

Simulation-based investigation on optical 3D surface measurement with composite spectral patterns

Chen Zhang^{*a}, Paul-Gerald Dittrich^a, Maik Rosenberger^a, Gunther Notni^{a,b}

^aTechnische Universität Ilmenau, Gustav-Kirchhoff-Platz 2, D-98693 Ilmenau, Germany;

^bFraunhofer Institute for Applied Optics and Precision Engineering, Albert-Einstein-Straße 7, D-07743 Jena, Germany

ABSTRACT

This paper presents the design and simulation of a single-shot optical 3D sensor based on multispectral pattern projection and a stereo-vision setup of two multispectral snapshot cameras. The performances of various combinations of available multispectral cameras, spatial light patterns, and 3D reconstruction algorithms as well as the geometric arrangements of the stereo-vision camera setup are simulated and analyzed. This simulation-based investigation delivers two optimized combinations of sensor components in terms of hardware and algorithm as orientations for practical sensor development in future, and an appropriate arrangement of the stereo-vision setup is determined. Moreover, the influences of sensor noise on 3D reconstruction are also estimated.

Keywords: 3D imaging, structured light, pattern projection, multispectral imaging

1. INTRODUCTION

High-precision optical 3D shape measurement technique is playing an increasingly important role in diverse fields such as in industrial machine vision, human-machine interaction, biomedicine, and security applications etc. Among various 3D measuring methods, the structured light technique, e.g. the phase-shift fringe projection profilometry,¹ presents itself as a well-established method for its outstanding depth resolution and completeness of obtained 3D point clouds. A major drawback of conventional structured light-based 3D methods is the real-time capability, which is primarily restricted by the switching rate of patterns using conventional projection devices and the large number of used light patterns. Hence, solution approaches to this real-time capability challenge can be proposed by raising the frequency of pattern switching, which are realized with high-speed projection techniques such as laser speckle projector,² multi-aperture projector,^{3,4} and GOBO projector,⁵⁻⁷ or reducing the number of patterns used for the calculation of a single 3D frame. For this approach, it is clear that the minimal pattern number is $N = 1$. Single-shot structured light 3D techniques could be divided into two classes according to the use of monochromatic or composite multispectral patterns. A typical technique in the first class is the Fourier profilometry.⁸ But it lacks in measurement stability and will confront with a lot of problems at surface with sharp edges and non-uniform optical properties. With the use of composite multispectral patterns, it is possible to project multiple different spatial light patterns simultaneously at different wavelengths. Using multi-channel image sensor these patterns are detected from a single image acquisition. In some previous works⁹⁻¹¹ the phase-shift color fringe pattern is proposed. In this pattern the three-step phase-shift is coded with RGB colors, so phase values can be calculated from one color image. The challenges with this color-based phase-shift method is the unwrapping of phase map without additional patterns (e.g. gray-codes) because of the limitation of only three channels.

Recently, miniaturized snapshot multispectral cameras¹²⁻¹⁴ offer novel possibilities in multi-channel image acquisition. These cameras realize simultaneous capturing of image data at different spectral bands, and thus exhibit the capability to extract more three light patterns from one acquisition. However, it should be taken into account that the most snapshot multispectral cameras are based on the sensor principle of on-chip multispectral filter array (MSFA) and hence have a reduced spatial resolution despite image demosaicing. Recently, we proposed a concept of single-shot optical 3D sensor based on multispectral array projection and snapshot multispectral cameras.¹⁵ First experiments showed the feasibility of the new measuring approach in principle. Here we conduct a fundamental investigation on the use of spectral patterns and MSFA-based multispectral cameras for 3D measurements. Using the mathematical model of multispectral imaging, we simulate 3D measurements using various multi-wavelength patterns and multispectral cameras with different number of spectral channels. Moreover, the influences of the size of the stereo baseline and sensor noise are also analyzed.

