High-speed pattern projection for three-dimensional shape measurement using laser speckles

Martin Schaffer, Marcus Grosse,* and Richard Kowarschik

Institute of Applied Optics, University of Jena, Froebelstieg 1, 07743 Jena, Germany

*Corresponding author: martin.schaffer@uni-jena.de

Received 19 March 2010; revised 28 April 2010; accepted 6 May 2010;
posted 18 May 2010 (Doc. ID 124803); published 18 June 2010

We propose a high-speed projection system that is able to project statistical speckle patterns at a rate of 500 Hz. Its purpose is to generate structured light for a real-time photogrammetry stereo vision setup. As conventional digital light projector (DLP) projection setups are limited in their maximum projection rate to 250 Hz for gray-value patterns, stripe projection systems are usually applied for real-time three-dimensional (3D) measurements. However, these techniques can only be used on steady surfaces as phase unwrapping has to be done. In contrast, the proposed setup is able to measure the shape of multiple spatially separated objects at once. We compare the speckle setup with a system using a DLP projector and with other fast 3D shape measurement setups, like the widely used stripe projection methods, qualitatively and quantitatively. © 2010 Optical Society of America

OCIS codes: 120.4630, 150.2950, 110.6880.

1. Introduction

Recently there has been growing interest in fast real-time three-dimensional (3D)-shape/scene measurements [1]. This development has been driven by several improvements and applications. The increased camera-to-computer bandwidth of modern industry interface standards created affordable and easy to access high-speed cameras (acquisition rate >100 fps). Additionally, the multicore architectures of current computer systems can be effectively exploited by many methods, and the parallel computing power of modern graphic processing units (GPUs) enable real-time processing by means of simple interfacing languages, such as OpenCL and CUDA. Several groups have shown tremendous speed-ups by adapting standard methods to GPUs [1–4].

Commercial real-time 3D measurement systems on the basis of time of flight (TOF) are already available, but they remain inconvenient in respect to point density and accuracy. Many techniques for real-time 3D measurements are based on stereo-photogrammetry and pattern matching. Methods without illumination, which analyze image features inherited from the object or scene structure itself, exist. Such methods offer a very short measurement time at the cost of limited accuracy. They may have reduced 3D point density and require objects or scenes that are cooperative. To overcome these limitations, so-called coded-light techniques [5] are available. The object/scene is illuminated with structured light to create textures in the resulting images. Several methods exist that only project one light pattern onto the object [6–10]. These methods can be applied to a broad class of objects. The measurement accuracy improves if more patterns are projected and captured consecutively for a 3D measurement [11]. A commonly used technique is digital fringe projection. The fringe projection setups proposed for real-time measurements are not able to measure spatially separated objects, as they are only projecting stripes and omit the usual gray-code sequence in order to realize a fast measurement time. A phase-unwrapping algorithm has to be applied in order to acquire a steady phase map. Phase unwrapping will yield incorrect phases at object discontinuities.

Furthermore, all structured light techniques are limited by the available projection speed. Systems

0003-6935/10/183622-08$15.00/0
© 2010 Optical Society of America

were set up with many efforts using conventional DLP projectors as 180 Hz projection systems [2]. As the projection limit is given by the current projection hardware, it cannot be easily addressed, therefore, it is critical to decrease the number of images for real-time measurements. Systems using only a single image or two images have been proposed. Takeda and Mutoh [10] proposed a single image 3D shape measurement system, but the accuracy was low for objects with varying surface reflectivity. By adding a homogeneously illuminated image, Guo and Huang [12] were able to suppress this effect. However, as already stated by Zhang [1], "It seems that the ultimate challenge is the fringe projection speed that is mainly limited by the current hardware technology," the projection system is the bottleneck for real-time applications.

Another major problem arises due to object movement while the patterns are projected and images are captured. In a static scene, a pixel of the image sequence of one camera always represents the same area of the object. In a dynamic scene, a pixel represents different parts of the object. Depending on measurement time and object velocity, this will reduce the accuracy or will hinder the 3D reconstruction of parts of the object/scene completely (depending on the coded-light technique used). Thus, the motion of the object has to be compensated [2] or the technique will be limited to a certain object speed.

The structured light method used for the proposed fast projection system is based on a combination of the temporal correlation technique (TCT) [13] and the areal correlation technique (ACT) commonly used for stereo matching. In [14], band-limited statistical patterns were used in combination with the TCT. Siebert and Marshall [15] projected speckle patterns via a standard projector and found stereo matches by the ACT. Instead of using a projector, a laser and a diffuser were taken to project speckle patterns onto the object [16], which were then analyzed with the ACT; however, it was not a setup for fast speckle-pattern projection. Furthermore, they generated only sparse correspondences and had huge computational costs. In [17], it is shown that a combination of the TCT and the ACT reduces the number of band-limited patterns necessary for stereo matching compared to the TCT alone.

