

Design and quality metrics of point patterns for coded structured light illumination with diffractive optical elements in optical 3D sensors

Ralf Vandenhouten^a, Andreas Hermerschmidt^b, and Richard Fiebelkorn^a

^aWildau University of Applied Sciences, Hochschulring 1, 15745 Wildau, Germany

^bHOLOEYE Photonics AG, Albert-Einstein-Straße 14, 12489 Berlin, Germany

ABSTRACT

Structured light has become a widespread technique for the development of camera-based 3D sensors. The structured illumination provides texture to homogeneous objects and thus allows for the reliable determination of the disparity of each object point in a stereo-camera setting. Even a monocular 3D sensor is possible if the light projector has a fixed relative position to the camera and if the structured light is coded, i.e. the position within the whole light pattern can be reconstructed uniquely from a small local window of the pattern, the uniqueness window. Coded patterns with such a uniqueness property are called Perfect SubMaps (PSM).

In our paper we focus on the design and evaluation of the subset of symmetric isolated binary toroidal PSMs (SIBTPSM) for structured light patterns, because of their beneficial properties with respect to the signal-to-noise ratio and the use with laser light sources and DOEs. We define several figures of merit that are relevant for the practical use of PSMs in a 3D sensor: the PSM size, the size of the uniqueness window, the Hamming distance, the density, and the homogeneity. We have created SIBTPSMs using our own dedicated algorithms and have designed and fabricated DOEs that produce these patterns with large fan angles of $61^\circ \times 47^\circ$ when used with near-infrared diode lasers ($\lambda = 830nm$). We analyze the influence of these characteristics on the 3D measurement process by theory, simulations, and experiments. The patterns of publicly available DOEs based on SIBTPSMs are used for comparison and reference.

Our results show that the PSM width, the uniqueness window size, the minimum and average Hamming distances, and the uniformity have strong impact on either speed or quality of the 3D reconstruction, whereas the point density and the PSM height are of minor importance.

Keywords: structured light, 3D sensor, diffractive optical element, DOE, Perfect SubMap, PSM, symmetric PSM

1. INTRODUCTION

Due to advances in high-resolution and high-speed electronic imaging sensors and increasing computational power there has been considerable progress in the development of 3D surface imaging technologies over the past few years. The applications are numerous and range from optical position measurement, scanning of objects, and gesture recognition to the endoscopic detection of physiological parameters, among others.¹⁻⁴ Various technical approaches for 3D surface imaging have been established: stereo vision, time-of-flight cameras, laser triangulation, and structured light. In this paper we focus on structured light systems which are known for their good overall performance in terms of resolution, speed of measurement, and cost efficiency.

Techniques for surface imaging by structured light can be categorized as either sequential (multiple-shot) or single-shot. As sequential approaches are appropriate only for static objects we concentrate on single-shot techniques, which in turn can be divided into three categories: techniques using continuously varying structured-light patterns, techniques based on 1D encoding schemes (strip indexing),⁵ and techniques using 2D encoding

Further author information: (Send correspondence to R.V.)

R.V.: E-mail: ralf.vandenhouten@th-wildau.de, Telephone: +49 3375 508359

A.H.: E-mail: andreas.hermerschmidt@holoeeye.com, Telephone: +49 30 63923667

schemes (grid indexing).¹ We focus on the latter because 2D structured light encoding patterns are the most robust with respect to varying object reflectivity and distortions of the optical system.

A 3D surface imaging system based on a 2D structured light pattern at least consists of a camera and a pattern projector with a fixed relative position of both. The depth of each object spot can then be calculated from the disparity of the pattern detail the camera sees at this spot. As the structured illumination provides texture to homogeneous objects it can also help to improve the 3D surface imaging in a conventional stereo vision system because it increases the depth visibility of low-textured object parts.⁶ Recently, 2D encoding patterns have been used even for light field 3D imaging with structured light.⁷

A light pattern projector is needed in any case. In principle any kind of projector is suitable for this purpose (e.g. a conventional projector based on imaging) if the light spectrum meets the requirements of the camera sensor. However, laser diode light sources are preferable because they can be very small and inexpensive and have good projection properties due to their point-shaped beam. Due to the coherent laser light the pattern can then be created by sending the collimated laser beam through a diffractive optical element (DOE) that modifies the light wave in such a way that the pattern will appear by diffraction. In this article, we examine the properties of patterns that can be generated by such DOE projectors.

The paper is structured as follows: In section 2 we investigate the properties of 2D structured light patterns for DOEs and how they can be created. Section 3 defines the quality metrics of such DOE patterns. The properties of the according DOEs are presented in section 4. Some simulation results with publicly available real-world patterns and some live experiments are shown in section 5, before we will finish with a conclusion.

2. STRUCTURED LIGHT PATTERNS

The concept of 2D encoding pattern techniques is to uniquely label every subwindow in the projected pattern, such that the position within the whole light pattern can be reconstructed uniquely from a small local window of the pattern, the *uniqueness window* of size $u \times v$. Coded patterns with such a uniqueness property (or window property) are called Perfect SubMaps (PSM).⁸

2.1 Constraints imposed by utilization of DOEs

In principle, the PSMs can contain any number of brightness levels (i.e. the size of the alphabet of the code). The maximum signal-to-noise ratio, however, which is particularly important for monocular 3D surface imaging systems and for non-uniform object reflectivity, is achieved if only two brightness levels are used⁸ and if the light spots are *isolated*. In this case the PSM is a *binary* grid pattern with code symbols that are either 0 or 1 and, for a robust detection of the light spots, each 1 is surrounded by zeros (*isolation* property).

For most optical 3D sensors it is sufficient that only a repeating part of the structured light pattern is a PSM, i.e. the complete pattern consists of identical tiles that are *toroidal* PSMs (TPSM). A TPSM of size $W \times H$ is a PSM mapped on a torus with

$$p(i + nW, j + mH) = p(i, j) \quad \forall i \in \{0, \dots, W - 1\}, j \in \{0, \dots, H - 1\}, n, m \in \mathbb{N} \quad (1)$$

and satisfies the uniqueness condition on the torus, i.e. all sub windows of size $u \times v$ on the torus are unique (see fig. 1). The toroidal property ensures that the unique mapping does not get lost at the border of the tiles, i.e. the minimum horizontal distance of two identical subwindows of size $u \times v$ in the tiled pattern is W and the minimum vertical distance is H .