*chen.zhang@tu-ilmenau.de; phone +49 03677 693977; <http://www.tu-ilmenau.de/qualitaetssicherung/>

2. SENSOR CONCEPT

Figure 1 illustrates the proposed concept of optical 3D sensor based on multispectral array projection. It is composed of a stereo-vision setup of two multispectral snapshot cameras and an array projection device. The array projection device consists of N single projection units with N different spatial light patterns and light sources λ_1 to λ_N corresponding to the spectral characteristics of multispectral cameras. Considering that the alignment of the single projection units and the precise control of patterns' characteristics and their relation between each other are extremely difficult, it is advantage to use pseudo-statistical patterns to avoid too complex optics design. All the projection units are activated synchronously, so that different patterns are extracted from a snapshot multispectral image.

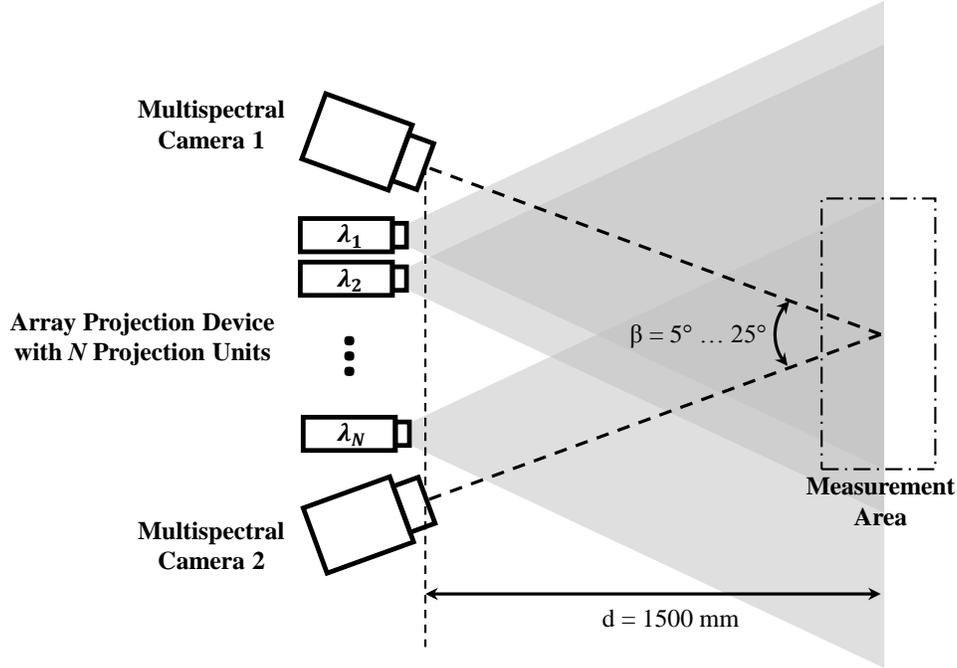


Figure 1. Concept of optical 3D sensor based on multispectral array projection.

As statistical patterns the speckle patterns can be firstly considered. Furthermore, Heist et al.¹⁶ proposed the aperiodic sinusoidal fringe patterns with statistical varying period lengths and phase shifts of fringes. These patterns exhibit an advantage regarding to the simplified fabrication. They can be fabricated as binary slides, while the sinusoidal intensity profile of fringes is produced with a minor defocusing of projection. The stereo matching at rectified stereo-vision setup can be realized at pixel level by calculating normalized cross-correlation between the sequence of N spectral values $I_1^{(1)}, \dots, I_N^{(1)}$ at each pixel in the first camera and the spectral value sequences $I_1^{(2)}, \dots, I_N^{(2)}$ at all pixels in the same row of the second camera. The pair of matched points should have the highest correlation coefficient ρ

$$\rho = \frac{\sum_{\lambda=1}^N [I_{\lambda}^{(1)} - \bar{I}^{(1)}][I_{\lambda}^{(2)} - \bar{I}^{(2)}]}{\sqrt{\sum_{\lambda=1}^N [I_{\lambda}^{(1)} - \bar{I}^{(1)}]^2} \sqrt{\sum_{\lambda=1}^N [I_{\lambda}^{(2)} - \bar{I}^{(2)}]^2}}, \text{ with } \bar{I}^{(k)} = \frac{1}{N} \sum_{\lambda=1}^N I_{\lambda}^{(k)}, k \in \{1,2\}. \quad (1)$$