In this paper, we present a speckle-pattern projection system that is able to project up to 500 different patterns per second. In combination with high-speed cameras, we are able to collect 207 stereo image pairs of an object with superimposed structured light per second. From this data, 17 independent 3D reconstructions per second can be generated (sequence length of 12 images per reconstruction). Additionally, by using a sliding window, a new 3D object state can be calculated for every new stereo image pair taken, so we can generate up to 207 different measured 3D states of the scene/object per second. Furthermore, our system is able to reconstruct complex nonsteady surfaces, as well as multiple spatially separated objects.

The paper is structured as follows. Section 2 explains the TCT and ACT techniques. In Section 3, the proposed setup and a reference setup are introduced, as well as the calibration method used for the stereo setup. In Section 4, we compare the performance of the proposed setup and the reference setup. Finally, in Section 5, we compare the proposed setup with other real-time setups and discuss advantages and disadvantages of existing methods.

2. Stereo Matching by Use of Correlation Techniques

A common method to find stereo correspondences is the correlation of subimages of the stereo view. In one view of the object, a rectangular window centered at a pixel (i,j) of the size $(2n_x+1)(2n_y+1)$ is chosen, with window length $2n_x+1$ and height $2n_y+1$ in pixels. This pixel area is now correlated with all possible windows of the size $(2n_x+1)(2n_y+1)$ of a second view. Several different ACTs exist. Although it has higher computational demand than other approaches, the normalized cross correlation is more robust and easier to handle [18] and, therefore, it is used in our setups. For all pixel pairs (i,j) and (i',j') of two views, the correlation can be calculated according to the normalized cross correlation to

$$\rho_{(i,j,i',j')} = \frac{\sum_{l=-n_x}^{n_x}\sum_{k=-n_y}^{n_y}(g(i+l,j+k)-\bar{g})(g'(i'+l,j'+k)-\bar{g'})}{\sqrt{\sum_{l=-n_x}^{n_x}\sum_{k=-n_y}^{n_y}(g(i+l,j+k)-\bar{g})^2}\sqrt{\sum_{l=-n_x}^{n_x}\sum_{k=-n_y}^{n_y}(g'(i'+l,j'+k)-\bar{g'})^2}}, \quad (1)$$

where $g(x,y)$ is the gray value of pixel (x,y) in one view and $\bar{g}$ is the average gray value of the window. $g'(x',y')$ and $\bar{g'}$ are the respective values of the window in the second view. The corresponding pixel of pixel (i,j) in the first view is the pixel (i',j') in the second view, which yields the highest correlation and surpasses a certain threshold. In the case of an untextured object, this method will fail because, for every window of the first view, there will be multiple high-correlating windows in the second view. This will result in reduced accuracy and false correspondences. If the same object is illuminated with

structured light, the accuracy of this method increases dramatically as an object texture is artificially introduced into the views. Depending on the chosen window size, this method will deliver bad or no correspondences, even with structured light, as every view shows a different perspective of the object. Therefore, the seen object surface of a rectangular window of one view will not map into a rectangular window in the second view. This problem can be solved by rectifying the image pair. Usually, ACT delivers inaccurate correspondences for borders of objects [19]. To minimize this influence, the correlation area should be chosen to be as small as possible.

In [13] the correlation area was reduced to one pixel, although Eq. (1) is not applicable in this case. As a sequence of N images per view is acquired, the correlation is not done via spatial correlation windows. Instead, a temporal correlation window is used. For every pixel of one view, the temporal gray-value sequence $g(i,j,n)$, where $n \in [1,N]$, is correlated with every gray-value sequence $g'(i',j',n)$ of the second view:

$$\rho_{(i,j,i',j')} = \frac{\sum_{t=1}^{N}(g(i,j,t)-\bar{g})(g'(i',j',t)-\bar{g'})}{\sqrt{\sum_{t=1}^{N}(g(i,j,t)-\bar{g})^2}\sqrt{\sum_{t=1}^{N}(g'(i',j',t)-\bar{g'})^2}}, \quad (2)$$

where $\bar{g}$ and $\bar{g}'$ are now the average gray values of the temporal gray-value sequence. As a constant temporal gray-value sequence will correlate with $\rho = 0$, this approach assumes a varying illumination over time. The best results are achieved if a sequence of structured light patterns is projected onto the object. Several different illumination patterns were used in combination with the TCT [14]. Recently, there has been increasing interest in combining temporal and areal image information, so-called space–time stereo, to solve the correspondence problem [20]. Using space–time stereo in combination with statistical patterns, it is possible to reduce the length of the image sequence compared to the TCT alone [17]. Depending on the desired accuracy and complexity of the object, at least six statistical patterns were necessary to yield a dense and outlier-free reconstruction with our reference setup. However, for higher accuracy, the sequence length has to be increased to 20 images. As the TCT is flexible in the length of the image sequence, as well as in the start and end image of a single sequence, a sliding temporal window technique can be used to generate a new 3D reconstruction for every new frame captured.