The most efficient way to build powerful structured light projectors with a small size is to combine a collimated diode laser with a diffractive optical element (DOE) that generates the PSM from the collimated laser beam. In applications, large projection angles, high contrast, and a good suppression of the undesired undiffracted (zeroth-order) laser beam are required. This currently can be achieved only with binary DOEs. However, such DOEs only can generate patterns that are *central symmetric* with respect to the 0th order of diffraction, i.e. the center of the pattern.⁹ Therefore, in this work we focus on the design and the evaluation of (central) symmetric isolated binary toroidal PSMs (SIBTPSM) for structured light DOEs.

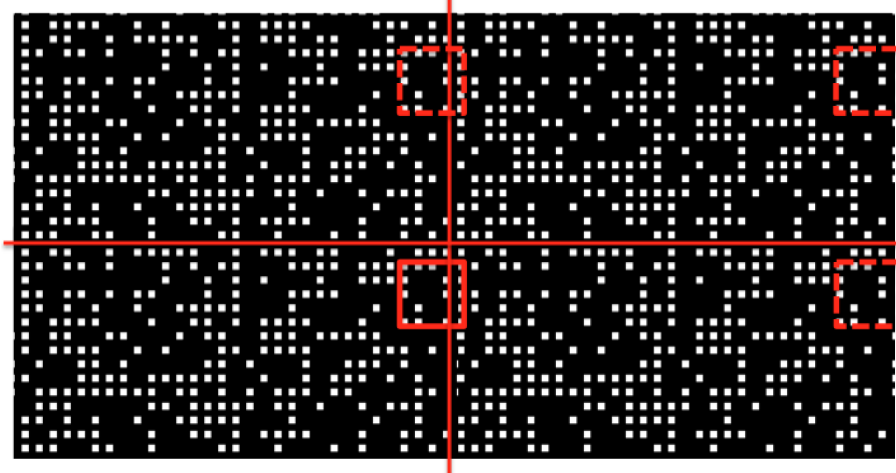


Figure 1. Four tiles of a toroidal PSM (TPSM) with a 9×9 uniqueness window. The uniqueness condition is satisfied even at the border of the tiles.

2.2 Generation of PSMs

Various approaches have been described for generating PSMs. Some of them are derived from De Bruijn sequences which can be regarded as a one-dimensional PSM with a k window property that contains each possible code in its subsequences of length k exactly once. De Bruijn sequences can easily be constructed e.g. by Galois fields¹⁰ or graph theory¹¹ and can be used for 1D encoding schemes (strip indexing).¹²

2D toroidal PSM tilings can be created by combining two one-dimensional De Bruijn sequences appropriately, either by calculating the outer product, by a shifting scheme, or by folding.^{10,13–16} The construction of 2D toroidal PSMs not based on De Bruijn sequences is more difficult. Several approaches have been suggested using evolutionary algorithms^{14,17} or brute force exhaustive search.^{8,18,19} These algorithms are computationally very expensive and only allow for the creation of relatively small binary PSM patterns in a reasonable time.

However, there is a drawback that is even worse. All of the PSMs in the cited publications are not SIBTPSMs and thus cannot be used in our case. They are neither central symmetric nor do they meet the isolation property. Most of them are even not binary but use either more than 2 code symbols or color codes not suitable for DOEs. The same holds true for the PSMs presented in Refs. 20–22. A central symmetric monochrome PSM is created in Ref. 23. However, it uses 3 symbols, has a size of only 27×29 , and it does not satisfy the isolation property. The PSMs suggested in Ref. 24 have a considerable size of 200×200 and satisfy the isolation property, but they are not SIBTPSMs because they lack the central symmetry.

To the best knowledge of the authors, the only published sufficiently large SIBTPSM is the proprietary DOE pattern developed by the company PrimeSense, that was used in their PS 1080 sensor, the popular Microsoft Kinect sensor, and the ASUS Xtion Pro Live. It was reverse engineered and analyzed in Ref. 26, and in the following we refer to the SIBTPSM of this pattern as KIN. The three mentioned sensors based on these patterns have been widely used by the scientific community for various projects in which 3D data acquisition was required and therefore represent a suitable benchmark for our efforts presented in this paper.

For lack of available SIBTPSM patterns with considerable size we used our own proprietary algorithms (which are beyond the scope of this paper) to generate three SIBTPSMs on a massively parallel high-performance computer (referred to as VAN2, VAN3, and VAN4, respectively). Sections of these patterns are shown in fig. 2. In the following sections the properties of these patterns are analyzed and compared to the KIN pattern.

3. QUALITY METRICS

For the evaluation and comparison of different structured light patterns in general, and (SIBT)PSMs in particular, appropriate quality metrics have to be defined. One obvious quality feature of a PSM is its size, i.e. its *width* and

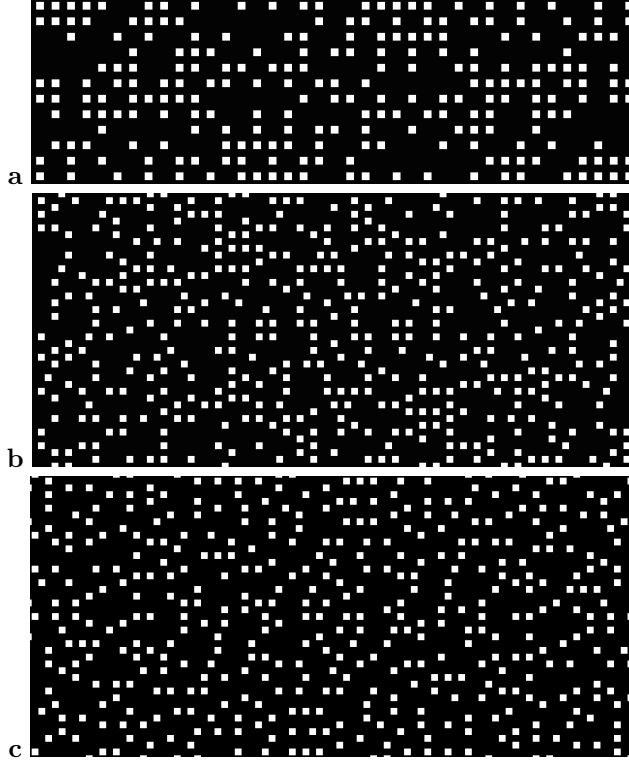


Figure 2. Central sections of the three SIBTPSMs generated by one of the authors: (a) VAN2, (b) VAN3, and (c) VAN4.