Moreover, the pixel matching could be further extended to block matching in order to improve the stability. In this case, the correlation coefficient ρ is calculated based on all the pixels of a pixel block with pre-defined block size of $h \times w$

$$\rho = \frac{\sum_{y=1}^h \sum_{x=1}^w \sum_{\lambda=1}^N [I_{\lambda,x,y}^{(1)} - \bar{I}^{(1)}][I_{\lambda,x,y}^{(2)} - \bar{I}^{(2)}]}{\sqrt{\sum_{y=1}^h \sum_{x=1}^w \sum_{\lambda=1}^N [I_{\lambda,x,y}^{(1)} - \bar{I}^{(1)}]^2} \sqrt{\sum_{y=1}^h \sum_{x=1}^w \sum_{\lambda=1}^N [I_{\lambda,x,y}^{(2)} - \bar{I}^{(2)}]^2}}, \text{ with } \bar{I}^{(k)} = \frac{1}{h \cdot w \cdot N} \sum_{y=1}^h \sum_{x=1}^w \sum_{\lambda=1}^N I_{\lambda,x,y}^{(k)}, k \in \{1,2\}. \quad (2)$$

3. SIMULATION-BASED INVESTIGATION

The simulation of the optical 3D sensor proposed in section 2 is implemented using Autodesk 3ds Max and Matlab. In this investigation, we examine four combinations of two kinds of multispectral cameras that are available on the market with two kinds of pseudo-statistical patterns: band-limited speckle patterns and aperiodic fringe patterns. The simulation of 3D reconstruction is performed for three different baselines between both cameras. In the short-baseline arrangement of the stereo-vision camera setup, the camera triangulation angle β is 5° , while it is 15° and 25° in the middle- and wide-baseline camera arrangement, respectively. Figure 2 shows the spectral characteristics of the simulated 4-channel RGB-NIR cameras of PIXELTEQ¹⁷ and 8-channel multispectral NIR cameras of SILIOS TECHNOLOGIES.¹⁸ The simulated cameras have the same resolution (1280×1024 pixels) of sensor chip, so the spatial resolution of each spectral channel of the 8-channel cameras is 2/3 of the resolution of the channels of the 4-channel cameras. For image demosaicing the bicubic interpolation is used. LEDs with full width at half maximum of 50 nm and central wavelengths corresponding to the camera spectral channels are simulated as light sources for pattern projection. The middle working distance of the simulated sensor system is 1500 mm, and the lateral measurement field at the middle working distance is about $200 \text{ mm} \times 250 \text{ mm}$. 3D surface measurements at a sphere with 50 mm radius and diffuse surface is simulated. Figure 3 illustrate the speckle and aperiodic pattern projected onto this sphere object and their presentation in frequency domain. Disparity maps are constructed with pixel matching and block matching methods up to a block size of 5×5 pixels.

