Vol.2017-CG-168 No.33
Vol.2017-DCC-17 No.33
Vol.2017-CVIM-209 No.33
2017/11/9

Data-based 3D shape reconstruction using light field constructed by multiple projectors

Yuki Shiba[1,a)] Satoshi Ono[1,b)] Ryo Furukawa[2] Shinsaku Hiura[2] Hiroshi Kawasaki[†1]

Abstract: Combination of a pattern projector and a camera is widely used for 3D measurement. From the pattern projection, various kinds of depth queues are extracted from the captured image such as disparities for active stereo, projector defocus for depth from defocus, or intensity variations for photometric stereo. To increase the depth queue information, while widening the working space and reducing the occlusions, increasing the number of projectors can be a promising solution. However, multiple projectors form a complicated light field where the captured image is difficult to process analytically. In this paper, we use the configuration of multiple projectors and a camera. To process the complicated light field, matching-based analysis is applied. In the preprocess of the 3D measurement, virtual sample images of planar board with various depths are generated with CG techniques. Then, their image features of small patches are extracted with PCA. In the 3D measurement step, the same image features of patches of the captured image are extracted and compared with the sample images. By utilizing the dimensional reduction of PCA and ANN search algorithm, the matching is processed effectively. Since our approach is sampling based, the proposed technique can handle arbitrary patterns for projection.

Keywords: Computational Photography, 3D scan, Light filed, Structured Light

1. Introduction

By the recent progress on technology and cost effectiveness of video projector, they are now wide-spread even in a household and used not only for a image presentation purpose just on a white screen, but also for a projection mapping on complicated shapes and/or textured objects, 3D reconstruction etc. Among them, 3D scan system using a projector and a camera is one of important and promising topics. Previously, such system usually consists of a single projector and a single camera. 3D Shapes are then reconstructed by either stereo, depth from defocus (DfD) or photometric stereo techniques. On the other hand, multi-view stereo system using only cameras becomes popular for the increasing demand on capturing a large scale scene and/or entire shape of moving objects, typically human activity in the scene. Recently, several researchers are trying to increase the number of the projector as well as the camera. However, there is one severe problem on increasing the projector, *i.e.*, multiple patterns projected onto the same object surface from multiple projectors interfere each other and it is difficult to analytically handle such overlapped patterns. A simple solution is to use different colors for each projector to decompose them to be used as independent patterns. However, only three colors can be used with commercial products and color cross-talk is also a problem.

The key idea to overcome the problems is that multiple projected patterns in the space can be considered as an implicit light field [5], [22] constructed by the bundle of light rays. Our solution is to store the light field as a reference and conduct matching between captured image to reconstruct 3D shapes of the scene. Note that since the light field includes all the optical effects, such as defocus of projector, lens distortion, etc, there is a high possibility that those non-linear phenomena can be solved with our approach. The method can be divided into two phases: a learning phase and a shape estimation phase. At the learning phase, since there is no medium which reflect the light ray in the air and the patterns are invisible and unobservable, we put a planar board in the scene and capture the reflected patterns on it. Since data size of the light field is tremendous, we just store a subset of it, *i.e.*, only a single value at each 3D point is stored where originally there are two rotation angles ϕ and θ. Such reduction is compensated by interpolation in parametric space in our method. Further, we conduct a principal component analysis (PCA) to eliminate the redundant dimension to make compact data.

At the shape estimation phase, depth of the object is estimated by matching the observed pattern and the reference pattern which is stored at the learning phase. To reduce the calculation time we use approximated nearest neighbour (ANN) search combined with a Belief Propagation (BP) to efficiently remove noise in final output. Such a light field based reconstruction can be understood as a combination of stereo and DfD simultaneously. A simple idea using a specially designed projector equipped with coded aperture is proposed [8]; however, utilization of multiple projectors to construct light field is not mentioned.

The main contribution of the paper can be summarized as follows.

[1] Kagoshima University, Kagoshima, ,Japan, Chiyoda, Tokyo 101–0062, Japan
[2] Hiroshima City University, Hiroshima, ,Japan
[†1] Presently with Kyushu University
[a)] sc112031@ibe.kagoshima-u.ac.jp
[b)] ono@ibe.kagoshima-u.ac.jp

(1) Unlike the previous multiple projector system for 3D scan, our technique has no severe limitation on setup of multiple projectors in the scene.
(2) Since the technique is based on data centric approach, non-linear effect such as defocus blur of projector can be naturally handled.
(3) Unlike the common data centric approach which requires huge data storage and long computational time, our technique can drastically reduce the computational time by PCA and ANN.

2. Related works

Active 3D measurement systems utilize various types of keys of projected patterns that appear in the captured images and show the depth information of the scenes. Such keys include correspondences between the image and the projected patterns used for active stereo systems, projector defocus of the patterns [24], or projector defocus modulated by coded apertures of the projector [8].