its *height*, as these quantities are directly related to the maximum disparity (and thus the maximum depth range) that can be determined unambiguously by a stereo correspondence algorithm. So in general a larger pattern is better. However, in the usual and often preferable setup where the horizontal axis of the pattern is parallel to the epipolar axis between the pattern projector and the camera(s) only the width of the pattern is almost directly proportional to the maximum unambiguous disparity. The PSM height is usually less important as the vertical pattern axis is orthogonal to the stereo disparity. Thus the effect of the pattern height depends more or less on the distortion of the projected pattern, i.e. the geometric imaging quality of the DOE. If the epipolar configuration is perfect and the distortion is very low the PSM height can be reasonably small (typically a few times the uniqueness window height for a robust 3D sensor). If the distortion is large, as e.g. with the Kinect DOE,²⁷ the PSM should be correspondingly higher in order to avoid a stereo correspondence mismatching.

As the uniqueness window defines the smallest subwindow of the pattern that can uniquely be identified for a reliable stereo matching, the *size of the uniqueness window* is an indicator of the lateral resolution of the depth measurement. The smaller the uniqueness window, the higher the resolution of the depth image. Another advantage of a small uniqueness window is the better decoding performance, particularly in a monocular setup where the code patterns in the subwindows must be mapped to code positions directly. The number of bits of a codeword in a binary pattern is equal to the size of the uniqueness window. This is of practical relevance, since, for example, a codeword of a PSM with 7×7 window property needs only 60% of the bit length of a PSM with 9×9 window property and can be stored in a simple 64-bit integer variable.

Real world recordings of the projected pattern may contain errors which weaken the robustness of the uniqueness property. This can be described by the Hamming distance which is defined for two subwindows of size $u \times v$ at the positions (i_1, j_1) and (i_2, j_2) , respectively:⁸

$$H_{u,v}((i_1, j_1), (i_2, j_2)) = \sum_{i=0}^{u-1} \sum_{j=0}^{v-1} \delta(p(i_1 + i, j_1 + j), p(i_2 + i, j_2 + j)) \quad \text{with} \quad \delta(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases} \quad (2)$$

The *minimum Hamming distance* $H_{u,v}^{min}$ and the *average Hamming distance* $\bar{H}_{u,v}$ of the PSM with size $W \times H$ can then be defined as

$$H_{u,v}^{min} = \min_{\forall (i_1, j_1) \neq (i_2, j_2)} H_{u,v}((i_1, j_1), (i_2, j_2)) \quad (3)$$

and

$$\bar{H}_{u,v} = \frac{1}{WH(WH-1)} \sum_{(i_1, j_1) \neq (i_2, j_2)} H_{u,v}((i_1, j_1), (i_2, j_2)). \quad (4)$$

In particular, a PSM with $u \times v$ window property must satisfy the condition $\bar{H}_{u,v} \geq H_{u,v}^{min} \geq 1$. A larger Hamming distance means a more robust pattern against noise or other error, so the minimum and the average Hamming distance are expected to be useful quality metrics of PSMs.

Another pattern metric is the *spot density* ρ which can be defined as the number of light dots divided by the total number of grid positions in the pattern, i.e.

$$\rho = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{H-1} p(i, j)}{W \cdot H}. \quad (5)$$

Finally, the quality of the 3D reconstruction can depend on the uniformity of the light dots (i.e. the 1's) in the pattern. Thus we define the *n-uniformity* γ_n as the standard deviation of the light dot frequency in a sliding window of size $n \times n$, i.e.

$$\gamma_n = \sqrt{\frac{1}{WH-1} \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \left(\sum_{k=0}^{n-1} \sum_{l=0}^{n-1} p(i+k, j+l) \right)^2 - \mu} \quad \text{with } \mu = n^2 \rho. \quad (6)$$

With this definition, a lower value of γ_n corresponds to a more uniform distribution of the light spots in the pattern.

Table 1. Properties of the SIBTPSMs 'VAN2', 'VAN3', 'VAN4', and 'KIN', respectively. The best values in each category are printed in bold face.

SIBTPSM	VAN2	VAN3	VAN4	KIN
Size	126 x 28	212 x 96	211 x 97	211 x 165
Spot number	440	2516	2282	3860
Corresponding spot density ρ	12.5%	12.4%	11.1%	11.1%
Min. uniqueness window size $u \times v$	8x8	7x7	7x7	9x9
Min. Hamming distance				
$H_{7,7}^{min}$	0	1	1	0
$H_{8,8}^{min}$	1	1	1	0
$H_{9,9}^{min}$	1	2	3	1
Average Hamming distance				
$\bar{H}_{7,7}$	10.7	10.6	9.7	9.7
$\bar{H}_{8,8}$	14.0	13.9	12.7	12.6
$\bar{H}_{9,9}$	17.7	17.6	16.0	16.0
n-Uniformity γ_7	2.1	1.7	1.1	1.1

The calculated metrics of the patterns VAN2, VAN3, VAN4, and KIN, respectively, are shown in table 1. It should be noted that the various metrics are not completely independent. For example, with a growing PSM size it becomes more difficult to ensure a large average Hamming distance. This is the reason why the relatively small pattern VAN2 gets the best marks in this category. The average Hamming distance also naturally decreases with a decreasing spot density since the number of possible codewords is reduced then. This accounts for the very good average Hamming distances of the relatively large VAN3 pattern with a high spot density compared to VAN4 and KIN.

4. DIFFRACTIVE OPTICAL ELEMENTS

The diffractive optical element for the generation of the structure light pattern needs to fulfill a few conditions in order to be usable in practical application. The projection should be created with no or low geometrical distortion, it should provide high light efficiency, good uniformity of the power distribution among the desired spots, and a high contrast. In addition, for the use with higher power lasers, a very good suppression of the undesired zeroth diffraction order is required, in order to permit the classification of the laser-based projection unit incorporating the DOE into a low laser safety class, ideally as an eye-safe device. Last but not least it is desirable to use only few surface levels in the microstructure in order to limit the fabrication effort, and we have therefore focused from the start on using binary DOEs with just two surface levels.