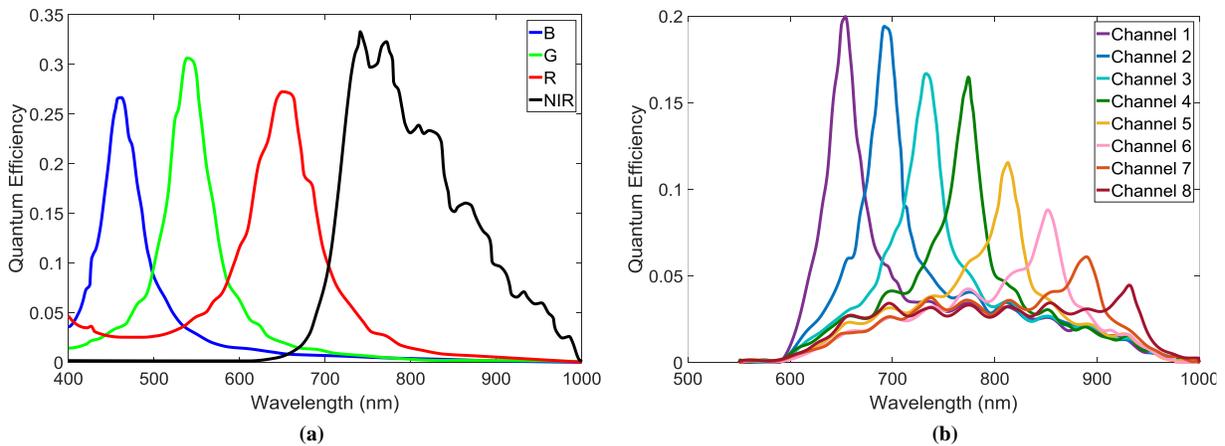


Figure 2. Spectral characteristics of 4-channel RGB-NIR camera (a) and 8-channel multispectral NIR camera (b).

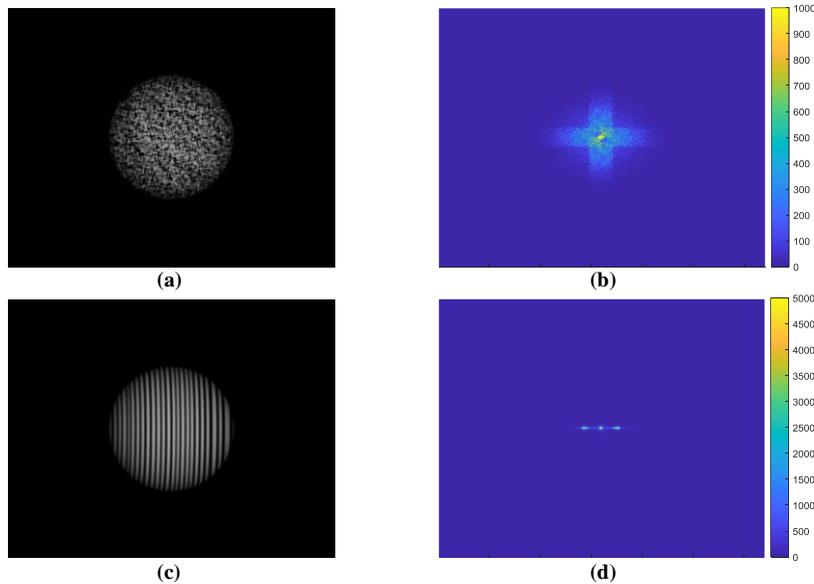


Figure 3. Spatial light patterns: band-limited speckle pattern in spatial domain (a) and frequency domain (b), and aperiodic fringe pattern in spatial domain (c) and frequency domain (d).

3.1 Simulation without noise

At first, simulation of 3D reconstruction without measurement noise is performed. Considering the limitations of signal-to-noise ratio (SNR) in practical applications, a sub-pixel accuracy of 0.1 pixels is calculated in the stereo matching with interpolation of spectral intensity values in each channel of spectral image cubes. For the quantitative evaluation of the reconstructed 3D point clouds, we calculate the root-mean-square error (RMSE) of sphere fitting, as illustrated in Figure 4, and the completeness of obtained 3D point cloud, which is defined as the ratio between the number of valid 3D points and the number of all object pixels in the 2D image. The results of the simulation of various measurement arrangements are shown in Table 1 to Table 3.

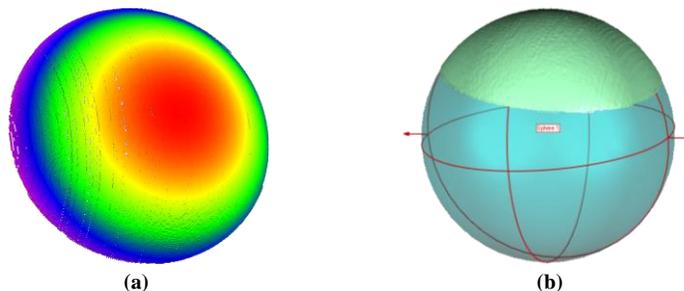


Figure 4. Simulated 3D measurement of sphere object: reconstructed 3D point cloud with false-color representation (a) and sphere fitting (b). This example is obtained using the combination of 8-channel multispectral NIR cameras, aperiodic fringe patterns, and pixel matching at the middle-baseline camera arrangement.