For estimating correspondences between the image and the pattern numerous approaches have been proposed. Typically, they are categorized into temporal and spatial coding techniques [16]. Temporal coding active stereo methods [17] have a long history of practical use, since pixel-by-pixel dense capture of the depth images can be realized with relatively simple algorithms. However, many of such systems generally assumes static scenes because multiple images are required for a single measurement. To cope with this problem, spacial coding systems, which require a single image for a single measurement have been actively researched [1], [6], [9], [10], [11], [18], [21], [25].

Spacial coding techniques of active stereo methods use various types of pattern features, such as color information [1], [9], [18], [25], or geometrical characteristics [6], [11], [21]. The problem of color information features is that the results may be affected by the object colors or textures. The geometrical characteristics also have problems such as difficulty of matching between the image and the pattern because of deformation caused by disparities, or projector defocus.

Projector defocus has been a big problem of 3D scanners that utilize pattern projection. Since normal projectors are built with large apertures for brightness, the depth-of-field is much narrower than that cameras. Some of the previous works deals with the problem by using the projector defocus as depth cues [8], [24].

In most of the above methods, the projected depth cues are analyzed by the process that specifically designed for each patterns. Recently, some researchers use matching based approaches, where the captured images are directly compared with the projected patterns [3], [7], or the sampled patterns captured for each depths [8]. In this paper, we proceed this approach one step further. In the proposed method, we no longer assume a specific pattern, and the algorithm is universal and independent of the projected pattern. Naturally, specialized modeling for the projected pattern is no longer necessary.

For such a matching-based approach, comparing image patches allowing the variations caused by depth, texture, and normal changes is a key technique. One of the commonly-used methods for such purpose is matching in eigen-space domain (representation by PCA), which has been used for face recognition [2], [19], or object recognition[13].

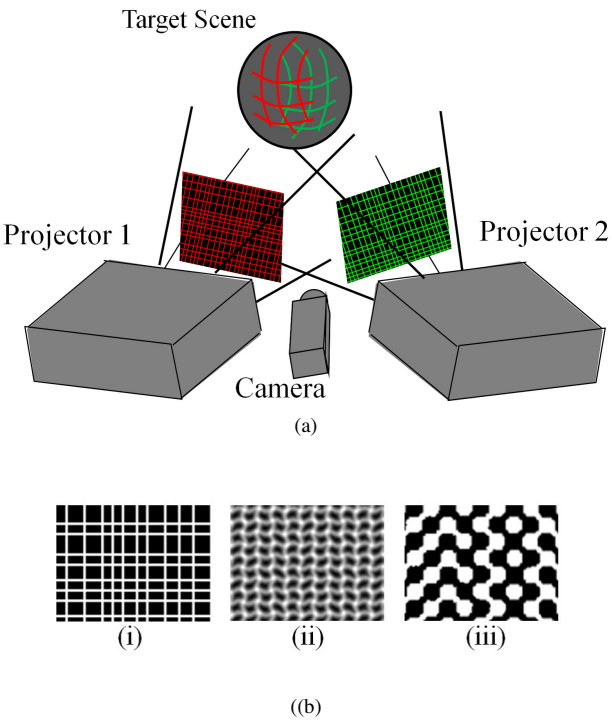


Fig. 1 (a) System configuration of projector and camera system. (b) Several patterns for one-shot scan. (i) random grids, (ii) ,wave grids [15], and (iii) wave grids2[23].

About projecting light field using multiple projectors, several systems have been proposed [4], [5], [14]. Jurik *et al.* proposed an array of a large number of laser projectors for generating a light field observed directly by human eyes [5]. Nagano *et al.* constructed a similar projector array except that a lenticular screen with anisotropic reflection is placed at the center of the projector array [14]. Hirsch *et al.* proposed a method using lenticular lenses inside the optics of a video projector [4].

3. Overview

3.1 System configuration

We assume the system which consists of multiple pattern projectors and a camera as shown in Fig. 1(a). The camera and the projector are assumed to be calibrated (*i.e.*, the intrinsic parameters of the devices and their relative positions and orientations are known). Since each projector casts a static pattern, no synchronization is required and it is a important feature to realize using multiple projectors simultaneously. Further, it is suitable to acquire a 3D shape of dynamic scene. In terms of the projecting pattern, since we do not know what kind of pattern is suitable for constructing light field to better reconstruct the 3D shape of the scene, we tested several well known patterns in our experiment, grid pattern, wave grid pattern, and patterns proposed by previous works as shown in Fig. 1. Similarly, we assign different colors to each projector, respectively.

We explain the projected pattern used in our system in Sec. 3.2 and show the overview of our algorithm in Sec. 3.3.