Diffractive optical elements for high-angle patterns require that each spot is chosen as a suitable diffraction order, so that the directions of the created diffraction beams form an undistorted pattern.²⁸ After selection of the desired diffraction orders to define the target intensity distribution among the spatial frequencies represented by the diffraction pattern of the DOE, well-known design algorithms like the IFTA²⁹ can be applied. A good angular accuracy can already be achieved by using a sufficiently large grating period of the DOE, which results in a fine spatial frequency grid. Moreover, it is possible to create arbitrary diffraction angles even with binary DOEs.³⁰

For zero order reduction, one option can be to use two DOEs in a sequence.²⁵ Nevertheless it is more attractive to use only a single DOE, which makes the diffractive projector simpler and allows for more degrees of freedom in the pattern design, particularly with respect to repetitions of the SIBTPSM and the correction of geometric distortions. In order to have a low zero order, there is a simple model for binary DOEs, provided that the feature sizes are large enough to justify the application of scalar diffraction theory to the problem. Both partial surfaces of the DOE can be considered the source of two coherent light waves which will then interfere in the far field. In order for this interference to be destructive, which corresponds to a suppression of the zeroth order, both partial waves need to carry the same share of the energy (i.e. 50%), and have a mutual phase delay of π . Fourier diffraction theory⁹ provides also an analytic expression for the zeroth order efficiency η_0 . It is simply given as the modulus of the integral of the light field amplitude over the aperture of the DOE, or under the simplifying assumption of a incident plane wave with constant intensity, as

$$\eta_0 = \cos^2 \frac{\phi}{2} + (1 - 2f)^2 \sin^2 \frac{\phi}{2}, \quad (7)$$

where ϕ denotes the phase delay introduced between the partial waves by the DOE, and f denotes the share of the area of the DOE microstructure surface that is air, compared to the remaining area being a dielectric material, or vice versa.

In our case however, we need to mention that due to the larger diffraction angles we have targeted, the features sizes of the binary DOE get so small, that scalar diffraction theory is not valid anymore. And because both the grating period and the number of created diffraction orders are quite large, a rigorous simulation is not possible. Nevertheless for other, simpler binary DOEs which can be simulated, it can be shown that the quantities that need to be optimized in order to suppress the zero order and to optimize the diffraction efficiency are in fact still the same.³¹ After design and fabrication of the DOEs using laser lithography, galvanofarming and compression molding with polycarbonate as the material of both the substrate and the microstructure of the DOE, we measured the zeroth order experimentally for a range of wavelengths (see figure 3). It is evident that we were able a very good suppression of the zeroth order to 0.1% at the laser wavelength λ_D of 830nm targeted by our design, and that for other wavelengths the zeroth order increases, as predicted by formula 7.

When changing the laser wavelength λ , in addition to the increase in zeroth order power, also scaling and geometrical distortion effects will occur. A laser spot which in any plane perpendicular to the optical axis is incident at a distance proportional to $\tan \theta(\lambda_D)$, will as a function of λ move to

$$\tan \theta(\lambda) = \left(\sqrt{\frac{\lambda_D^2}{\lambda^2} \left(1 + \frac{1}{\tan^2 \theta(\lambda_D)} \right) - 1} \right)^{-1}, \quad (8)$$

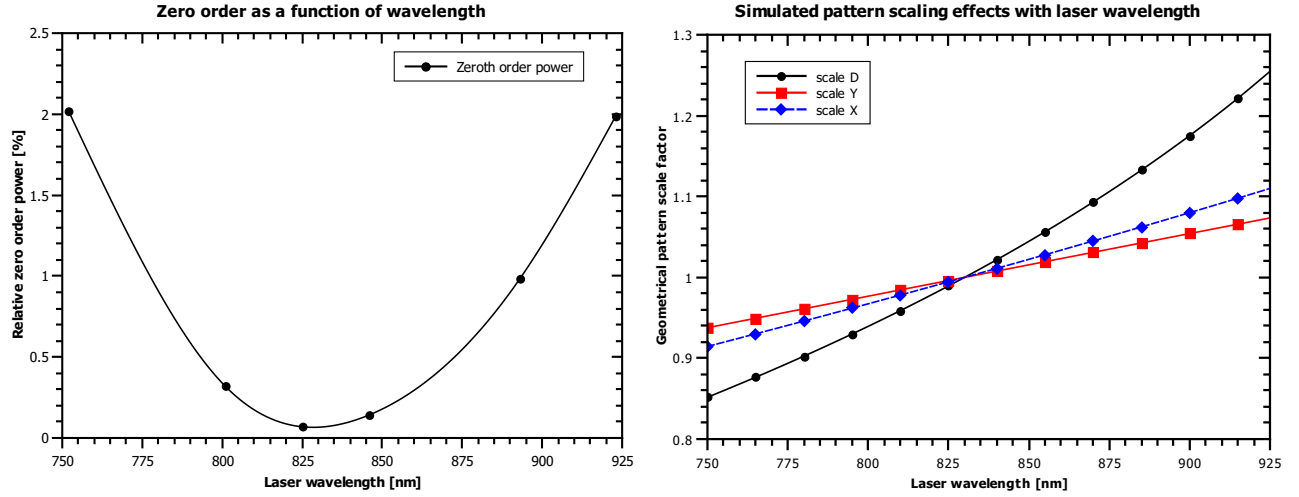


Figure 3. Left: Change of the relative power of the undiffracted (i.e. zeroth) order with the wavelength, measured with laser diode sources in the NIR range. Right: Change of pattern boundary coordinates with the wavelength, in x -, y - and diagonal direction.

which leads to different (and for large changes $|\lambda - \lambda_D|$ increasingly nonlinear) scaling effects, as can be seen in figure 3. For wavelengths smaller than λ_D , a barrel distortion of the pattern will be observable, and for larger wavelengths a cushion distortion. Nevertheless, for the rather small changes of wavelength in the application, in many cases mostly governed by the change of laser diode wavelength with temperature (typically 0.3nm/K), the distortion effects are small, and if necessary can be taken into account.

Table 2. Properties of the DOEs based on the SIBTPSMs 'VAN2', 'VAN3' and 'VAN4', respectively

DOE	DE-R 332 (VAN2)	DE-R 372 (VAN3)	DE-R 375 (VAN4)
Base Tile Repetitions	5x15	3.4 x 4.7	2.85 x 4.55
Total size	629 x 419	721 x 451	601 x 441
Aspect Ratio	1.5 (3:2)	1.6 (16:10)	1.36 (4:3)
Total spot number	33000	40100	29594
DOE design wavelength λ_D [nm]	830	850	830
Field of view angles [deg]	59.6 x 41.7	59.7 x 39.6	61.3 x 46.8
Grid spacing at 1m working distance [mm]	1.825 x 1.825	1.6 x 1.6	1.975 x 1.975

As already mentioned in section 2.1, patterns often contain repetitions of tiles, and for the sake of achieving a large enough field of view, the three SIBTPSMs (compare figure 2) were used as base tiles for three larger patterns, which were then created by three dedicated DOEs. Table 2 summarizes the properties of the DOEs.