Overall, a good compromise between 3D accuracy and completeness of obtained 3D measurement result can be achieved with the middle-baseline arrangement of the stereo-vision camera setup, in which the camera triangulation angle is 15° . Comparing to the 3D results obtained with the short-baseline camera arrangement, the RMSEs are lower by at least 40% with the middle-baseline arrangement, while the completeness of 3D result remains almost the same. A further increase of the triangulation angle to 25° (wide-baseline camera arrangement) leads to more invalid 3D points and a deterioration of the completeness of measurement result. The reason to this may be that large camera view angle results in more high frequency components in the sensor plane and hence more artifacts by the image restoration to original sensor resolution due to the low-pass acquisition of raw image data using MSFA-based sensors. Moreover, the large difference of camera perspective leads to different demosaicing artifacts at both cameras and therewith increased errors in stereo matching.

Table 1. Results of simulation with short-baseline camera arrangement ($\beta = 5^\circ$).

Patterns	Multispectral cameras	Stereo matching method	RMSE of sphere fitting (mm)	Completeness of 3D measurement result
Speckle patterns	4-channel RGB-NIR cameras	Pixel matching	0.379	76.7%
		Block matching (3×3)	0.161	100%
		Block matching (5×5)	0.130	100%
	8-channel multispectral NIR cameras	Pixel matching	0.362	99.2%
		Block matching (3×3)	0.314	100%
		Block matching (5×5)	0.276	100%
Aperiodic fringe patterns	4-channel RGB-NIR cameras	Pixel matching	0.236	79.4%
		Block matching (3×3)	0.132	100%
		Block matching (5×5)	0.114	100%
	8-channel multispectral NIR cameras	Pixel matching	0.151	100%
		Block matching (3×3)	0.145	100%
		Block matching (5×5)	0.149	100%

Table 2. Results of simulation with middle-baseline camera arrangement ($\beta = 15^\circ$).

Patterns	Multispectral cameras	Stereo matching method	RMSE of sphere fitting (mm)	Completeness of 3D measurement result
Speckle patterns	4-channel RGB-NIR cameras	Pixel matching	-	-
		Block matching (3×3)	0.073	99.2%
		Block matching (5×5)	0.074	99.2%
	8-channel multispectral NIR cameras	Pixel matching	0.156	97.6%
		Block matching (3×3)	0.140	99.2%
		Block matching (5×5)	0.137	99.4%
Aperiodic fringe patterns	4-channel RGB-NIR cameras	Pixel matching	0.126	65.6%
		Block matching (3×3)	0.076	100%
		Block matching (5×5)	0.075	100%
	8-channel multispectral NIR cameras	Pixel matching	0.053	98.7%
		Block matching (3×3)	0.081	99.1%
		Block matching (5×5)	0.099	99.0%

Table 3. Results of simulation with wide-baseline camera arrangement ($\beta = 25^\circ$).

Patterns	Multispectral cameras	Stereo matching method	RMSE of sphere fitting (mm)	Completeness of 3D measurement result
Speckle patterns	4-channel RGB-NIR cameras	Pixel matching	-	-
		Block matching (3×3)	0.142	89.1%
		Block matching (5×5)	0.099	93.8%
	8-channel multispectral NIR cameras	Pixel matching	0.293	73.6%
		Block matching (3×3)	0.181	90.6%
		Block matching (5×5)	0.166	90.9%
Aperiodic fringe patterns	4-channel RGB-NIR cameras	Pixel matching	-	-
		Block matching (3×3)	0.075	90.0%
		Block matching (5×5)	0.073	89.7%
	8-channel multispectral NIR cameras	Pixel matching	0.038	94.9%
		Block matching (3×3)	0.084	93.5%
		Block matching (5×5)	0.098	90.2%