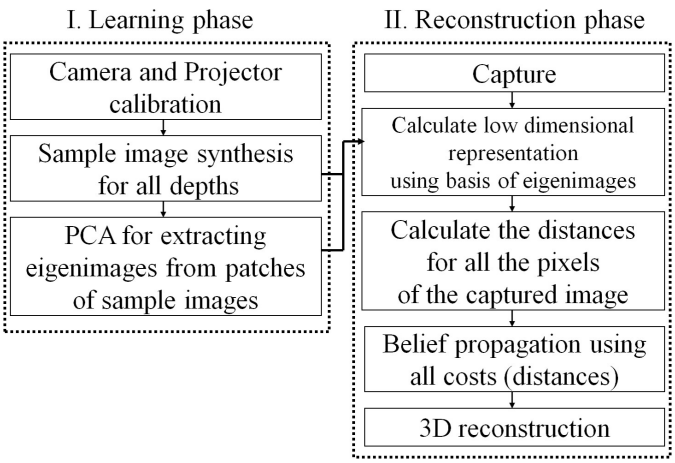


Fig. 2 Flow chart of our shape reconstruction algorithm.

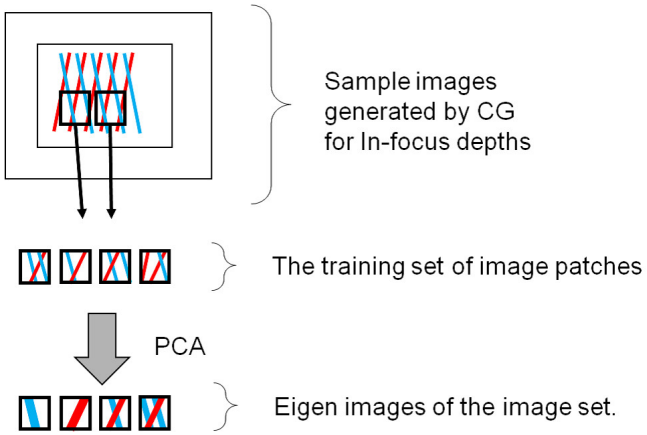


Fig. 3 Extracting image features of sample images.

3.2 Projection pattern

It is important to use suitable patterns to create an efficient light field in order to realize the best performance on the purpose, *i.e.*, accurate and robust 3D reconstruction. Since there is no previous researches on a light field based 3D reconstruction, we used several well-known patterns for one-shot scan, *i.e.*, spatial encoding patterns, in our experiment. The patterns we used in the experiment is shown in Fig. 1. Fig. 1(i) is a random grid pattern, Fig. 1(ii) is a wave grid pattern [15], Fig. 1(iii), is also a well known pattern for single color one-shot scan [23].

3.3 Algorithm overview

Our method consists of two phases: pattern learning phase and 3D reconstruction phase as shown in Fig 2.

In the pattern learning phase, first, the projector and the camera are calibrated. Then, by using the calibration parameters, virtual images are synthesized by assuming a planer board placed at a certain depth, pattern projected by a virtual projector and captured by a virtual camera. The images are synthesized by changing the depth. Finally, parameters and eigen images are estimated by PCA in order to map the input image to low dimensional values, typically 30D in our case and the values are stored in database.

In the 3D reconstruction phase, input image is converted to low dimensional representation by calculating coefficient for eigen images. Then, for each position at the captured image, costs for all the depths are acquired by calculating the inner product of the coefficient vector of the input image and the reference images at each depth in the stored data. The created cost volume is then used to apply BP in the final step. To reduce the cost of calculating the cost volume, ANN search is used to calculate the cost for just top n depths in our method. Once the depth values of all the pixels are estimated, 3D shapes are recovered using the camera and projector calibration parameters.

4. Active stereo based on synthetic light field

4.1 Light field creation by virtual projector and camera system

In the pattern learning phase, we need to store the entire light field created by multiple projectors. Since the light field created by multiple projectors is not static because some patterns are occluded by objects in the scene, such variations should be captured and stored. In addition, since the light field itself is not visible, some objects should be placed in the scene to reflect the light ray to be observed by the camera and the precise position and the shape of the object should be also recorded. Conducting such complicated processes to store the light field is obviously impractical. In this paper, to efficiently store the entire light field, we create a virtual light field in the computer to synthesize the image where virtual planar board is placed in the virtual space and captured by the virtual camera.

It is simply achieved by GPU based implementation. We prepare the pattern images for each projectors as well as intrinsic and extrinsic parameters for each projectors and cameras in advance. Then, scene images are rendered by using pixel shader where color values are sampled from each projector calculating the line of sights from the object surface point to each projector and blending them. To capture the light field considering the occlusion effect, we turn on and off each projector based on all the combination of them.