5. RESULTS

To evaluate our SIBTPSMs and DOEs with regard to their usefulness for optical 3D sensing we tested them in two different experiments: a simulation of an idealized, distortion-free projection of the PSM, recorded with a perfect and distortion-free stereo camera pair, and a real-world experiment with a laser projector using the real DOEs and an NIR stereo camera setup. In order to get quantitative results we use a simple virtual sloping wall as a scene for the simulation and a slightly more complex scene with two boxes in the foreground for the live tests. The details and the results are presented in the following sections.

5.1 Simulation

In our simulation experiment we assumed the following setup: Two identical coplanar cameras with distortion-free lenses and a relative distance of 70mm are directed to the object, a sloping planar wall in a distance of 1200mm that is rotated by 15 degrees around the vertical axis. The epipolar geometry is supposed to be perfect, i.e. the line connecting the two centers of the image planes is part of the common x-axis, and thus the images are already rectified for stereo matching. The laser projector is coplanar to the cameras and on the x-axis as well. The distance from the left camera is 50mm, and 20mm from the right camera. By means of ray tracing and triangulation we computed the two images recorded by the cameras.

These synthetic images were used for stereo matching to calculate the depth image. For this we chose the semi-global stereo block matching algorithm `StereoSGBM` of the OpenCV library³² in the mode `MODE_SGBM_3WAY` with the following parameters (i.e. with all filtering options switched off): `minDisparity = 0`, `numDisparities = 96`, `blockSize = 7`, `P1 = P2 = 0`, `disp12MaxDiff = 0`, `preFilterCap = 0`, `uniquenessRatio = 1`, `speckleWindowSize = 0`, `speckleRange = 0`.

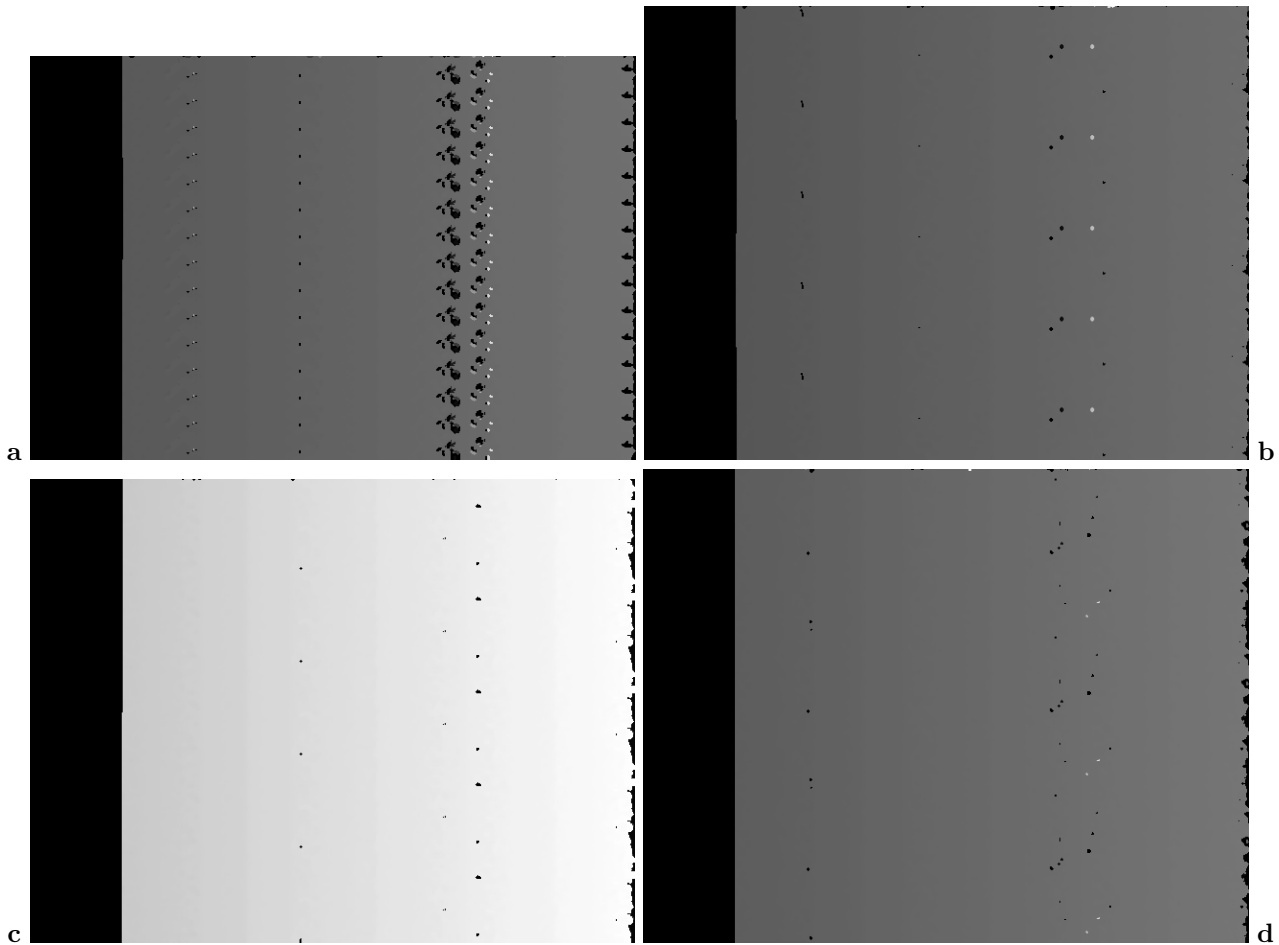


Figure 4. Normalized disparity images from unfiltered semi-global block matching of simulated undistorted stereo image pairs using the following SIBTPSMs: (a) VAN2, (b) VAN3, (c) VAN4, (d) KIN. Presumably due to the large minimum Hamming distance of 3 VAN4 is completely free of mismatching-related outliers.