Table 1 and Table 2 show that the pixel matching method gives a poor performance at the 4-channel RGB-NIR cameras. Either this method delivers very noisy and incomplete 3D results, or it fails with large triangulation angles. This is due to the lower number (here $N = 4$) of spatial patterns used for the identification of matched image pixels. This problem can be mitigated by taking the spatial information of patterns into the stereo camera correlation. With the 8-channel cameras, characterized by $N = 8$, a robust pixel matching is feasible in general, whereby the utilization of aperiodic fringe patterns provides a few more valid 3D points. At these cameras, the block matching results in a reduction of RMSE at the short-baseline arrangement, but in an increase of RMSE at the middle-baseline arrangement.

With the middle-baseline arrangement, high completeness of 3D measurement result ($> 95\%$) can be achieved with the most combinations of patterns, cameras, and stereo matching methods except using the pixel matching method at 4-channel RGB-NIR camera, as aforementioned. Using speckle patterns and block matching method with a block size of 3×3 pixels, a RMSE of 0.073 mm and a 3D point cloud completeness of 99.2% are achieved with the 4-channel cameras. With the 8-channel NIR cameras, an even lower RMSE of 0.053 mm is achievable by the use of aperiodic fringe patterns and pixel matching, but the completeness of 3D result becomes fractionally lower. These both combinations of hardware sensor components and 3D reconstruction algorithms are identified as “good sensor settings”, for which the influences of fixed-pattern-noise of image sensor are further investigated in section 3.2.

3.2 Simulation with fixed-pattern-noise

In practice, the image quality is limited by sensor noise. For the approach to optical 3D measurement proposed in this work, a major error source is the fixed-pattern-noise (FPN), or rather photo response non-uniformity (PRNU) according to the EMVA 1288 standard¹⁹ in each spectral channel of multispectral cameras. For the investigation on the influences of FPN on the quality of 3D measurement, we simulate this noise at different levels and for each camera separately. For each image pixel, a Gaussian distributed spectral value magnification factor η is pseudo-randomly generated with mean value $\mu = 1$ and standard deviation σ . Thereafter, the original image data are multiplied with the generated magnification factors.

In this investigation the range of spatial standard deviation σ of FPN is from 0.1% to 5.0%. Based on the image data with simulated FPN, 3D reconstructions are performed for the both “good sensor settings” chosen in section 3.1. The results of this investigation are illustrated in Figure 5. Here the combination of 4-channel cameras, speckle patterns, and 3×3 block matching is abbreviated as “4C-SP-BM”, while the combination of 8-channel cameras, aperiodic fringe patterns, and pixel matching is abbreviated as “8C-AFP-PM”.