4.2 Low dimensional representation by PCA

In the proposed method, for each pixel of the captured image, image patch around the pixel is compared to the sample images generated by virtual projector. This process is time-consuming if implemented naively, and should be reduced for practical use. To deal with this problem, we use low dimensional representation by PCA (*i.e.*, eigenspace representation).

Utilizing PCA for reducing dimension of image data has been used in computer vision for a long time. For example, PCA-based representation is used for facial image recognition or analysis for representing or separating changes of illuminations or facial expressions [2], [19], for object recognition with various image changes caused by variations of 3D viewpoints [13], or for fast image matching realized by reducing the dimension of raw image data vector [20].

Figure 3 shows the flow to extract the feature basis for the projected pattern. For applying PCA, we must first collect the training set from the sample images for calculating the eigenspace of the data set. In the proposed method, the sample images are obtained in the calibration process, where the images of the fronto-

parallel planes are captured by a real setup of a projector-camera system or generated by virtual projector simulation. From the sample images, samples of image patches with the same patch sizes are extracted.

The image patches are first represented as column vectors of $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N$ by simple rasterization. If the image patches are M by M pixels, the dimension of the column vectors are M^2. Here, we use $L = M^2$ for simplicity. The average image patch is calculated by $\bar{\mathbf{p}} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{p}_k$, and the deviations from the average data are $\mathbf{q}_k = \mathbf{p}_k - \bar{\mathbf{p}}$. A set of orthonormal basis for representing $\mathbf{q}_k$ can be calculated with PCA.

In normal PCA, eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_L$ of the $L \times L$ covariance matrix

$$\mathbf{C} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{q}_k \mathbf{q}_k^\top = \mathbf{A}\,\mathbf{A}^\top, \tag{1}$$

where $\mathbf{A} = [\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_N]$, are used for the orthogonal basis set. However, in computer vision problems, often $N < L$, then, we can use eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ of the $N \times N$ matrix $\mathbf{L} = \mathbf{A}^\top \mathbf{A}$ for forming the basis set to save the computational cost of eigenvector calculation[19]. The basis vectors, then, can be calculated by $\mathbf{u}_i = \sum_{k=1}^{L} (\mathbf{v}_i)_k\ (\mathbf{q}_k)$ for $i = 1, \cdots, N$, where $(\mathbf{v}_i)_k$ is the k-th element of vector $\mathbf{v}_i$.

From the obtained basis, the representation of a new image patch $\mathbf{r}$ is $(w_1\ w_2\ \cdots\ w_N)^\top$, where $w_i = \mathbf{u}_i^\top(\mathbf{r} - \bar{\mathbf{p}})$. Let eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ be sorted by the descending order of the associated eigenvalues. Then, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_L$ are aligned in the order of optimally representing the training set $\{\mathbf{q}_1, \cdots, \mathbf{q}_N\}$ for minimizing the sum of errors of l^2 norm. Thus, if the image patch $\mathbf{r}$ is similar to the training set, $(w_1\ w_2\ \cdots\ w_L)^\top$ where $L \leq N$, is a good L-dimensional representation of $\mathbf{r}$. The process of deciding the basis set using PCA can be regarded as a process of learning the image features for representing the training set of the pattern.

Fig. 4 shows the eigenvector basis for different pattern projections. The top row images are extracted for a grid pattern by aprojector, and the middle row images are for grid patterns by two projectors, and the bottom row images are for a grid patterns by three projectors. It is shown that the basis represents the features of the training image set for different types of light field.

In the calibration step, we calculate the low-dimensional (L-D) representations for the sample patch images. In the measurement step, we calculate the L-D vectors for the image patches around each pixel. These patches are matched with the images sampled for each depth in the L-D vector space.

4.3 Efficient depth estimation by ANN and MRF

By using PCA, high dimensional patch information is efficiently reduced to low dimensional vector. Those vectors are stored at the learning phase. At the depth estimation phase, first, input images are converted to low dimensional representation by calculating the coefficient of each eigen image. Then, maximum similarity between input and reference vector is searched. Although the vector size is just 10 to 40, it still requires heavy computational cost to find the maximum if all the similarity values for all the depths are calculated. In this paper, to reduce the calculation cost, we use approximate nearest neighbor (ANN) search [12]. It reduces the processing time with a little sacrifice in accuracy.