The resulting disparity images are shown in fig. 4. The brightness of a point is proportional to the corresponding disparity and thus inversely proportional to the depth. Positions that could not be matched are marked in black. However, the disparity images were normalized to show maximum contrast, so they cannot be compared directly. Three of the patterns (VAN3, VAN4, KIN) produce reasonable results with only a few missing points.

VAN2 shows larger spots of missing points due to its non-uniformity: in some areas of the PSM the density of the spots is very low which can lead to problems in the stereo matching process.

Apart from some missing points all of the four PSMs produce reasonable results with regard to the depth measurement of by far most of the points. However, VAN2, VAN3, and KIN show a few very light dots indicating extreme disparity outliers. These outliers are due to stereo mismatching, presumably because the minimum Hamming distance of these patterns is very low (≤ 1) and thus the patterns are not very robust against errors at some positions. Only VAN4 (with a minimum Hamming distance of 3) is completely free of such outliers.

Table 3. Accuracy of the simulated depth recordings with 'VAN2', 'VAN3', 'VAN4', and 'KIN', respectively. The best values in each category are printed in bold face.

SIBTPSM	VAN2	VAN3	VAN4	KIN
RMS deviation [mm]	26.63	11.70	5.03	8.23
Max. deviation [mm]	672.8	679.4	23.7	719.4

This is in line with the results of the accuracy analysis shown in table 3. The more uniform patterns VAN4 and KIN yield a better overall RMS accuracy than the less uniform patterns VAN2 and VAN3. The best results are obtained with the VAN4 pattern.

5.2 Live Tests

The experimental stereo setup for the live test consisted of two coplanar NIR cameras (IDS UI-3240ML-NIR-GL) with a longpass 800nm filter (MIDOPT FIL LP800/34) mounted on a rig with a relative distance (baseline) of 168mm. For the structured light we used an 830nm diode laser (LaserComponents, 40mW) in the middle of the baseline together with our DOEs described in table 2.

As we did not have access to a separate DOE producing the KIN pattern, we used an ASUS Xtion Pro Live stereo camera for projecting the KIN structured light pattern, without using its camera capabilities. The integrated diode laser of this device has significantly more power than the laser used with the other DOEs. With a LaserComponents XLP12 power detector we measured a light power past the KIN DOE that was more than twice as high as the light power behind the VAN x DOEs with the 40mW laser. It can therefore be assumed that the diode laser of the ASUS device has a minimum power of 80mW. Thus a direct comparison of the results is at least problematic since the signal-to-noise ratio with the KIN configuration is better than the others by at least the factor 2. This should be kept in mind when evaluating the results.

The cameras were directed to a scene consisting of a sloping wall with an angle of $\approx 45^\circ$ relative to the baseline, and two cardboard boxes in the foreground. The closest point of the right box was about 2m from the cameras, and the farthest point of the wall was at a distance of about 4m. Fig. 5a shows an infrared image of the scene from the perspective of the left camera.

The cameras were calibrated using a chessboard calibration pattern, and the depth calculations were performed by stereo matching with the undistorted and rectified camera images using again the semi-global stereo block matching algorithm `StereoSGBM` of the OpenCV library³² in the mode `MODE_SGBM_3WAY`. For this real-life configuration, however, we used the following more 'realistic' parameters, including a moderate filtering: `minDisparity = 16`, `numDisparities = 128`, `blockSize = 3`, `P1 = 0`, `P2 = 300`, `disp12MaxDiff = 1`, `preFilterCap = 63`, `uniquenessRatio = 10`, `speckleWindowSize = 100`, `speckleRange = 32`.

The results are presented in fig. 5. As the objects in the scene have low texture of their own the stereo camera yields very poor results without structured light illumination (see fig. 5b). With structured light, all the four patterns (VAN2, VAN3, VAN4, KIN) give very good results in the near field where the signal-to-noise ratio of the projected pattern is sufficiently high. Only the KIN pattern produces very few missing or error points in the lower left part of the wall and on the frontal box. The reasons for this could be the relatively large minimum uniqueness window of the KIN pattern (9x9) and the large distortion of the KIN DOE.²⁷

