Hz. A high-speed B/W CCD camera synchronized with the projector captures the three images, from which the 3D shape of the object is reconstructed. A color CCD camera is also used to capture images for texture mapping. The maximum 3D shape acquisition speed is 120 Hz (532 $\times$ 500 pixels), which is high enough for capturing the 3D shapes of moving objects. Two coding methods, sinusoidal phase-shifting method and trapezoidal phase-shifting method, were tested and results with good accuracy were obtained. The trapezoidal phase-shifting algorithm also makes real-time 3D reconstruction possible.

1. Introduction

3D shape acquisition has applications in such diverse areas as computer graphics, virtual reality, medical diagnostics, robotic vision, industrial inspection, reverse engineering, etc. [1, 2, 3, 4, 5, 6]. With the recent technological advances in digital imaging, digital projection display, and personal computers (PCs), it is becoming increasingly possible for 3D shape acquisition to be done in real time.

Among all the existing ranging techniques [7], stereovision is probably the most studied method. Traditional stereovision methods estimate shape by establishing spatial correspondence of pixels in a pair of stereo images. Recently, Zhang *et al*. [8] and Davis *et al*. [9] developed a new concept called spacetime stereo, which extends the matching of stereo images into the time domain. By using both spatial and temporal appearance variations, it was shown that matching ambiguity could be reduced and ac-

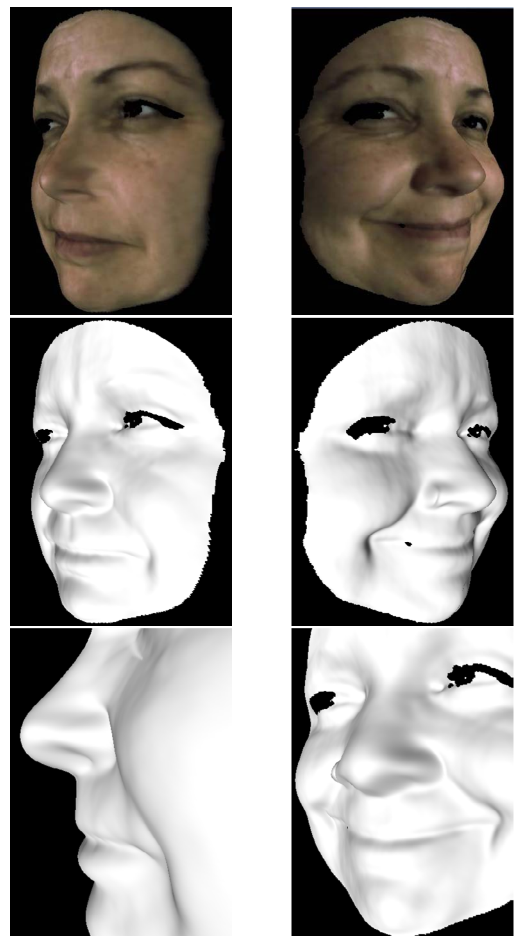


Figure 1. 3D models of a human face acquired by our real-time 3D shape acquisition system. The acquisition frame rate was 120 Hz and the resolution was 532 $\times$ 500 pixels. Two separate frames selected from a sequence of over 60 frames are displayed here. Top row: models with color texture mapping. Middle row: models in a shaded mode. Bottom row: zoom-in views of the models.

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)
1063-6919/04 $ 20.00 IEEE

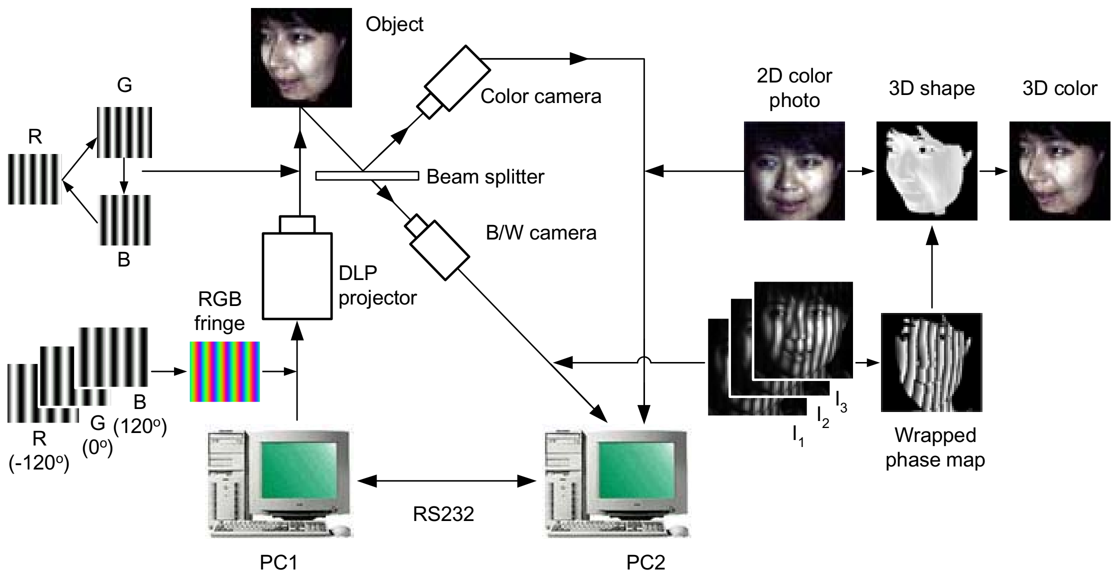


Figure 2. Schematic diagram of our real-time 3D shape acquisition system. A color fringe pattern is generated by a PC and is projected onto the object by a DLP video projector (Kodak DP900). A high-speed B/W CCD camera (Dalsa CA-D6-0512W) synchronized with the projector is used to capture the images of each color channel. Then image processing algorithms are used to reconstruct the 3D shape of the object. A color CCD camera (Uniq Vision UC-930) also synchronized with the projector and aligned with the B/W camera is used to capture color images of the object for texture mapping.

curacy could be increased. As an application, Zhang *et al.* demonstrated the feasibility of using spacetime stereo to reconstruct the shapes of dynamically changing objects [8]. The shortcoming of spacetime stereo or any other stereo vision method is that matching of stereo images is usually time-consuming. Therefore, it is difficult to reconstruct 3D shapes from stereo images in real time.