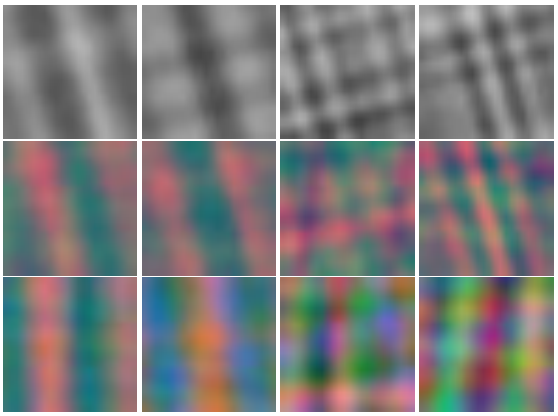


Fig. 4 Samples of eigenvectors visualized in image format (eigenimages) for grid patterns. The top row images are eigenimages for a grid pattern to used one projector, and the middle row images are for grid patterns by two projectors, and the bottom row images are for grid patterns by three projectors. From left colums to right, the 2nd, 3rd, 11th and 12th eigenimages in the order of eigenvalues.

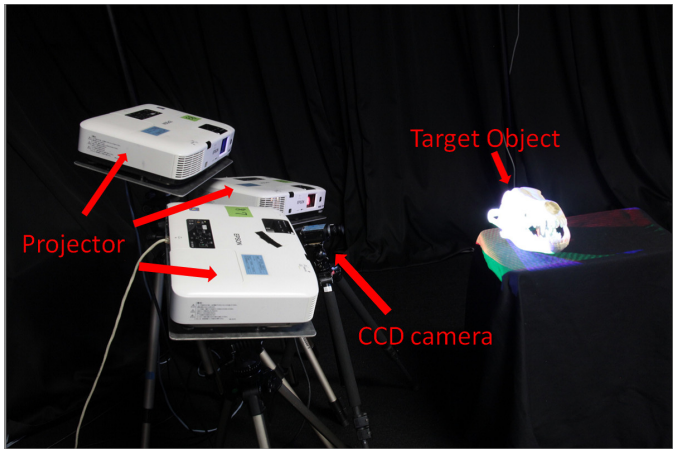


Fig. 5 Experimental system with a camera and a video projector.

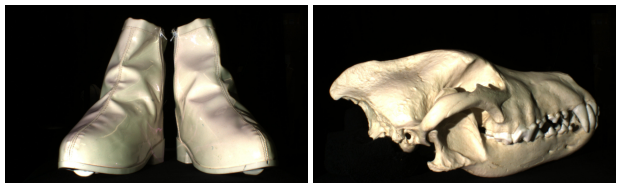
Although initial depth estimation result consists of many noises with wrong depth, such wrong depths are efficiently removed and corrected by MRF based approach. For MRF, cost volume is usually required, however, ANN originally returns just single cost for maximum similarity. Since top 10% of high similalrity includes the correct depth with 90% possibility based on our survey, we modified ANN to output top 10% depth values with cost (inverse similarity). We use BP to solve MRF.

5. Experiments

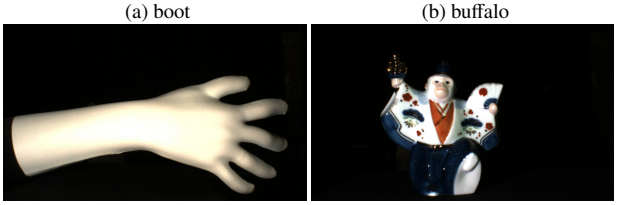
Experiments were conducted to examine the effectiveness of the proposed 3D measurement system. First, the comparison of the reconstruction accuracy between the systems where the number of the projectors varies from one up to three was made (Sec. 5.2). Next, dimension reduction of feature values using PCA was evaluated (Sec. 5.3). Finally, the experiments using various projected patterns were performed (Sec. 5.4).

5.1 Experimental setup

Fig. 5 shows our actual experimental setup. The number of projectors were varied depending on the experiment's purposes.



(a) boot (b) buffalo



(c)hand (d)monke

Fig. 6 Measurement object

Reference images were synthesized and stored by the virtual projector and camera system using the pattern information and the calibration parameters. For synthesis, the virtual planar object was placed in the scene and its position is moved along z direction with 1mm interval, and a virtual camera captured reference images at each depth.

Because of the experimental environment condition, we put a close-up lens to change the scale as to be 1/3 of real length. With this scale, the motion range of the screen is 150mm-625mm from the projector and the camera, in-focus distance is 250mm ±100mm for the projector, the reference plane capturing interval is 0.5mm.

Four arbitrary shape objects shown in Figure 6 were used in this experiment: mannequin hand, buffalo head with complicated shape, boots with glossy surface and monkey with glossy, painted surface. Ground truth of these object shape were defined by measuring them with gray-code scan.