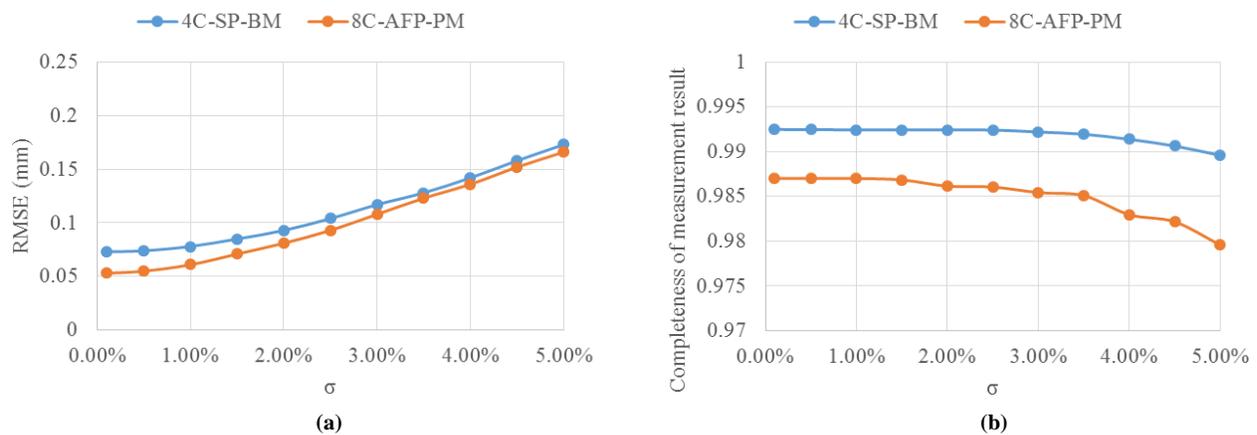


Figure 5. Quality scores of 3D measurement with “good settings” at different levels of fixed-pattern-noise: RMSE (a) and completeness of measurement result (b).

In Figure 5(a) a growth of RMSE can be recognized from $\sigma = 0.5\%$, from which an almost linear relation between σ and RMSE can be seen. From $\sigma = 3.5\%$ there is only a minimal difference between RMSE values obtained with both sensor settings. Figure 5(b) shows the decreases of completeness of 3D measurement result with the rise of fixed-pattern-noise. The setting “4C-SP-BM” exhibits a better robustness of stereo matching to noise. With this setting the completeness falls from 99.2% to 99% at $\sigma = 0.5\%$, while it falls from 98.7% to 98% with the setting “8C-AFP-PM”.

4. SUMMARY AND FUTURE WORK

In this work, we presented the design and simulation of a single-shot optical 3D sensor based on multispectral pattern projection and stereo snapshot multispectral cameras. In this theoretical investigation we simulated the performances of various combinations of multispectral snapshot cameras, spatial light patterns, and algorithms for stereo matching, as well as the influences of the geometric arrangement of the stereo-vision setup. This investigation shows that proposed sensor concept is fundamentally possible for both 4-channel and 8-channel cameras with suitable choices of projection patterns and stereo matching methods. From the simulation two combinations of sensor components can be suggested for the practical sensor development. Moreover, the simulation indicates that the triangulation angle should be around 15° in

order to achieve a good compromise between accuracy of 3D measurement and completeness of 3D measurement result. From a practical point of view, the 8-channel multispectral NIR cameras could exhibit the advantage that the optical 3D measurements in the near-infrared spectral range are less influenced by objects' own color.

The proposed sensor concept enables the extraction of multiple light patterns from a single image acquisition, hence, the quality of single-shot 3D imaging can be greatly improved in terms of depth resolution and stability of stereo matching. Using this approach, high-quality 3D image data could be captured at the frame rate of used 2D sensors, so the 3D frame rate can be raised significantly with the use of high-speed image sensors. However, it may be restricted by the necessary exposure time in practice because of the low quantum efficiency of most snapshot multispectral cameras.

One of the main systematic error sources is the fixed-pattern-noise of image sensors. However, the fixed-pattern-noise is a deterministic non-uniformity of the pixel sensitivity over the active sensor area and hence could be eliminated based on a sensor characterization following the EMVA1288 standard.¹⁹ Figure 5 shows that the impact of fixed-pattern-noise is practically negligible if it is suppressed to a level $\sigma < 1\%$.

A further challenge in practical applications is the difference of spectral characteristics between two multispectral image sensors due to the uncertainties in fabrication process. There could be differences in central wavelengths and shapes of spectral sensitivity curves of the camera spectral channels. Investigation on influences of such effects and experimental evaluation of the proposed approach to 3D measurement will be continued in the future.

ACKNOWLEDGMENTS

This work has been sponsored by the German Federal Ministry of Education and Research in the program Zwanzig20 – Partnership for Innovation as part of the research alliance 3Dsensation (project number: 03ZZ0462).