Another major group of ranging techniques is structured light, which includes various coding methods and employs varying number of coded patterns. Unlike stereo vision methods, structured light methods usually use processing algorithms that are much simpler. Therefore, it is more likely for them to achieve real-time performance (acquisition and reconstruction). For real-time shape acquisition, there are basically two approaches. The first approach is to use a single pattern, typically a color pattern. Harding proposed a color-encoded Moiré technique for high-speed 3D surface contour retrieval [10]. Geng developed a rainbow 3D camera for high-speed 3D vision [11]. Wust and Capson [12] proposed a color fringe projection method for surface topography measurement with the color fringe pattern printed on a color transparency film. Huang *et al.* [13] implemented a similar concept but with the color fringe pattern produced digitally by a digital-light-processing (DLP) projector. Zhang *et al.* developed a color structured light technique for high-speed scans of moving objects [14]. Since the above methods all use color to code the patterns, the shape acquisition result is affected to varying degrees by the variations of the object surface color. In general, the more patterns used in a structured light system, the better accuracy that can be achieved. Therefore, the above methods sacrifice accuracy for improved acquisition speeds.

The other structured light approach for real-time shape acquisition is to use multiple coded patterns but switch them rapidly so that they can be captured in a short period of time. Raskar *et al.* proposed an imperceptible structured light method for real-time depth extraction, which was based on rapidly switching binary-coded patterns [15]. Rusinkiewicz *et al.* [16] and Hall-Holf and Rusinkiewicz [17] developed a real-time 3D model acquisition system that uses four patterns coded with stripe boundary codes. The acquisition speed achieved was 60 Hz, which is good enough for scanning slowly moving objects. However, like any other binary-coding method, the spatial resolution of these methods is relatively low because the stripe width must be larger than one pixel. Moreover, switching the patterns by repeatedly loading patterns to the projector limits the switching speed of the patterns and therefore the speed of shape acquisition. Huang *et al.* recently proposed a high-speed 3D shape measurement technique based on a rapid phase-shifting technique [18]. This technique uses three phase-shifted, sinusoidal grayscale fringe patterns to provide pixel-level resolution. The patterns are projected to the object with a switching speed of 240 Hz. However, limited by the frame rate of the camera used, the acquisition speed achieved was only 16 Hz.

In this paper, we describe an improved version of the system developed by Huang *et al*. [18] for real-time 3D shape acquisition. This system takes full advantage of the single-chip DLP technology for rapid switching of three coded fringe patterns (Perlin *et al*. [19, 20] and McDowall *et al*. [21] used similar techniques for 3D display). A color fringe pattern with its red, green, and blue channels coded with three different patterns is created by a PC. When this pattern is sent to a single-chip DLP projector, the projector projects the three color channels in sequence repeatedly and rapidly. To eliminate the effect of color, color filters on the color wheel of the projector are removed. As a result, the projected fringe patterns are all in grayscale. A properly synchronized high-speed black-and-white (B/W) CCD camera is used to capture the images of each color channel from which 3D information of the object surface is retrieved. A color CCD camera, which is synchronized with the projector and aligned with the B/W camera, is also used to take 2D color pictures of the object at a frame rate of 26.8 Hz for texture mapping. With this prototype system, 3D shape acquisition can be done at a speed of up to 120 Hz in a continuous measurement mode.

So long as the number of patterns needed for 3D reconstruction is not more than three, any coding method can be used to achieve the same 3D acquisition speed. In this research, we tried two coding methods, the sinusoidal phase-shifting method and the trapezoidal phase-shifting method. For the sinusoidal phase-shifting method, we use three phase-shifted sinusoidal fringe patterns with a phase shift of 120°. Since in a phase-shifting method, the calculation of an arctangent function, which is relatively slow, is necessary to determine the phase, it is difficult to realize real-time 3D reconstruction with an ordinary PC. To increase the calculation speed, we developed a novel coding method, namely, the trapezoidal phase-shifting method. With this new coding method, it was demonstrated experimentally that high-resolution, real-time 3D shape acquisition and reconstruction were possible.

In Section 2, we discuss the working principle of the structured light system for real-time 3D shape acquisition. Section 3 describes the two coding methods, the sinusoidal phase-shifting method and the trapezoidal phase-shifting method, implemented in this research. Section 4 presents some experimental results and Section 5 discusses the advantages of the proposed system. Section 6 summarizes the results obtained in this research.