5.2 Evaluation of multiple projectors

The first experiment was performed in order to evaluate the effectiveness by increasing the number of applied projectors from one, two and three. The wave pattern which has relatively high reconstruction accuracy with unicolor was used. The unicolor wave pattern was projected by each projector. Monochrome pattern was projected when one projector was used. Red and green patterns were projected for two projectors. In the case of three projectors, Blue was added to the two-projector setup. The window size was 32*32 for the coarse process of NCC and 24*24 for fine process, respectively, to match captured images with the referenced image.

Figure 7 shows the results. As the number of projectors increases, the accuracy improved in all four objects. This is because more projectors could make more complicated patterns that produces better depth queue information.

5.3 Evaluation of PCA representation

Next, to evaluate the dimension reduction of feature values using PCA, the following two methods were compared: 1) verify the data directly with the referenced image using NCC, 2) verify the data after dimension reduction by PCA.

In this experiment, three projectors were used. Four projec-

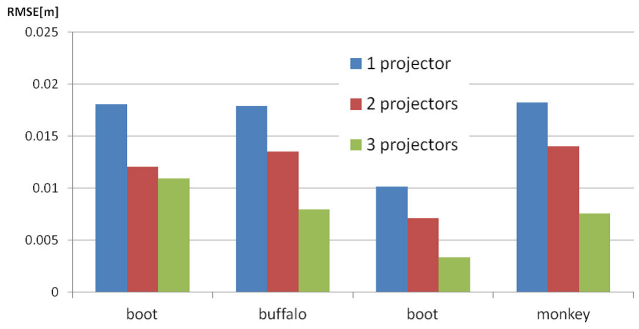


Fig. 7 Comparison on RMSE with varying the number of projectors.

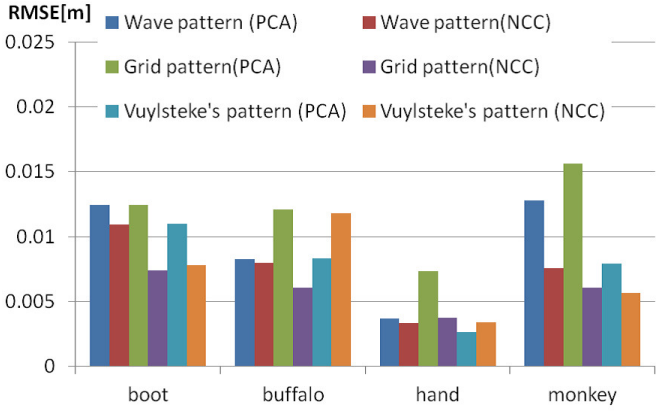


Fig. 8 Comparison on RMSE with varying reconstruction method (NCC and PCA), and projection pattern.

tion patterns were used: Vuylsteke's pattern [23], random grid, and wave [15]. The window size was 32*32 for coarse level and 24*24 for fine level for NCC, as the same condition with the experiment of previous section and 24*24 for PCA. In case of PCA, principal components up to the 30th level was applied. To evaluate the results, the comparison of RSME and that of reconstruction time were made, respectively.

Fig. 8 shows the comparison result for RMSE, and figure 9 shows the comparison results of reconstruction time, respectively. The result shows that there is no significant difference for RSME between NCC- and PCA-based methods, which means that the dimension reduction by PCA did not sacrifice the reconstruction accuracy. On the other hand, the result of reconstruction time clearly shows PCA reduced the reconstruction time drastically. PCA requires just 30 dimensional features, whereas NCC requires $24 \times 24 = 512$ dimensions. In addition, the computational complexity of the reconstruction time of NCC and PCA is $O(n)$ and $O(\log n)$ against the number of reference images n, respectively. Therefore, the difference of the reconstruction time becomes larger in the case that higher depth resolution with more reference images.

5.4 Reconstruction of textured and curved surface object with various patterns

In the final experiment, the dependency on the projected patterns were dissected with the results obtained in the experiments in Sec. 5.3. Figures 10 through 12 show depth maps of the reconstruction results with tested four projection patterns. We can