REFERENCES

- [1] Kühmstedt, P., Munkelt, C., Heinze, M., Bräuer-Burchardt, C., and Notni, G., "3D shape measurement with phase correlation based fringe projection," *Proc. of SPIE* 6616, 66160B (2007).
- [2] Schaffer, M., Grosse, M., Harendt, B., and Kowarschik, R., "High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection," *Opt. Lett.* 36(16), 3097–3099 (2011).
- [3] Heist, S., Mann, A., Kühmstedt, P., Schreiber, P., and Notni, G., "Array projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement," *Opt. Eng.* 53(11), 112208 (2014).
- [4] Brahm, A., Ramm, R., Heist, S., Rulff, C., Kühmstedt, P., and Notni, G., "Fast 3D NIR systems for facial measurement and lip-reading," *Proc. SPIE* 10220, 102200P (2017).
- [5] Hyun, J.-S., Chiu, G. T.-C., and Zhang, S., "High-speed and high-accuracy 3D surface measurement using a mechanical projector," *Opt. Express* 26(2), 1474-1487 (2018).
- [6] Heist, S., Lutzke, P., Schmidt, I., Dietrich, P., Kühmstedt, P., Tünnermann, A., and Notni, G., "High-speed three-dimensional shape measurement using GOBO projection," *Opt. Lasers Eng.* 87, 90–96 (2016).
- [7] Heist, S., Dietrich, P., Landmann, M., Kühmstedt, P., Notni, G., and Tünnermann, A., "GOBO projection for 3D measurements at highest frame rates: a performance analysis," *Light: Science & Applications* 7(1), 71 (2018).
- [8] Takeda M. and Mutoh K., "Fourier transform profilometry for the automatic measurement of 3-D object shapes," *Appl. Opt.* 22(24), 3977-3982 (1983).
- [9] Zhang, Z., Jing, Z., Feng, H., Ma, H., Guo, T., and Zhang, S., "Absolute phase calculation from one composite RGB fringe pattern image by windowed Fourier transform algorithm," *Proc. SPIE* 7855, 78550I (2010).
- [10] Wang, Z., Zhang, Z., Guo, T., Zhang, S., and Hu, X., "Absolute phase calculation from one composite RGB fringe pattern image by wavelet transform algorithm," *Proc. SPIE* 8200, 82000H (2011).
- [11] Wang, Y., Sun, Y., Zhang, T., Chen, D., and Yu, X., "A Novel Color Encoding Fringe Projection Profilometry based on Wavelet Ridge Technology and Phase-Crossing," *International Journal of Performability Engineering* 14(5), 917-926 (2018).
- [12] Geelen, B., Tack, N., and Lambrechts, A., "A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic," *Proc. SPIE* 8974, 89740L (2014).
- [13] Lapray, P.-J., Wang, X., Thomas, J.-B., and Gouton, P., "Multispectral Filter Arrays: Recent Advances and Practical Implementation," *Sensors* 14(11), 21626-21659 (2014).

- [14] Heist, S., Zhang, C., Reichwald, K., Kühmstedt, P., Notni, G., and Tünnermann, A., “5D hyperspectral imaging: fast and accurate measurement of surface shape and spectral characteristics using structured light,” *Opt. Express* 26(18), 23366–23379 (2018).
- [15] Zhang, C., Brahm, A., Breitbarth, A., Rosenberger, M., and Notni, G., “Single-frame three-dimensional imaging using spectral-coded patterns and multispectral snapshot cameras,” *Opt. Eng.* 57(12), 123105 (2018).
- [16] Heist, S., Kühmstedt, P., Tünnermann, A., and Notni, G., “Theoretical considerations on aperiodic sinusoidal fringes in comparison to phase-shifted sinusoidal fringes for high-speed three-dimensional shape measurement,” *Appl. Opt.* 54(35), 10541–10551 (2015).
- [17] PIXELTEQ. Available online: <https://pixelteq.com/pixelcam/> (accessed on 24 May 2019).
- [18] SILIOS TECHNOLOGIES, MICRO-OPTICS Supplier. Available online: <http://www.silios.com/> (accessed on 24 May 2019).
- [19] European Machine Vision Association, “EMVA Standard 1288 - Standard for Characterization of Image Sensors and Cameras,” (2016).