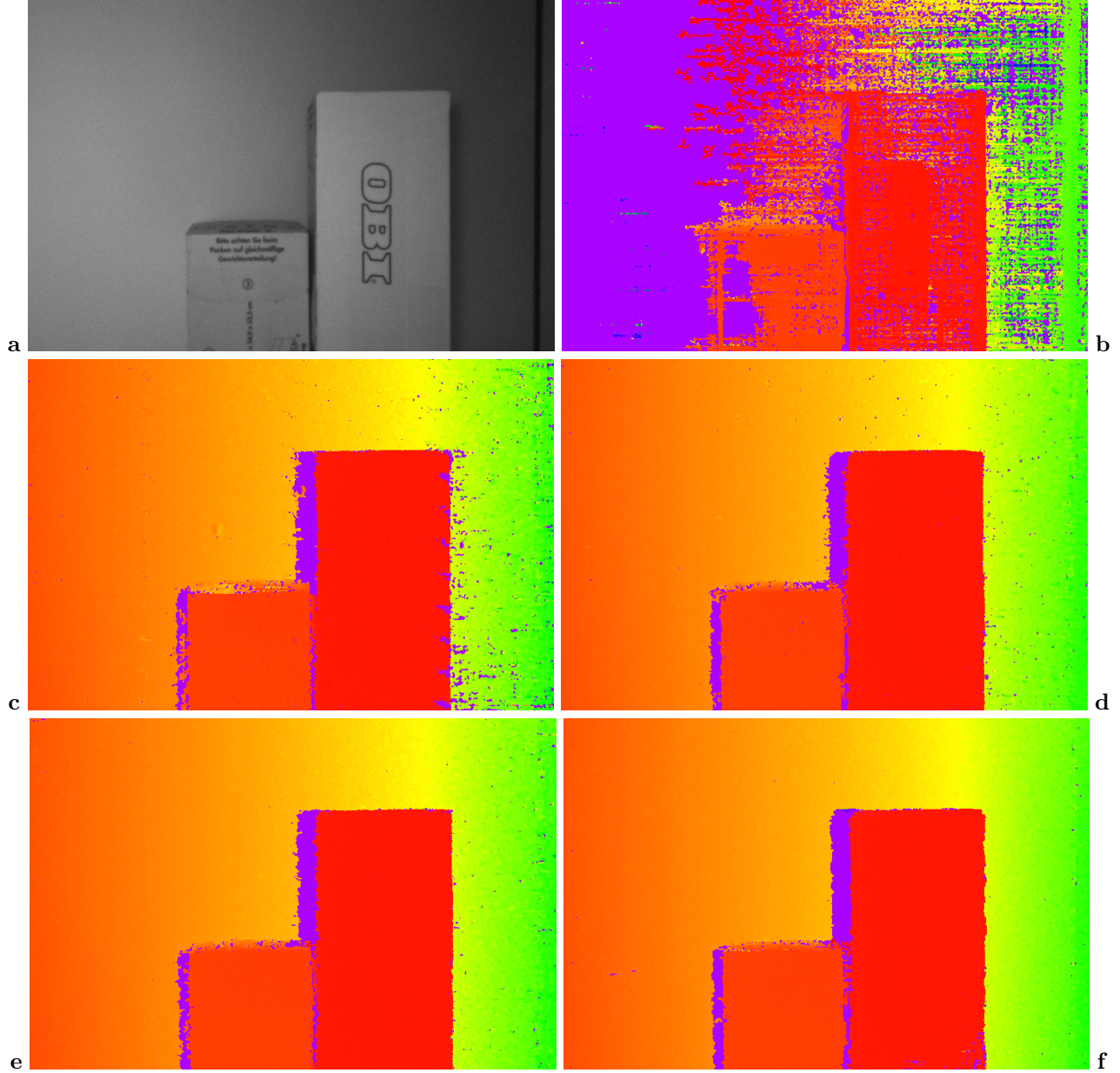


Figure 5. Depth recordings of a real scene with a sloping wall in the background and two boxes in the foreground using semi-global block matching with a stereo camera. (a) shows the left camera's infrared image of the scene. (b) is the depth image of the scene without structured light. The images (c), (d) and (e) were recorded using a 40mW NIR laser (830nm) and the DOEs (c) DE-R 332 (VAN2), (d) DE-R 372 (VAN3), and (d) DE-R 375 (VAN4), respectively. For image (f) the ASUS Xtion Pro Live with a $\approx 80\text{mW}$ NIR laser (830nm) and the KIN DOE was used for the structured light illumination.

As expected, the number of errors generally increases with the distance from the cameras. Particularly VAN2 shows a larger number of errors or missing points in the far field that can be attributed to the relatively non-uniform spot distribution of the VAN2 pattern. The patterns VAN3, VAN4, and KIN, however, yield robust depth measurement even in the far field up to the maximum distance of 4m in this experiment. Interestingly, the results of VAN3 and VAN4 in the far field are almost as good as the results of the KIN setup although the latter was projected with roughly the double laser power and thus a much better signal-to-noise ratio. As in the

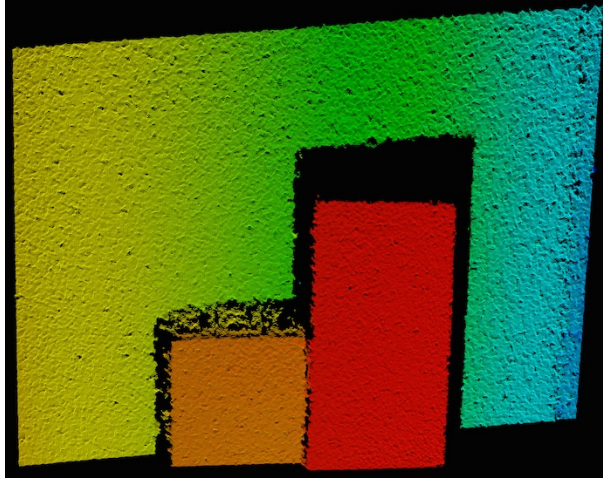


Figure 6. Point cloud of the real scene reconstructed using the VAN4 pattern.

simulation test, the best overall results are obtained with the VAN4 pattern, probably due to its good Hamming properties, the good uniformity, and the almost distortion-free DOE projection. The reconstructed point cloud of the scene using this pattern is shown in fig. 6.

6. CONCLUSIONS

In this work we first identified properties of patterns for the structured light illumination of a 3D sensor that are projected using a laser and a diffractive optical element (DOE), and found that such a pattern should consist of tiles of a (central) symmetric, isolated, binary, toroidal perfect sub-map (SIBTPSM). We then discussed the generation of such patterns and selected four SIBTPSMs for further evaluation. For this, we defined several quality metrics: the width and the height of the SIBTPSM, the size of the minimum uniqueness window, the minimum and the average Hamming distance, the spot density, and the n-uniformity. We also discussed the physical properties of DOEs and their effects on the pattern projection. The sample patterns and the corresponding quality metrics are evaluated with a simulation test and with a stereo camera live test.

Our results suggest that the uniqueness window size, the minimum and average Hamming distances, and the uniformity of the SIBTPSMs have strong impact on the quality of the depth measurement and thus the 3D reconstruction. A small uniqueness window is also beneficial for a fast decoding of the projected pattern. The PSM width is important, too, as it corresponds with the maximum possible disparity that can be detected unambiguously. For DOEs with low distortion, the PSM height is of minor importance. The same holds true for the spot density, which is (due to the restrictions of SIBTPSMs) varying only in a very small range anyway.

In our experiments we used a relatively weak 40mW laser for a reliable depth measurement of up to 4m and more with two of our SIBTPSMs, VAN3 and VAN4. The diffractive elements would permit significantly stronger laser sources of up to several 100mW optical power to be used in the application, if this would be required for applications with larger working distances, and the resulting projectors would still be eye-safe. The virtually distortion-free projection based on a single DOE approach contributes to the quality of the depth reconstruction, particularly for SIBTPSMs with limited height.

The developed SIBTPSMs can be used for other applications as well in which different parameters like field-of-view angles, laser wavelengths, aspect ratios or local grid sizes are required, because the projection by DOEs offers a reasonable degree of flexibility in this regard. The results of our work should encourage researchers currently still using the common 3D sensors based on the PrimeSense platform to go for application specific diffractive structured light projectors that are based on competitive SIBTPSMs and potentially can offer better performance for their applications.