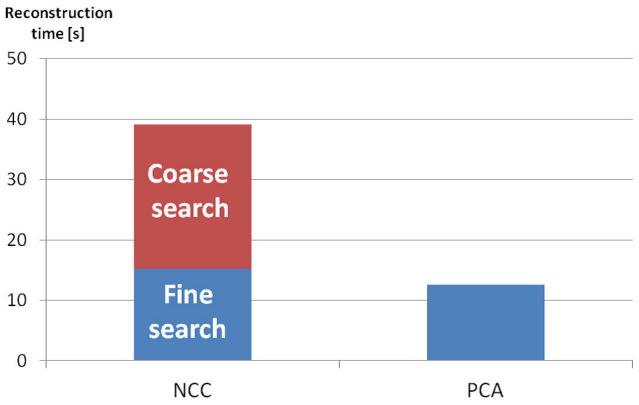
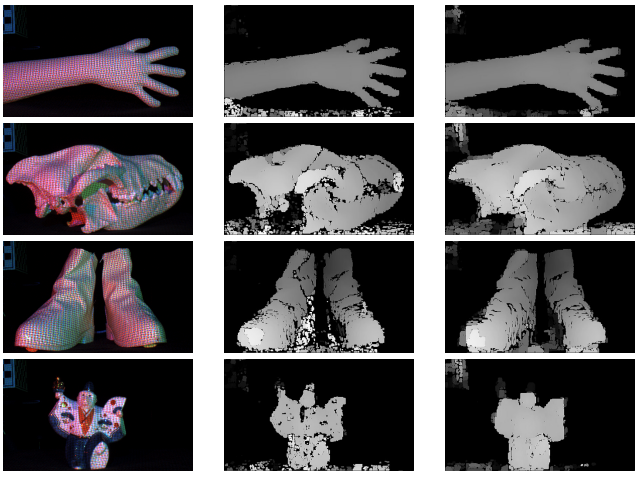
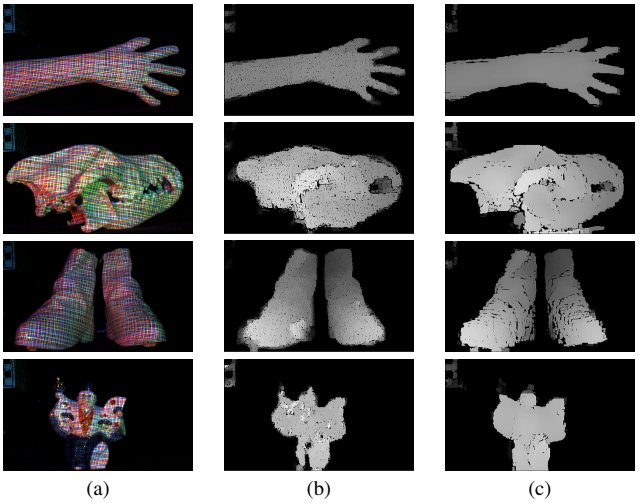


Fig. 9 The reconstruction time by PCA and NCC.



(a) (b) (c)

Fig. 10 Reconstruction results of textured and curved surface object with wave pattern. (a) input, (b) PCA, (c) NCC

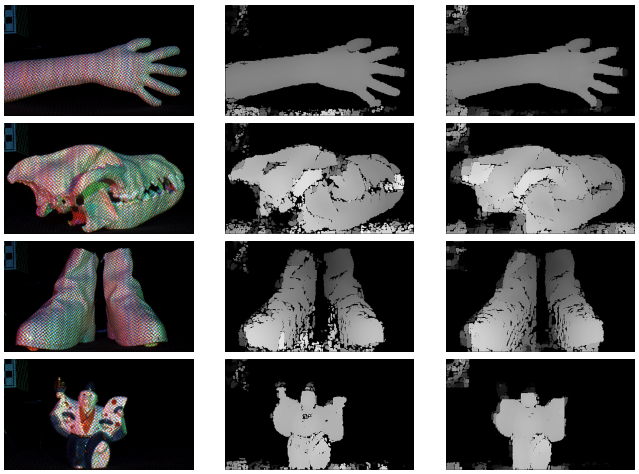


(a) (b) (c)

Fig. 11 Reconstruction results of textured and curved surface object with grid pattern. (a) input, (b) PCA, (c) NCC

confirm that the proposed method could reconstruct the object shapes with curved surfaces and non-uniform texture using any tested projection pattern.

On the other hand, some projection pattern-specific tendencies were observed. In the case using random grids, PCA-based reconstruction was affected by the complicated curve and texture, as shown in Figure 11.



(a) (b) (c)

Fig. 12 Reconstruction results of textured and curved surface object with pattern proposed by [23]. (a) input, (b) PCA, (c) NCC

6. Conclusion

In this paper, we propose an active stereo technique based on light field approach. With our technique, light field is formed by multiple projectors that project arbitrary patterns and are located in arbitrary poses. The resulting light field forms patterns on 3D surfaces with rich depth queues of depth information. Since these queues are not easy to extract analytically, we use matching-based approach, which can be universally applied for arbitrary types of image queues. The set of the sample images for varying depths are generated by CG simulation. To realize efficient matching process for the 3D measurement, the dimensionality of the raw data of the image patches in the sample image set is reduced by PCA. Then, the comparison of the patches in the captured image with the sample image set is done in the low dimensional space. Experimental results are shown to prove that our technique is stable irrespective to material of target objects, lighting condition, noise of the sensor and projection patterns. In the future, real-time implementation using SIMD will be considered as an important task for us.

Acknowledgment

This work was supported in part by JSPS KAKENHI Grant No. 15H02758, 15H02779 and 16H02849, MIC SCOPE 171507010 and MSR CORE12.