Although the correlation techniques are computationally more demanding than phase retrieval, it is possible to realize real-time applications using GPUs [21]. In this paper, we use the TCT to solve the correspondence problem.

3. Reference Coded-Light Setup Based on DLP Projector and Proposed Setup Using Speckle Patterns

A. Reference Coded-Light Setup

Because we propose a new projection system, it is necessary to compare its performance, e.g., measurement accuracy, sequence length for TCT, and projection rate, to an existing setup. At our institute, a stereo setup (see Fig. 1) using the TCT based on [13] was realized using band-limited statistical patterns (BLP) [14,22]. About 20 BLP (see Fig. 2) are projected onto the object and synchronously captured with a stereo setup. For every pixel of one view, the corresponding pixel in the second view is found by correlating the temporal intensity sequence with every pixel on the epipolar line of the second view. The pixel pair yielding the highest correlation and reaching a given threshold is accepted as correspondent. The system can be used as a hybrid space–time stereo system [17]. This reference system uses a 60 Hz DLP with 1024 × 768 pixels. Therefore, the projection of 20 statistical patterns takes at least 0.33 s. Practically, we cannot surpass 0.66 s as image formation takes 16 ms and every single capturing process has to wait until the initial image formation has finished. Although special DLP hardware kits are available that offer a projection rate of up to 250 gray-value patterns per second, the cost of such a system would hinder many potential applications. The reference system is able to measure height deviations of about 40 μm and has a noise level of about 70 μm using two AVT Pike 640 × 480 CCD cameras, which are looking at an object at a 1.2 m distance. The measurement volume is about 25 cm × 30 cm × 30 cm. The intrinsic parameters are derived using the Zhang calibration [23] and the extrinsic parameters are derived by using the eight-point algorithm [24]. Figure 3 shows the results of two face measurements, as well as an arbitrary object, with optical properties different from those of than human skin. The color information was added by consecutively projecting a homogeneous red, green, and blue illumination via the DLP and synchronous capturing with both cameras. Therefore, the color values perfectly match every point, and no interpolation or mapping is needed.

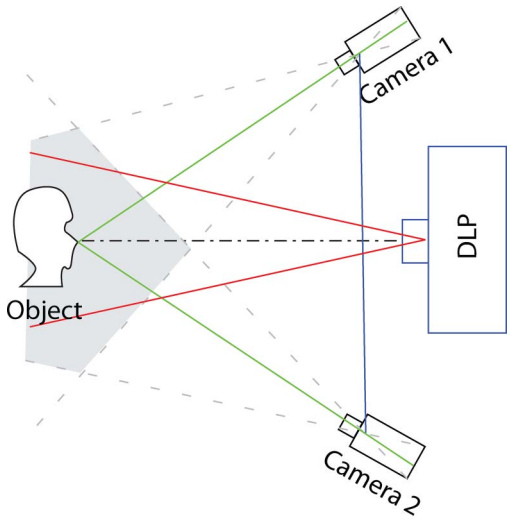

Fig. 1. (Color online) Reference stereo setup.

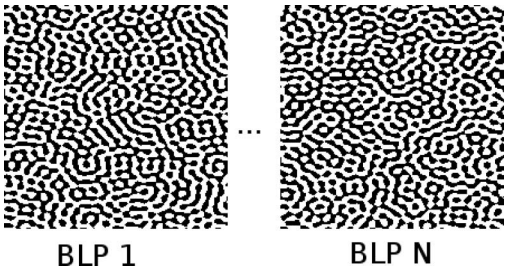

Fig. 2. Two BLP out of a sequence of N BLP patterns, which are projected onto the object per measurement via the DLP.

B. Speckle Projection Setup

As we apply statistical patterns in our reference setup, it was our idea to use laser-induced speckles instead of DLP projected patterns. The setup is shown in Fig. 4. We employ a 200 mW Nd:YAG laser emitting coherent laser light at $\lambda = 532$ nm. The beam is focused onto the diffuser using a conventional lens ($f = 7.5$ cm). The interference of the scattered waves at the diffuser surface creates an objective speckle field, which is then imaged by a second lens onto the object to be measured.