REFERENCES

- [1] Geng, J., “Structured-light 3d surface imaging: a tutorial,” *Advances in Optics and Photonics* **3**, 128–160 (2011).
- [2] Spoelder, H. J. W., Vos, F. M., Petriu, E. M., and Groen, F. C. A., “Some aspects of pseudo random binary array-based surface characterization,” *IEEE Transaction on Instrumentation and Measurement* **49**(6), 1331–1336 (2000).
- [3] Gedat, E., Fechner, P., Fiebelkorn, R., and Vandenhouten, R., “Multiple human skeleton recognition in RGB and depth images with graph theory, anatomic refinement of point clouds and machine learning,” in [*IEEE International Conference on Systems, Man, and Cybernetics (SMC)*], 000627–000631 (2016).
- [4] Vandenhouten, R. and Wendlandt, O., “High-precision optical position measurement in indoor environments,” *Optik & Photonik* **10**(3), 24–27 (2015).
- [5] Yamazaki, S., Nukada, A., and Mochimaru, M., “Hamming color code for dense and robust one-shot 3D scanning,” in [*Proceedings of the British Machine Vision Conference*], Hoey, J., McKenna, S., and Trucco, E., eds., BMVA Press (2011).
- [6] Alhwarin, F., Ferrein, A., and Scholl, I., “IR stereo kinect: Improving depth images by combining structured light with IR stereo,” in [*PRICAI 2014: Trends in Artificial Intelligence*], DN Pham, S. P., ed., *Lecture Notes in Computer Science* **8862**, Springer (2014).
- [7] Cai, Z., Liu, X., Peng, X., Yin, Y., Li, A., Wu, J., and Gao, B. Z., “Structured light field 3d imaging,” *Opt. Express* **24**, 20324–20334 (2016).
- [8] Morano, R. A., Ozturk, C., Conn, R., Dubin, S., Zietz, S., and Nissanov, J., “Structured light using pseudorandom codes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**(3), 322–327 (1998).
- [9] Goodman, J. W., [*Introduction to Fourier optics, 3rd edition*], Roberts and Company Publishers (2005).
- [10] MacWilliams, F. J. and Sloane, N. J. A., “Pseudo-random sequences and arrays,” *Proc. of the IEEE* **64**(12), 1715–1729 (1976).
- [11] Chung, F., Diaconis, P., and Graham, R., “Universal cycles for combinatorial structures,” *Discrete Mathematics* **110**, 43–59 (1992).
- [12] Pages, J., Salvi, J., and Forest, J., “A new optimised De Bruijn coding strategy for structured light patterns,” in [*Proceedings of the 17th International Conference on Pattern Recognition, 2004*], (2004).
- [13] Cock, J. C., “Toroidal tilings from De Bruijn-good cyclic sequences,” *Discrete Mathematics* **70**, 209–210 (1988).
- [14] Ozturk, C., Nissanov, J., and Dubin, S., “Generation of perfect map codes for an active stereo imaging system,” in [*Proceedings of the 1996 IEEE Twenty-Second Annual Northeast Bioengineering Conference*], (1996).
- [15] Bruckstein, A. M., Etzion, T., Giryas, R., Gordon, N., Holt, R. J., and Shuldiner, D., “Simple and robust binary self-location patterns,” *IEEE Transactions on Information Theory* **58**(7), 4884–4889 (2012).
- [16] Etzion, T., “Sequence folding, lattice tiling, and multidimensional coding,” *IEEE Transactions on Information Theory* **57**(7), 4383–4400 (2011).
- [17] Kapinya, J. B., *Evolutionary Computing Solutions for the de Bruijn Torus Problem*, Master’s thesis, Vrije Universiteit Amsterdam (2004).
- [18] Claes, K. and Bruyninckx, H., “Robot positioning using structured light patterns suitable for self calibration and 3D tracking,” in [*International Conference on Advanced Robotics*], (2007).
- [19] Maurice, X., Graebing, P., and Doignon, C., “A pattern framework driven by the Hamming distance for structured light-based reconstruction with a single image,” in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], IEEE (2011).
- [20] Wijenayake, U. and Park, S.-Y., “Dual pseudorandom array technique for error correction and hole filling of color structured-light three-dimensional scanning,” *Opt. Eng.* **54**(4) (2015).
- [21] Wijenayake, U., Baek, S.-H., and Park, S.-Y., “An error correcting 3D scanning technique using dual pseudorandom arrays,” in [*Proceedings of the Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization & Transmission*], (2012).

- [22] Wijenayake, U., Choi, S.-I., and Park, S.-Y., “Combination of color and binary pattern codification for an error correcting m-array technique,” in [*Proceedings of the Canadian Conference on Computer and Robot Vision (CRV)*], 139–140 (2012).
- [23] Albitar, C. and Graebing, P., “Robust structured light coding for 3D reconstruction,” in [*IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007*], IEEE (2007).
- [24] Wijenayake, U. and Park, S.-Y., “An M-array technique for generating random binary pattern based on a connectivity constraint,” in [*Workshop on Image Processing and Image Understanding*], (2012).
- [25] Shpunt, A. and Pesach, B., “Optical pattern projection.” Patent Application No. US20100284082 (2010).
- [26] Reichinger, A., “Kinect Pattern Uncovered.” azt.tm’s Blog <https://azttm.wordpress.com/2011/04/03/kinect-pattern-uncovered/> (2011). (Accessed: 8 April 2017).
- [27] Nordmark, A., *Kinect 3D Mapping*, Master’s thesis, Linköpings Universitet (2012).
- [28] Hermerschmidt, A. and Krüger, S., “Diffractive optics: Diffractive elements generate complex light patterns.” www.laserfocusworld.com <http://www.laserfocusworld.com/articles/print/volume-42/issue-12/features/diffractive-optics-diffractive-elements-generate-complex-light-patterns.html> (2006).
- [29] Skeren, M., Richter, I., and Fiala, P., “Iterative fourier transform algorithm: comparison of various approaches,” *Journal of Modern Optics* **49**, 1851–1870(20) (September 15, 2002).
- [30] Hermerschmidt, A., Krüger, S., and Wernicke, G., “Binary diffractive beam splitters with arbitrary diffraction angles,” *Opt. Lett.* **32**(5), 448–450 (2007).
- [31] Clausnitzer, T., Kämpfe, T., Kley, E.-B., Tünnermann, A., Peschel, U., Tishchenko, A. V., and Parriaux, O., “An intelligible explanation of highly-efficient diffraction in deep dielectric rectangular transmission gratings,” *Opt. Express* **13**(26), 10448–10456 (2005).
- [32] Alekhin, A. et al., “Opencv 3.2.0.” <http://opencv.org/opencv-3-2.html> (December 2016).