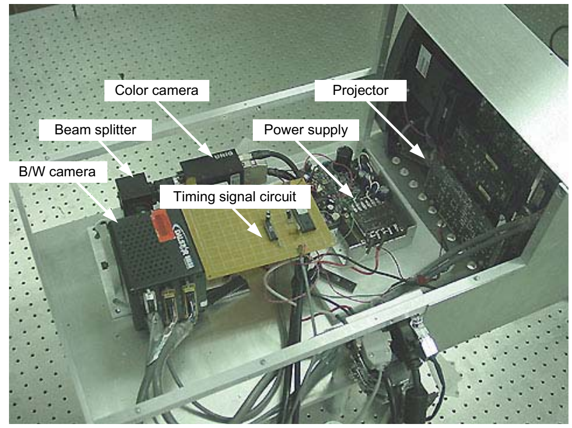


Figure 3. Photograph of our real-time 3D shape acquisition system (Box size: 24”×14”×10”).

2. Real-time structured light system

Figure 2 shows the layout of the proposed structured light system for 3D shape acquisition and Figure 3 shows a picture of the developed hardware system. For the projection of computer-generated patterns, a single-chip DLP projector is used, which produces images based on a digital light switching technique [22]. The color image is produced by projecting the red, green, and blue channels sequentially and repeatedly at a high speed. Our eyes then integrate the three color channels into a full color image. To take advantage of this unique projection mechanism of a single-chip DLP projector, we create a color pattern which is a combination of three patterns in the red, green, and blue channels respectively. In the mean time, we remove the color filters on the color wheel of the projector to make the projector operate in a monochrome mode. As a result, when the color pattern is sent to the projector, it is projected as three grayscale patterns, switching rapidly from channel to channel (240 Hz). A high-speed B/W camera, which is synchronized with the projector, is used to capture the three patterns rapidly for real-time 3D shape acquisition. An additional color camera is used to capture images for texture mapping.

2.1. Projector and camera synchronization

Figure 4 shows the timing signals of our system. The projector trigger signal is generated externally by a microcontroller-based circuit. The internal timing signal of the projector is disabled. The projection timing chart indicates the sequence and timing of the projection of the color channels. The projection cycle starts with the red channel at the falling edge of the projector trigger signal. The camera

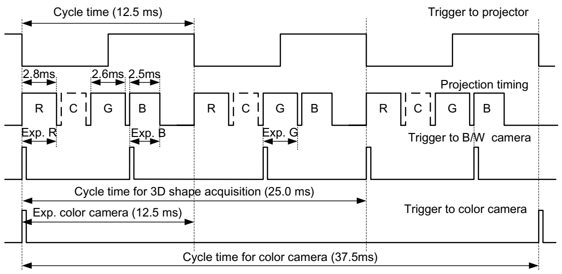


Figure 4. System timing chart. The waveform at the top is the trigger signal to the projector generated by a microcontroller based timing signal circuit. The second waveform from the top is the projection timing chart, where R, G, B, and C represent the red, green, blue, and clear channels of the projector. The clear channel is designed to enhance the brightness of the projected image. For the patterns used in this research, nothing is projected in this channel. The next two waveforms are the trigger signals to the B/W camera and the color camera respectively.

trigger signal is generated by the same circuit, which guarantees its synchronization with the projector trigger signal. Because of the limitation in the frame rate of our B/W camera (maximum 262 fps), two projection cycles are needed for the camera to capture the three phase shifted fringe patterns, which results in a frame rate of 40 Hz for 3D shape acquisition. If a higher speed camera is used, a maximum frame rate of 80 Hz can be achieved. However, since the phase relationship between any neighboring two patterns is the same, any newly captured pattern can be combined with its preceding two patterns to produce a 3D shape image. Therefore, the real frame rate can be three times higher, or 120 Hz for the current system setup and 240 Hz if a higher speed camera is used. The color picture taken by the color CCD camera (Uniq Vision UC-930) can be mapped directly onto the 3D shape once the correspondence of the two cameras is determined. To obtain 3D and color information of the object simultaneously, multi-threading programming was used to guarantee that two cameras work independently and that the timing of image grabbing is only determined by the external trigger signal.

2.2. Texture mapping

For more realistic rendering of the object surface, a color texture mapping method is used. In the sinusoidal phase-shifting method, the three fringe patterns have a phase shift of 120° between neighboring patterns. Since averaging the three fringe patterns washes out the fringes, we can obtain a color image without fringes by setting the exposure time of the color camera to one projection cycle or 12.5 ms. If the sinusoidal patterns are not truly sinusoidal due to nonlinear effects of the projector, residual fringes will exist. However, our experimental results show that such residual fringes are negligible and the image is in general good enough for texture mapping purpose.

Since aligning the two cameras perfectly is difficult, it is necessary to perform coordinate transformation in order to match the pixels of the two cameras. For this purpose, we use projective transformation, which is good enough for texture mapping in this system. That is:

$$I_{bw}(x, y) = PI_c(x, y) \tag{1}$$

where I_{bw} is the intensity of the B/W image, I_c is the intensity of the color image, and P is a 3×3 2D planar projective transformation matrix. The parameters of the matrix P depend on the system setup, which only need to be determined once through calibration. Once the coordinate relationship between the two cameras is determined, we can identify the corresponding pixel of any 3D image pixel in the color image for texture mapping.

3. Coding methods

The proposed system provides us with the capability of projecting and capturing three coded patterns rapidly. As long as three patterns are sufficient for 3D reconstruction, any coding method can be used to achieve real-time acquisition speed. In this research, two coding methods, sinusoidal phase-shifting and trapezoidal phase-shifting methods, were implemented to demonstrate the capability of the system.

3.1. Sinusoidal phase-shifting method

Sinusoidal sinusoidal phase-shifting method has been used extensively in optical metrology to measure 3D shapes of objects at various scales. In this method, a series of phase-shifted sinusoidal fringe patterns are recorded, from which the phase information at every pixel is obtained. This phase information helps determine the correspondence between the image field and the projection field. Once this correspondence is determined, the 3D coordinate information of the object can be retrieved based on triangulation.