The statistical intensity distribution leads to patterns that are similar to the band-limited patterns of the reference setup; nevertheless, the structure of speckle patterns is naturally given in such a way that the size is the only variable parameter. There are several ways to change the appearance (speckle size) of the speckles. Initially, their size depends on the focal laser spot size and the grain number of the diffuser. The focal laser spot radius can be calculated by

$$L = 0.61\frac{\lambda f}{A}, \tag{3}$$

with aperture A, focal length f and wavelength λ. Hence the objective speckle size is given by [25]

$$s = \frac{\lambda z}{2\,L} \tag{4}$$

where z stands for the free-space propagation length marked in Fig. 4 and s is the speckle radius. In our setup, the aperture of 2 mm is given by the beam diameter. The speckle size of the pattern can be calculated by the FWHM of the autocorrelation function as proposed by [25] and amounts to ≈1.5 mm. To adjust the speckle size, the focal length of the lens is changed, which, in turn, changes the focal spot size of the laser beam. The speckle size is optimized so that it is similar to the size of the band-limited statistical patterns of the reference setup.

Another initial important factor of the setup is the diffuser itself. To project statistically independent patterns at a high rate, it is necessary that small diffuser displacements lead to different patterns. Therefore, the goal is to realize a small area of scattering in the axial as well as in the lateral direction. The diffuser's axial positioning is done accurately, because, otherwise, disturbing speckle movement effects can occur. To limit the scattering area axially, a ground-glass diffuser is used that has no impact on the speckles by, e.g., refractive index variations in the glass substrate like opal diffusers. Depending on the axial positioning of the diffuser, it is possible to switch between speckle movements or speckle boiling. Speckle boiling can be enabled by placing the scattering surface directly into the focal plane so that the beam propagates nearly collimated through the scattering area. In the case of speckle boiling, the speckle field is very sensitive to a diffuser displacement and all speckle patterns are different, which guarantees the best correlation results. In this case, the speckle size is usually larger due to a smaller illuminated area, which causes object points located near each other to have the same intensity. The advantage of placing the diffuser before or behind the focal plane is a smaller speckle size. However, due to convergent or divergent beam propagation, the speckles do not change sensitively but rather move against (divergent) or with (convergent) the moving direction of the diffuser, which leads to worse correspondences. At the moment, we are using an additional telescopic setup to decrease the beam diameter/aperture for larger illuminated areas in the focal spot and to get highly sensitive (speckle boiling) speckles of the right size.

Another challenge was to create a setup that can shift the diffuser quickly in a step and settle regime. By using a goniometrical stepping motor with high acceleration that is able to start and stop quickly

Fig. 3. (Color online) Results of two 3D measurements of human faces using the reference setup.

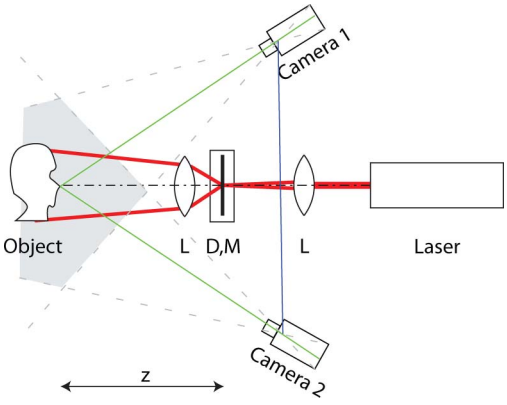

Fig. 4. (Color online) Setup consisting of laser, diffuser (D), step motor (M), lens (L), and object.

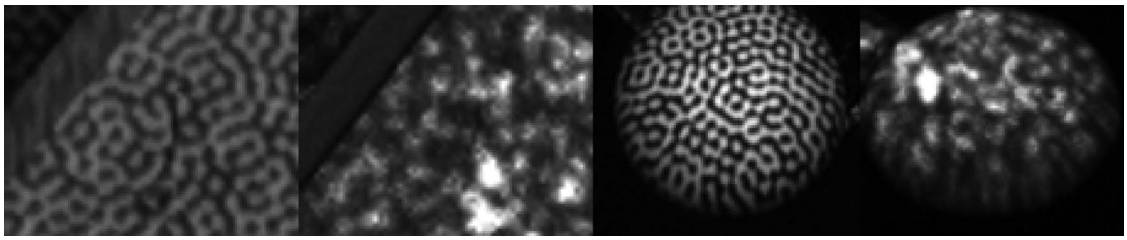

Fig. 5. Images from left to right: plane with DLP-pattern, plane with speckle pattern, sphere ($r = 1.52$ cm) with DLP pattern, and sphere with speckle pattern.

(1.7 ms/step($1° \approx 1$ mm)), nearly 600 different statistical patterns can be created per second. Higher projection rates are now limited by the displacement system used.