References

[1] R. Benveniste and C. Ünsalan. A color invariant based binary coded structured light range scanner for shiny objects. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 798–801. IEEE, 2010.
[2] K. W. Bowyer, K. Chang, and P. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. *Computer vision and image understanding*, 101(1):1–15, 2006.
[3] V. Couture, N. Martin, and S. Roy. Unstructured light scanning to overcome interreflections. In *2011 International Conference on Computer Vision*, pages 1895–1902, Nov 2011.
[4] M. Hirsch, G. Wetzstein, and R. Raskar. A compressive light field projection system. *ACM TOG*, 33(4):58, 2014.
[5] J. Jurik, A. Jones, M. Bolas, and P. Debevec. Prototyping a light field display involving direct observation of a video projector array. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, pages 15–20. IEEE, 2011.
[6] H. Kawasaki, R. Furukawa, , R. Sagawa, and Y. Yagi. Dynamic scene shape reconstruction using a single structured light pattern. In *CVPR*, pages 1–8, June 23-28 2008.
[7] H. Kawasaki, H. Masuyama, R. Sagawa, and R. Furukawa. Single colour one-shot scan using modified penrose tiling pattern. *IET Computer Vision*, 7(5):293–301, October 2013.
[8] H. Kawasaki, S. Ono, Y. Horita, Y. Shiba, R. Furukawa, and S. Hiura. Active one-shot scan for wide depth range using a light field projector based on coded aperture. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3568–3576, 2015.
[9] Q. Li, M. Biswas, M. R. Pickering, and M. R. Frater. Accurate depth estimation using structured light and passive stereo disparity estimation. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 969–972. IEEE, 2011.
[10] Q. Li, M. Biswas, M. R. Pickering, and M. R. Frater. Dense depth estimation using adaptive structured light and cooperative algorithm. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, pages 21–28. IEEE, 2011.
[11] Microsoft. Xbox 360 Kinect, 2010. http://www.xbox.com/en-US/kinect.
[12] M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. 2:331–340, 2009.
[13] H. Murase and S. K. Nayar. Visual learning and recognition of 3-d objects from appearance. *International journal of computer vision*, 14(1):5–24, 1995.
[14] K. Nagano, A. Jones, J. Liu, J. Busch, X. Yu, M. Bolas, and P. Debevec. An autostereoscopic projector array optimized for 3d facial display. In *ACM SIGGRAPH 2013 Emerging Technologies*, page 3. ACM, 2013.
[15] R. Sagawa, K. Sakashita, N. Kasuya, H. Kawasaki, R. Furukawa, and Y. Yagi. Grid-based active stereo with single-colored wave pattern for dense one-shot 3D scan. In *Proc. 2012 Second Joint 3DIM/3DPVT Conference*, pages 363–370, Zurich, Switzerland, Oct. 2012.
[16] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern recognition*, 43(8):2666–2680, 2010.
[17] K. Sato and S. Inokuchi. Three-dimensional surface measurement by space encoding range imaging. *Journal of Robotic Systems*, 2:27–39, 1985.
[18] S. Tang, X. Zhang, and D. Tu. Fuzzy decoding in color-coded structured light. *Optical Engineering*, 53(10):104104–104104, 2014.
[19] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.
[20] M. Uenohara and T. Kanade. Use of fourier and karhunen-loeve decomposition for fast pattern matching with a large set of templates. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(8):891–898, 1997.
[21] A. O. Ulusoy, F. Calakli, and G. Taubin. Robust one-shot 3d scanning using loopy belief propagation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 15–22. IEEE, 2010.
[22] M. Visentini-Scarzanella, T. Hirukawa, H. Kawasaki, R. Furukawa, and S. Hiura. A two plane volumetric display for simultaneous independent images at multiple depths. In *PSIVT workshop Vision meets Graphics*, pages 1–8, 2015.
[23] P. Vuylsteke and A. Oosterlinck. Range image acquisition with a single binary-encoded light pattern. *IEEE Trans. on PAMI*, 12(2):148–164, 1990.
[24] L. Zhang and S. Nayar. Projection defocus analysis for scene capture and image display. *ACM Transactions on Graphics (TOG)*, 25(3):907–915, 2006.
[25] X. Zhang, Y. Li, and L. Zhu. Color code identification in coded structured light. *Applied optics*, 51(22):5340–5356, 2012.
[26] Furukawa, Ryo and Morinaga, Hiroki and Sanomura, Yoji and Tanaka, Shinji and Yoshida, Shigeto and Kawasaki, Hiroshi. Shape Acquisition and Registration for 3D Endoscope Based on Grid Pattern Projection. *European Conference on Computer Vision*, 399–415, 2016.