Many different sinusoidal phase-shifting algorithms have been developed. In this research, the simple three-step algorithm is used [23], which requires three phase-shifted images. The intensities of the three images with a phase shift of 120° are as follows (See Figure 5 for the plots):

$$I_r(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) - 2\pi/3] \ , \tag{2}$$

$$I_g(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y)] \ , \tag{3}$$

$$I_b(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + 2\pi/3] \ , \tag{4}$$

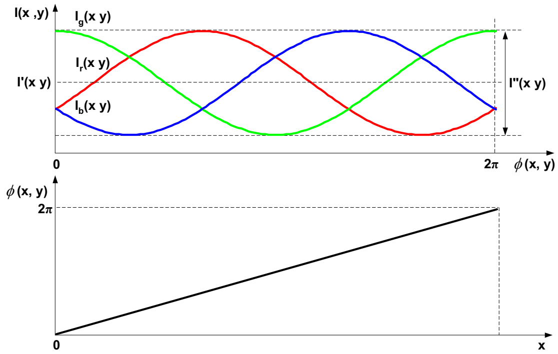


Figure 5. Sinusoidal phase-shifting method. The plot at the top is the cross section of the fringe pattern and the plot at the bottom is the calculated phase.

where $I'(x, y)$ is the average intensity, $I''(x, y)$ is the intensity modulation, and $\phi(x, y)$ is the phase to be determined. Solving these three equations simultaneously, we obtain $\phi(x, y)$:

$$\phi(x, y) = \arctan\left[\sqrt{3}(I_r - I_b)/(2I_g - I_r - I_b)\right] . \quad (5)$$

This equation provides the so-called modulo 2π phase at each pixel whose values range from 0 to 2π. By removing the 2π discontinuity of the image using a phase unwrapping algorithm, a continuous 3D phase map is obtained. This phase map can be converted to the depth map by a phase-to-height conversion algorithm based on triangulation. A simple phase-to-height conversion algorithm, which is used in this research, is described in Refs. [18, 24]. In this algorithm, surface height is considered proportional to the difference between the phase maps of the object and a flat reference plane with the scale factor determined through calibration. More complex but more accurate calibration methods are discussed in Refs. [25, 26, 27].

3.2. Trapezoidal phase-shifting method

Our ultimate goal is to build a true real-time 3D shape acquisition system in which not only the images are captured in real time, the reconstruction of 3D shape is also done in real time. This requires that the processing of the images be performed rapidly. The traditional sinusoidal phase-shifting algorithm described in the previous section works well in terms of measurement accuracy. However, due to the need of calculating an arctangent function in order to obtain the phase, the reconstruction of 3D shape is relatively slow. Therefore it is difficult to capture and process the images simultaneously in real time with an ordinary PC. Currently, we can only record the images in real time. The processing of the images has to be done off-line. In this section, we describe a newly proposed coding method, namely the trapezoidal phase-shifting method, which combines the concept of phase shifting with the intensity ratio methods for improved processing speed.

Several intensity-ratio based methods have been developed in the past, which use a very simple calculation for 3D shape reconstruction [28, 29, 30]. In these methods, two patterns are projected onto the object, a gray-level ramp and a constant illumination or a flat pattern. To eliminate the effect of nonuniform spatial illumination, Savarese *et al.* used three patterns instead of two [31]. These methods had the advantage of high processing speed but their spatial resolution was limited. To alleviate the problem, Chazan and Kiryati used a sawtooth like pattern to replace the ramp pattern [32]. The spatial resolution was improved, but as a trade-off, ambiguity was introduced which could be a problem when measuring objects with discontinuous surface shapes. The discontinuities in the sawtooth pattern may also introduce large errors when the pattern is defocused.

To improve the spatial resolution without introducing the ambiguity problem, we developed an improved method called trapezoidal phase-shifting method. This method uses three phase-shifted trapezoidal patterns which can be projected rapidly using our real-time structured light system. The concept is similar to the sinusoidal method. However, instead of phase, intensity ratio is calculated, which is much faster in terms of processing speed. Following are the intensity equations for the three color channels (See Figure 6 for the plots):

$$I_r(x, y) = \begin{cases} I''(2 - \frac{6x}{T}) + I_0 & x \in [T/6, T/3) \\ I_0 & x \in [T/3, 2T/3) \\ I''(\frac{6x}{T} - 4) + I_0 & x \in [2T/3, 5T/6) \\ I_0 + I'' & \text{Otherwise} \end{cases} , \quad (6)$$

$$I_g(x, y) = \begin{cases} I''\frac{6x}{T} + I_0 & x \in [0, T/6) \\ I_0 + I'' & x \in [T/6, T/2) \\ I''(4 - \frac{6x}{T}) + I_0 & x \in [T/2, 2T/3) \\ I_0 & x \in [2T/3, T] \end{cases} , \quad (7)$$

$$I_b(x, y) = \begin{cases} I_0 & x \in [0, T/3) \\ I''(\frac{6x}{T} - 2) + I_0 & x \in [T/3, T/2) \\ I_0 + I'' & x \in [T/2, 5T/6) \\ I''(6 - \frac{6x}{T}) + I_0 & x \in [5T/6, T] \end{cases} , \quad (8)$$

where T is the stripe width for each color channel, I_0 is the minimum intensity level, and I'' is the intensity modulation. The stripe is divided into six regions, each can be uniquely identified by the intensities of the red, green, and blue channels. For each region, we calculate the intensity ratio similar to that in the traditional intensity ratio method.

Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)
1063-6919/04 $ 20.00 IEEE

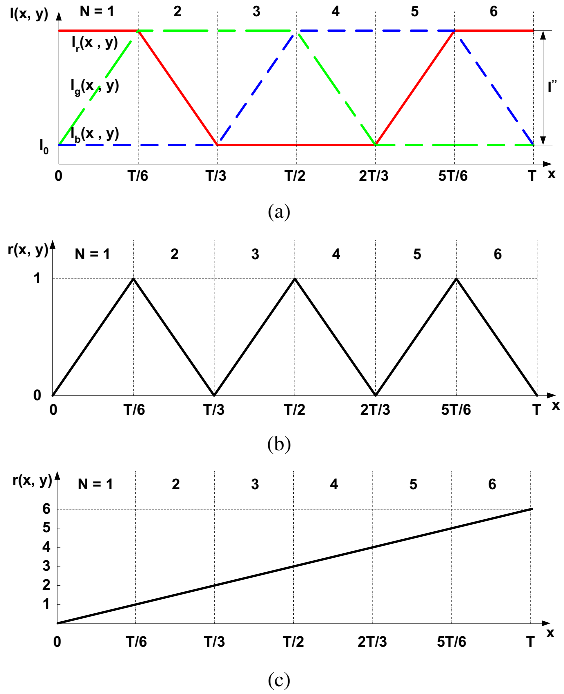


Figure 6. Trapezoidal phase-shifting method. (a) Cross section of the fringe pattern. (b) Intensity ratio in a triangular shape. (c) Intensity-ratio ramp after the removal of the triangular shape.

That is:

$$r(x, y) = \frac{I_{med}(x, y) - I_{min}(x, y)}{I_{max}(x, y) - I_{min}(x, y)} , \tag{9}$$

where $r(x, y)$ is the intensity ratio and $I_{min}(x, y)$, $I_{med}(x, y)$, and $I_{max}(x, y)$ are the minimum, median and maximum intensity values at point (x, y) respectively. $r(x, y)$ has a triangular shape whose value ranges from 0 to 1. This triangular shape can be converted to a ramp by identifying the region to which the pixel belongs using the following equation:

$$\begin{aligned} r(x, y) =& 2 \times \text{round}\left(\frac{N-1}{2}\right) \\ &+ (-1)^{N+1} \frac{I_{med}(x, y) - I_{min}(x, y)}{I_{max}(x, y) - I_{min}(x, y)} , \end{aligned} \tag{10}$$

where N is the region number. The value of $r(x, y)$ ranges from 0 to 6. 3D shape can be reconstructed by using the triangulation method as in the sinusoidal phase-shifting method.

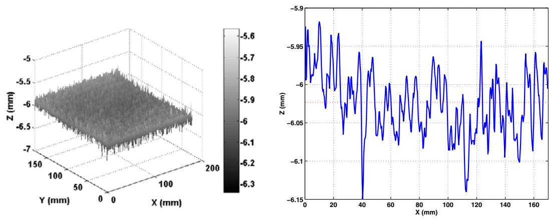
To obtain a higher spatial resolution, the pattern can be repeated. In this case, the value of the intensity ratio is periodical with a range of [0, 6]. The discontinuity can be removed by an algorithm similar to the phase unwrapping algorithm used in the sinusoidal method. The trade-off of using a repeated pattern is that potential height ambiguity is introduced.

Compared to the sinusoidal phase-shifting method, this method can process images much faster because it only needs to calculate the intensity ratio at each pixel, which is much faster than the calculation of the arctangent function (4.6 ms versus 20.8 ms on a P4 2.8GHz PC for an image size of 532×500). Compared to the previous intensity-ratio based methods, this method has a resolution that is six times higher. Even more important is the fact that the new method is much less sensitive to the blurring of the projected fringe pattern, which could occur when measuring objects with a large depth. Our simulation results show that even if the pattern becomes sinusoidal, which is the extreme of blurring, the maximum error is only 3.7%.

One disadvantage of this method, however, is in texture mapping. In the sinusoidal method, averaging the three phase-shifted fringe images cancels out the fringes and generates a grayscale image of the object. This grayscale image can be used for texture mapping, thus eliminating the need for capturing an additional image. For the trapezoidal pattern, theoretically, images for grayscale texture mapping can be obtained by choosing the maximum intensity value of the three patterns at each point. However, this approach is practically difficult because of image defocus, which results in residual lines in the image. Even with an additional color camera for texture mapping as in the system described in this paper, the new method will pose some problems because of the difficulty in eliminating the fringes in the image. In the sinusoidal method, we set the exposure time of the color camera to be equal to the projection cycle time. By doing so, we guarantee that the image captured by the color camera is the average of the three sinusoidal patterns and therefore no fringe will be visible in the image. With the trapezoidal pattern, it is obvious that the same technique can not be used. One potential solution is to capture the image when none of the color channels is projected, such as when the clear channel is projected. However, this will require a camera with a much faster sampling speed and possibly also an additional flash light. We are currently working on finding a better solution to this problem.

4. Experiments

After calibration, we first tested the measurement noise of the system. Figure 7 shows the measured results of a flat board using the sinusoidal phase-shifting algorithm. Since the surface is very smooth, the variations shown in the results are largely due to the noise of the system, which measures approximately at RMS 0.05 mm for the whole measured area (260×244 mm). We also measured the same



(a) (b)

Figure 7. Measured result of a flat board with a smooth surface using the sinusoidal phase-shifting method. The measured area is 260 $\times$ 244 mm and the noise is RMS 0.05 mm. (a) 3D plot. (b) One example cross section.

board with the trapezoidal phase-shifting method, which showed a noise level of RMS 0.055 mm for the same measured area.