4. Experimental Results and Comparison

A. Qualitative

Figure 5 shows a comparison of the projected patterns realized with the DLP reference setup to the ones created by the speckle system. Although the structure size is similar, the reference system delivers a better global contrast. The speckle patterns do not address the full dynamic range of the cameras at most pixels. Although bright DLP projectors are available, it will be a future challenge to decrease exposure times in a high-speed scheme. By using laser-induced speckles, the purchase of expensive high-sensitivity cameras can be avoided and brighter images can be created by simply taking more powerful lasers. By using a 2 W laser, we have realized bright images of a sample plane with a shutter time of 140 μs. As continuous laser systems with a light output of 15 W are commercially available, even uncooperative surfaces can be imaged using shutter times in the submillisecond time regime. Another disadvantage of DLP projectors in time-limited schemes is the image formation time during one projection. Frame buffers are filled sequentially and lead to a sequentially constructed image. As image formation in the proposed speckle pattern setup is realized with the speed of light, cameras can start the capturing process immediately after the shift of the diffuser is done. Furthermore, due to the fact that we did not disassemble the color wheel of our projector in the reference setup, as others did [1], a disturbing beat frequency between the projector illumination and the frame rate of the cameras was recognizable. The unequally illuminated images led to poorer reconstruction and to noisier point clouds. Lasers do overcome this problem. In previous work [26], we have shown that, due to the overlap of the discrete pixel matrices, the correspondence assignation is locally hindered in a moirélike fashion. Although this effect is dominated by the overlapping matrices of the stereo camera system, the discrete digital mirror device matrix of the DLP projector system worsens the effect. As there is no discrete pixel matrix of the projection in the case of the speckle projection system, it will perform better in terms of overall number of correspondences. Moreover, the depth of field in a classical optical system can limit the depth of the measurement and makes the system inappropriate for depth measurements in a large volume. Because of the statistically coherent phenomenon of speckles, the speckle field offers good contrast and sharp transitions over the full field of illumination. Another advantage of the proposed speckle pattern setup can be seen in its applications. Often 360° 3D shape measurements are desired. Usually such systems are implemented with multiple setups and the resulting point clouds are matched afterward. Lasers can be switched with fiber optical switches or split in multiple fibers to arrange a multiangle speckle illumination with just one laser.

The results of real-time measurements of a human face using speckle projection are shown in Fig. 6. The raw 3D point data is shown and no outlier removal or postprocessing is used. The stereo images were taken at a rate of 180 Hz. We used a sliding window size of four images; therefore, we are able to extract 45 3D states of the face for every second of measured data.

Fig. 6. Reconstructed 3D time sequence of a human face in motion.

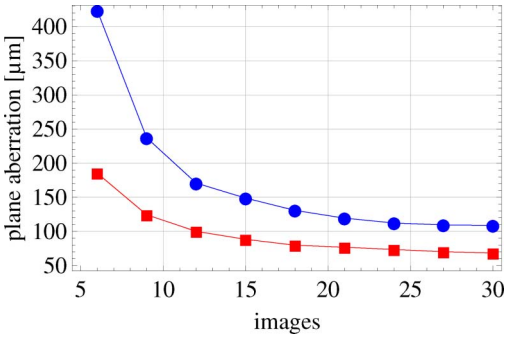

Fig. 7. (Color online) Standard deviation of measured 3D points to a fitted plane (dot, speckle; square, DLP) for different sequence lengths.

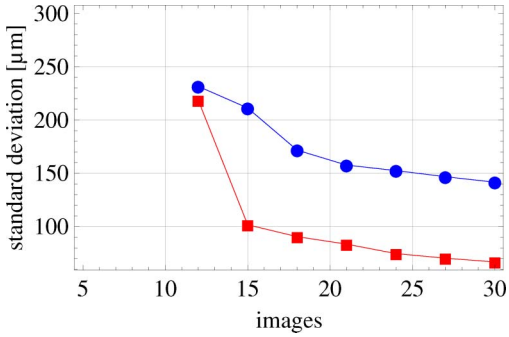

Fig. 9. (Color online) Standard deviation of measured 3D points to a fitted sphere (dot, speckle; square, DLP) for different sequence lengths.

At the moment, the reconstruction is done offline, at a rate of one 3D reconstruction per second.

B. Quantitative

To compare the speckle system to the reference setup, two measurements are conducted. For the first test, a plane sample (deviation of the plane is smaller than 10 μm over 10 cm diameter) is used. The standard deviation σ_P of measured points to a fitted plane are calculated. For the second test, a spherical sample is used (radius deviation smaller than 1 μm, sphere radius is 1.5002 cm). The standard deviation σ_S of the measured distances of all points to the center of the fitted sphere are calculated. In both tests, the conditions are kept constant (stereo setup geometry, aperture, focus, calibration data). Only the illumination source is changed and, therefore, the shutter time and frame rate of the cameras are adapted to the source. For the stereo setup, we use two AVT Pike CCD cameras with a resolution of 640×480, a pixel size of 7.4 μm, and a 17 mm lens. The baseline distance between the two cameras is 0.57 m, and both cameras are directed convergently at the objects that are 1.2 m away. The diagonal length of the measurement volume is 0.3 m. The DLP projector has a resolution of 1024×768 pixels. The results are shown in Figs. 7–9 and 11. Each graph shows the comparison between DLP and the speckle system in dependence of the number of images used for the TCT. The noise of the measured points to a given surface is higher using the speckle system compared to the DLP-system (see Fig. 7). As the pattern structure used is the only difference, the worse correspondences arise from this change. The DLP patterns deliver more homogeneous and higher contrast patterns, contain fewer high-frequency components, as they are band limited, and they illuminate the object with a constant intensity distribution over the entire area. The negative exponential intensity distribution of speckle patterns [25] causes more inaccurate point correspondences due to higher probabilities for low-intensity speckle structures. As the contrast of the speckle patterns is not as high as with the DLP system, the subpixel interpolation will deliver worse results with the speckle system. This increases the noise of the point-cloud measurement. If lasers with higher power for brighter illumination are available, it is possible to apply a spatial filter to realize a cut-off frequency in the speckle pattern. Furthermore, the speckle pattern could be optimized by using different diffusers for more homogeneous illumination. Another improvement can be achieved by using other microlenses, which create a smaller laser spot on the diffuser.

A sequence length of 12 images is an adequate compromise between overall measurement time and measurement accuracy (see Figs. 7 and 8). Almost 97% of all correspondences found are correct, and total noise is lower than 220 μm for a sphere; respectively, 170 μm for a plane measurement. In Fig. 10 one can see that there are holes in the reconstructed plane, which are caused by the moiré pattern of the two discrete camera matrices. This effect is reduced when using speckle patterns, as the DLP patterns form another discrete matrix on the object, which

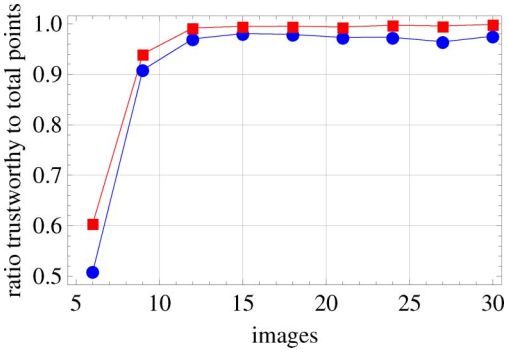

Fig. 8. (Color online) Ratio of correct correspondences to total correspondences of the plane measurement (dot, speckle; square, DLP).

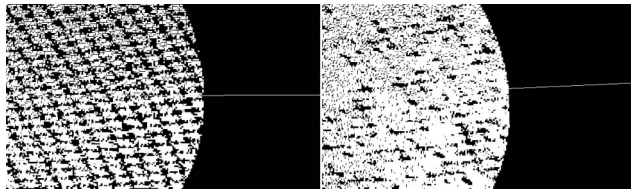

Fig. 10. Reconstruction of the sample plane using a sequence length of 30 images, only the illumination source was changed between measurements; left, DLP system; right, speckle system.

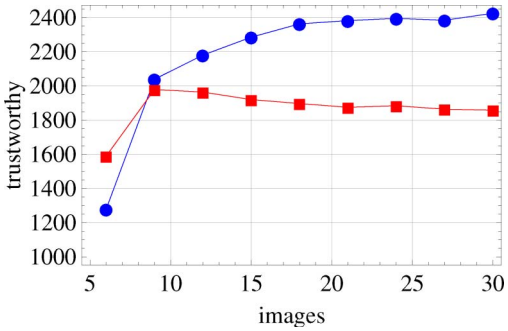

Fig. 11. (Color online) Total number of correct correspondences (dot, speckle; square, DLP).

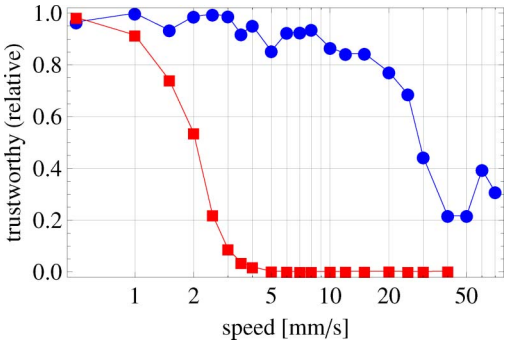

Fig. 12. (Color online) Relative amount of trustworthy points for different plane speed (dot, speckle; square, DLP).

increases the effect. By using the speckle system, between 5% and 25% correct correspondences more are found compared to the DLP system (see Fig. 11).