We then measured human faces with both methods to test their capability in capturing 3D dynamic facial changes. Figure 8 shows the results obtained by using the sinusoidal phase-shifting method. During the acquisition, the subject was asked to smile so that facial changes were introduced. The experimental results demonstrate that the system is able to measure moving objects with good accuracy. Figure 1 shows two of the frames of the reconstructed 3D models. It can be seen that the detailed facial expression changes have been successfully captured. As an application, this type of high quality models are being used for the tracking, learning, and transferring of facial expressions [33].

We also tried using the trapezoidal phase-shifting method to measure human faces. Figure 9 shows the measured results. The quality of the 3D data is similar to that of the sinusoidal phase-shifting method. However, texture mapping could not be easily implemented.

It should be noted that in both coding methods, we used multiple fringes or stripes in the patterns to improve resolution. Therefore, the potential ambiguity problem is there [34]. However, since the human face is smooth, we did not encounter serious difficulties in 3D reconstruction.

5. Discussion

The proposed real-time 3D shape acquisition system has several advantages over systems based on other methods:

1. *High resolution*. Both of the proposed coding methods provide a pixel-level resolution, which is much higher than that of the stereo vision and binary coding based systems. The depth resolution of the proposed trapezoidal phase-shifting method is also six times better than that of the previously developed intensity ratio methods when the same number of stripes are used.

2. *High acquisition speed*. Our system takes advantage of the unique operating principle of a single-chip DLP projector to project three coded patterns repeatedly with a cycle time of 12.5 ms. If the camera can keep up with the speed, the maximum achievable 3D shape acquisition speed is 240 Hz (360 Hz if a latest-generation DLP projector, which has a shorter cycle time, is used). With the camera currently used in our system, we were able to achieve a speed of 120 Hz with grayscale texture mapping. With color texture mapping, the speed is reduced to 26.8 Hz due to the limited frame rate of the color camera used. This acquisition speed is faster than that achieved by most of the other methods proposed.

3. *High processing speed*. With the new trapezoidal phase-shifting method, the total processing time for 3D shape reconstruction was reduced to around 24.2 ms per frame (532$\times$500 pixels) with a Pentium 4, 2.8 GHz PC. Therefore, if parallel processing is used, it is possible to realize real-time 3D shape acquisition and reconstruction with an ordinary PC.

4. *Large dynamic range*. Since the sinusoidal phase-shifting method is insensitive to the defocusing of the projected pattern, a relatively large depth can be measured, provided that the fringe visibility is good enough. The trapezoidal phase-shifting method is also much less sensitive to defocusing when compared to previously introduced intensity ratio methods.

5. *Simultaneous 2D and 3D acquisition*. This is one of the unique features of the sinusoidal phase-shifting method. From the three phase-shifted fringe images, we can capture not only the 3D information of the object, but also the 2D information, which is obtained by averaging the three fringe images.

6. *Diverse coding options*. Our system provides the basic capability of rapidly projecting and capturing three coded patterns for 3D shape acquisition. In this research, we successfully demonstrated the feasibility of two coding methods, the sinusoidal phase-shifting method and the trapezoidal phase-shifting method. However, other coding methods are also possible as long as three patterns are sufficient for 3D shape reconstruction.

6 Conclusions

A high-resolution, real-time 3D shape acquisition system has been developed. The acquisition speed can be up to 120 Hz for the current setup and 240 Hz if a higher speed

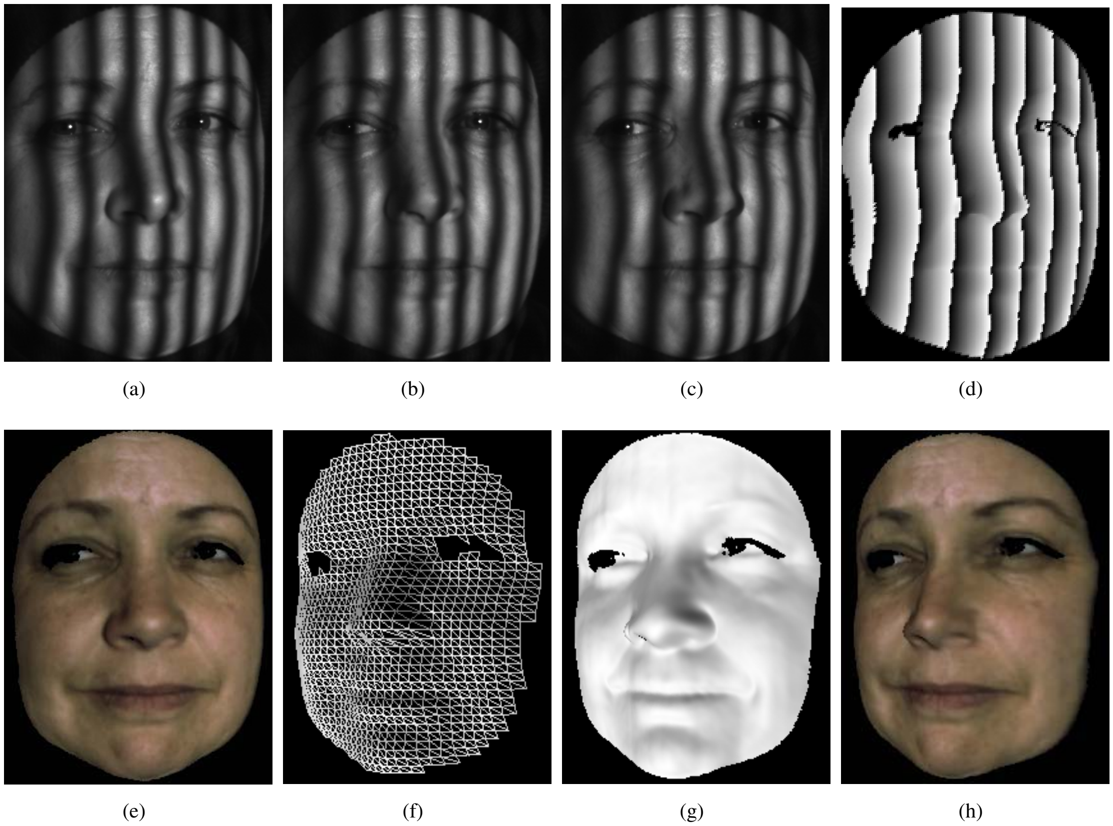


Figure 8. Shape acquisition of a human face using the sinusoidal phase-shifting method. (a)-(c) Phase-shifted fringe images. (d) Wrapped phase map. (e) 2D color image for texture mapping. (f) 3D wire-frame model. (g) 3D shaded model. (h) 3D model with color texture mapping.

camera is used. Two coding methods, the sinusoidal phase-shifting method and the trapezoidal phase-shifting method, were proposed. With the sinusoidal phase-shifting method, we were able to acquire dynamic 3D models with color texture mapping, but the processing speed was not fast enough for real-time 3D reconstruction. In contrast, the new trapezoidal phase-shifting method allowed for real-time 3D reconstruction with a similar resolution and noise level, but unfortunately texture mapping was more difficult. Both methods were tested experimentally, which demonstrated the feasibility of the methods for high-resolution, real-time 3D shape acquisition. The noise level was found to be RMS 0.05 mm in an area of 260×244 mm for the sinusoidal phase shifting method and 0.055 mm for the trapezoidal phase-shifting method. This real-time 3D shape acquisition system has potential applications in many different fields, such as computer graphics, medical diagnostics, plastic surgery, industrial inspection, reverse engineering, robotic vision, etc. If infrared light is used this system can also be used for security checks in homeland security applications.

Acknowledgments

The authors thank Susan Brennan, Xiaomei Hao, Alexander Mohr, Dimitris Samaras, Yang Wang, Lei Zhang, etc. for their willingness to serve as models for our experiments and Yang Wang and Liu Yang for their help in programming. This work was supported by the National Science Foundation under grant No. CMS-9900337 and the National Institute of Health under grant No. RR13995.

References

[1] D. Beymer and K. Konolige, "Real-time tracking of multiple people using continuous detection", in *IEEE Frame Rate Workshop*, 1999.

[2] S. B. Gokturk, J.-Y. Bouguet, C. Tomasi, and B. Girod, "Model based face tracking for view-independent fa-

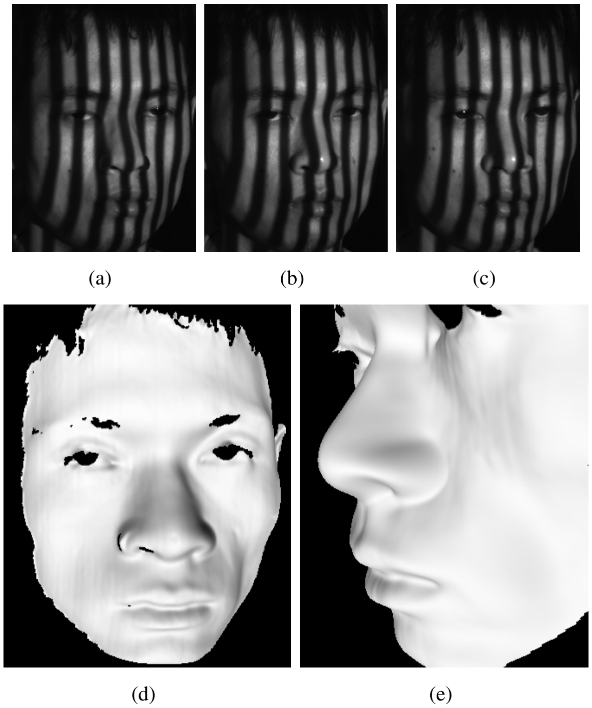


Figure 9. Shape acquisition of a human face using the trapezoidal phase-shifting method. (a)-(c) Captured fringe images. (d)-(e) Reconstructed 3D model. Texture mapping is difficult in this case.

cial expression recognition", in *The 5th International Conference On Automatic Face and Gesture Recognition*, 2002, pp. 20–21.

[3] C. Tomasi, J. Zhang, and D. Redkey, "Experiments with a real-time structured-from-motion system", in *4th International Symposium On Experimental Robotics*, 1995, pp. 123–128.

[4] C. Tomasi, A. Rafii, and I. Torunoglu, "Full-size projection keyboard for handheld device", in *Commun. ACM*, 2003, vol. 46, pp. 70–75.

[5] J. Neumann, C. Fermuller, Y. Aloimonos, and V. Brajovic, "Compound eye sensor for 3d ego motion estimation", in *IEEE International Conference on Robotics and Automation*, May 2004.

[6] H. Jin, P. Favaro, and S. Soatto, "Real-time feature tracking and outlier rejection with changes in illumination", in *The 8th IEEE International Conference on Computer Vision*, 2001, pp. I: 684–689.