To test the setups for their ability to measure moving objects, a plane was moved away from the cameras with a constant speed using a linear unit in a vertical direction to the baseline of both cameras. During the movement, 21 statistical patterns were projected using the reference DLP system and the proposed speckle system. The measurement results of 21 acquired and processed images are shown in Figs. 12 and 13. As expected, the noise of the measured 3D points correlates with the speed of the object. The average measurement time was 1119 ms for the DLP system and 102.5 ms for the speckle system. A speed-up of ≈10 times could be obtained with the speckle system, which is evident by the velocity where the number of trustworthy points halves. With the speckle setup it is possible to measure objects/

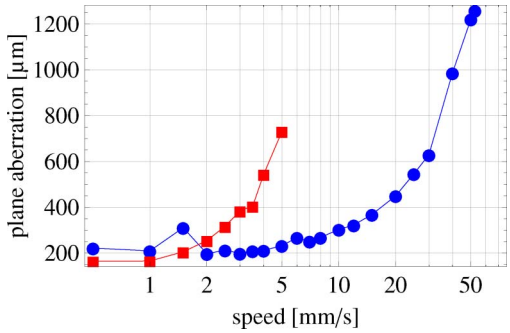

Fig. 13. (Color online) Plane aberration of the reconstructed plane for different speed (dot, speckle; square, DLP).

scene with a vertical speed of up to ≈20 mm/s. If only 12 images are used for reconstruction, the maximum tolerable speed almost doubles, to ≈40 mm/s.

5. Comparison with Other Real-Time Setups

Time-of-flight cameras realize a 3D scene reconstruction rate of up to 60 Hz and no processing on a CPU is necessary. At a distance of about 1 m, they have a noise value within the millimeter regime. Standard stereo view approaches using ACT employ high-resolution (>10 megapixel) cameras and are able to realize a noise level of about 0.1 mm; however, they have huge computational cost, mostly low point densities, and are not able to process the data for an individual 3D reconstruction within a time frame of seconds.

In the introduction, we discussed several coded-light approaches. Unfortunately, most papers do not publish measurement accuracies of their setups. In Table 1, we compare the performance of four different 3D measurement systems—our DLP reference setup (DLP 30 fps), the presented speckle setup (Speckle 207 fps), the setup presented in [27] (Stripe 120 fps) and the values of a TOF system [28]. All these systems were applied to a similar measurement volume and camera resolution (TOF has ≈1/10 of the resolution). The measurement accuracy of the speckle system is comparable to the stripe projection system, but its measurement time is about twice as long, due to the longer image sequence needed. As our current projection limit is about 500 Hz, we should be able to reach an equal measurement time in future work. At present we can only calculate one 3D reconstruction per second after the measurement data has been

Table 1. Comparison of Uncertainty of Measurement for Four Setups[a]

System	σ_P (μm) Plane	σ_S (μm) Sphere	3D data/s	3D recons/s	t_m (ms)	N_S	Density
DLP 30 fps	100	220	2.5	1	400	12	80%
Speckle 207 fps	170	240	17(52*)	1	59	12	90%
Stripe 120 fps	100–200	—	40	40	25	3	—
Time-of-flight	≈7000	—	≈60	≈60	—	—	—

[a] σ_P is the standard deviation of measured points to a fitted plane; σ_S is the standard deviation of measured points to a fitted sphere; 3D data/s is the number of possible 3D reconstructions from the captured data within 1 s (* sliding window of size four); 3D recons/s is the number of 3D reconstructions per second calculated in an on-line mode; the measurement time is denoted as t_m; N_S is the number of images used for one 3D reconstruction; the point density in the last column is the number of correspondences found divided by the number of points seen.

acquired, compared to an online 3D reconstruction rate of 40 fps with the stripe projection system. It may be possible to increase the reconstruction rate by using GPUs or more recent computer hardware. Although the performance of the speckle system is slightly worse compared to real-time stripe projection systems, it has no phase-unwrapping problems and, therefore, may be more suitable for certain applications, especially when it comes to scenes containing spatially separated objects.

6. Conclusion

We showed that it is feasible to use laser speckles to project statistical patterns at a rate of 207 fps for dense and accurate 3D reconstructions of various objects. Although the speckle projection system does have some disadvantages, e.g., patterns that are difficult to modify, it overcomes the speed limit of typical projection systems and, thus, may be more appropriate when it comes to moving or fast-changing objects. Additionally, it is able to measure spatially separated objects as well as objects containing unsteady height deviations. At present the setup is able to collect 17.25 independent, dense 3D reconstructions per second. When using a sliding window of size four for 3D reconstruction, up to 52 3D states of the object/scene can be calculated for every second of measured data. The accuracy of the setup is comparable to other real-time measurement systems. Compared to our reference system, the noise level of the 3D reconstruction is higher; however, up to 25% more correspondences can be found. In future work, we want to increase the accuracy by postprocessing the data as well as optimizing the speckle patterns. Furthermore, we want to increase the capturing rate to 400 Hz, which will enable us to collect dense 3D reconstructions of faster-moving objects. Additionally, faster processes will be addressed in the future by implementing a motion compensation algorithm.

References

1. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," Opt. Lasers Eng. **48**, 149–158 (2010).
2. T. Weise, B. Leibe, and L. V. Gool, "Fast 3D scanning with automatic motion compensation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)* (IEEE, 2007), pp. 1–8.
3. R. Yang, L. Wang, G. Welch, and M. Pollefeys, "Stereo vision on GPU," in *Workshop on Edge Computing Using New Commodity Architectures* (University of North Carolina, 2006).
4. S. Zhang, D. Royer, and S. Yau, "GPU-assisted high-resolution, real-time 3-D shape measurement," Opt. Express **14**, 9120–9129 (2006).
5. J. Pags, J. Salvi, R. Garcia, and C. Matabosch, "Overview of coded light projection techniques for automatic 3D profiling," in *IEEE Robotics and Automation (ICRA'03)* (IEEE, 2003), Vol. 1, pp. 133–138.
6. K. G. Harding, "Phase grating use for slop discrimination in moiré contouring," Proc. SPIE **1614**, 265–270 (1992).
7. Z. J. Geng, "Rainbow 3-D camera: new concept of high-speed three vision system," Opt. Eng. **35**, 376–383 (1996).
8. C. Wust and D. W. Capson, "Surface profile measurement using color fringe projection," Machine Vis. Appl. **4**, 193–203 (1991).
9. L. Zhang, B. Curless, and S. M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," in *IEEE Symposium on 3D Data Processing, Visualization, and Transmission* (IEEE, 2002), pp. 24–36.
10. M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," Appl. Opt. **22**, 3977–3982 (1983).
11. S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, "Real-time 3D model acquisition," in *SIGGRAPH 2002 Proceedings* (ACM, 2002), Vol. 21, pp. 438–446.
12. H. Guo and P. Huang, "3-D shape measurement by use of a modified Fourier transform method," Proc. SPIE **7066**, 70660E (2008).
13. B. Michaelis and P. Albrecht, "Stereo photogrammetry with improved spatial resolution," in *14th International Conference on Pattern Recognition* (IEEE, 1998), pp. 845–849.
14. A. Wiegmann, H. Wagner, and R. Kowarschik, "Human face measurement by projecting bandlimited random patterns," Opt. Express **14**, 7692–7698 (2006).
15. J. P. Siebert and S. J. Marshall, "Human body 3D imaging by speckle texture projection photogrammetry," Sensor Rev. **20**, 218–226 (2000).
16. M. Dekiff, G. Bischoff, Z. Borocz, D. Dirksen, G. von Bally, and C. Denz, "3d-formerfassung mittels korrelation projizierter specklemuster," in *Deutsche Gesellschaft für angewandte Optik Proceedings* (European Optical Society, 2008).
17. M. Grosse and R. Kowarschik, "Space-time multiplexing in a stereo-photogrammetry setup," in *Fringe 2009*, W. Osten and M. Kujawinska, eds. (Springer, 2009), pp. 755–759.
18. J. P. Lewis, "Fast normalized-cross-correlation," in *Vision Interface* (ACM, 1995).
19. H. Hirschmller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," Int. J. Comput. Vis. **47**, 229–246 (2002).
20. 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 (CVPR'03)* (IEEE, 2003), Vol. 2, pp. 359–366.
21. J. Mairal, R. Keriven, and A. Chariot, "Fast and efficient dense variational stereo on GPU," in *Proceedings of International Symposium on 3D Data Processing, Visualization, and Transmission* (IEEE, 2006), pp. 97–104.
22. H. Wagner, A. Wiegmann, R. Kowarschik, and F. Zollner, "3D measurement of human face by stereophotogrammetry," Proc. SPIE **5856**, 509–516 (2005).
23. Z. Zhang, "A flexible new technique for camera calibration," Tech. Rep. MSR-TR-98-71 (Microsoft Research, 1998).
24. R. I. Hartley, "In defense of the 8-point algorithm," IEEE Trans. Pattern Anal. Machine Intell. **19**, 580–593 (1997).
25. J. C. Dainty, ed., *Laser Speckle and Related Phenomena* (Springer Verlag, 1984).
26. M. Grosse, "Disturbing moiré effects in a stereo-photogrammetry setup," in *Deutsche Gesellschaft für angewandte Optik Proceedings* (European Optical Society, 2009).
27. S. Zhang, "High-resolution, real-time 3-D shape measurement," Ph.D. dissertation (Stony Brook University, 2005).
28. Centre Suisse d'Electronique et de Microtechnique, "Time of flight camera technology," Tech. Rep. (Centre Suisse d'Electronique et de Microtechnique, 2009).