[7] J. Salvi, J. Pagés, and J. Batlle, "Pattern codification strategies in structured light systems", Tech. Rep., Insititut d'Informática i Aplicacions, Girona, Spain, 2002.

[8] L. Zhang, B. Curless, and S.M. Seitz, "Spacetime stereo: Shape recovery for dynamic scenes", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2003, pp. II: 367–374.

[9] J. Davis, R. Ramamoorthi, and S. Rusinkiewicz, "Spacetime stereo: a unifying framework for depth from triangulation", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2003, pp. II: 359–366.

[10] Kevin G. Harding, "Phase grating use for slop discrimination in moiré contouring", in *Proc. SPIE*, 1991, vol. 1614 of *Optics, illumination, and image sensing for machine vision VI*, pp. 265–270.

[11] Z. J. Geng, "Rainbow 3-d camera: New concept of high-speed three vision system", *Opt. Eng.*, vol. 35, pp. 376–383, 1996.

[12] C. Wust and D. W. Capson, "Surface profile measurement using color fringe projection", *Machine Vision and Applications*, vol. 4, pp. 193–203, 1991.

[13] P. S. Huang, Q. Hu, F. Jin, and F. P. Chiang, "Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring", *Opt. Eng.*, vol. 38, pp. 1065–1071, 1999.

[14] Li Zhang, Brian Curless, and Steven M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming", in *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, June 2002, pp. 24–36.

[15] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs, "The office of the future: A unified approach to image-based modeling and spatially immersive displays", *Computer Graphics*, vol. 32, no. Annual Conference Series, pp. 179–188, 1998.

[16] S. Rusinkiewicz, O. Hall-Holt, and Levoy Marc, "Real-time 3d model acquisition", in *SIGGRAPH*. 2002, vol. 1281 of *3D acquisition and image based rendering*, pp. 438 – 446, ACM Press.

[17] O. Hall-Holt and S. Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects", in *The 8th IEEE International Conference on Computer Vision*, 2001, pp. II: 359–366.

[18] P. S. Huang, C. Zhang, and F. P. Chiang, "High-speed 3-d shape measurement based on digital fringe projection", *Opt. Eng.*, vol. 42, no. 1, pp. 163–168, 2003.

[19] K. Perlin, S. Paxia, and J. Kollin, "An autostereoscopic display", in *SIGGRAPH*. 2000, pp. 319–326, ACM Press.

[20] K. Perlin, C. Poultney, J. Kollin, D. Kristjansson, and S. Paxia, "Recent advances in the nyu autostereoscopic display", in *Proc. of the SPIE*, Jan. 2001, pp. 196–203.

[21] Ian McDowall, Mark Bolas, Dan Corr, and Terry Schmidt, "Single and multiple viewer stereo with dlp projectors", in *Proc. of the SPIE*, Jan. 2001, pp. 418–425.

[22] L. J. Hornbeck, "Digital light processing and mems: Timely convergence for a bright future (invited plenary paper)", in *Proc. SPIE*, 1995, vol. 2639, pp. 3–21.

[23] D. Malacara, Ed., *Optical Shop Testing*, John Wiley and Songs, NY, 1992.

[24] C. Zhang, P. S. Huang, and F. P. Chiang, "Microscopic phase-shifting profilometry based on digital micromirror device technology", *Appl. Opt.*, vol. 48, no. 8, pp. 5896–5904, 2002.

[25] Jean-Yves Bouguet, *Visual Method for Three-Dimensional Modelling*, PhD thesis, California Insititute of Technology, Pasadena, CA, 1999.

[26] Q. Hu, P. S. Huang, Q. Fu, , and F. P. Chiang, "Calibration of a 3-d shape measurement system", *Opt. Eng.*, vol. 42, no. 2, pp. 487–493, 2003.

[27] Qingying Hu, *3-D Shape Measurement Based on Digital Fringe Projection and Phase-Shifting Techniques*, PhD thesis, State University of New York at Stony Brook, Stony Brook, NY, 11794, 2001.

[28] B. Carrihill and R. Hummel, "Experiments with the intensity ratio depth sensor", *Computer Vision, Graphics and Image Processing*, vol. 32, pp. 337–358, 1985.

[29] T. Miyasaka, K. Kuroda, M. Hirose, and K. Araki, "Reconstruction of realistic 3d surface model and 3d animation from range images obtained by real time 3d measurement system", in *IEEE 15th International Conference on Pattern Recognition*, Sep. 2000, pp. 594–598.

[30] T. Miyasaka, K. Kuroda, M. Hirose, and K. Araki, "High speed 3-d measurement system using incoherent light source for human performance analysis", in *XIXth Congress of the international society for photogrammetry and remote sensing*, July 2000, pp. 16–23.

[31] S. Savarese, JY. Bouguet, and P. Perona, "3d depth recovery with grayscale structured lighting", Tech. Rep., Computer Vision, California Insititute of Technology, 1998.

[32] Gilia Chazan and Nahum Kiryati, "Pyramidal intensity-ratio depth sensor", Tech. Rep. No. 121, Israel Institute of Technology, Technion, Haifa, Israel, Oct. 1995.

[33] Yang Wang, Xiaolei Huang, Chan-Su Lee, Song Zhang, Zhiguo Li, Dimitris Samaras, Dimitris Metaxas, Ahmed Elgammal, and Peisen Huang, "High resolution acquisition, learning and transfer of dynamic 3d facial expressions", in *Eurographics 2004 (will appear)*, 2004.

[34] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light", in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2003, pp. 195–202.