Estimation of 3D Geometry Using Multi-View and Structured Circular Light System

Deokwoo Lee
Samsung Electronics / Camera R&D Lab, Division of Mobile Communications, Suwon, South Korea
Email: dwoolee@gmail.com

Hamid Krim
North Carolina State University / Department of Electrical and Computer Engineering, Raleigh, NC, USA
Email: ahk@ncsu.edu

***Abstract*—This paper proposes an approach to multiple view-projection systems planning for a 3D object reconstruction in a practical settings. A *I/O* system theoretic is adopted and, viewpoints and light sources which are located at different positions are respectively the output and the input. The system identification is defined by a relationship between input and output. Structured circular light patterns are generated by multiple light sources. The proposed approach may be applied to both multiple inputs(projectors)-viewpoints (MPV) or single input-viewpoint (*SPV*) systems. This contribution is chiefly for an approximate reconstruction sufficient to characterize / clarify a target object rather than a perfect 3D reconstruction. To that end, we propose a development of an efficient 3D system identification frameworks based around an efficient 3D camera system we have developed in the laboratory. We show that these are closely related to the geometric information of an object, and with minimal prior information, an efficient *MPV* system is achieved.**

***Index Terms*—multiple view-projection, 3D reconstruction, Structured light system, 3D camera system**

I. INTRODUCTION

3D image reconstruction has been extensively researched in a numerous areas including, computer vision, surveillance, security, biometrics, etc. *Passive* and *active* methods have been widely used to recover 3D real information [1], [2], [3]. The active method which uses structured light systems has been employed due to its advantages such as efficient 2D data acquisition process, low cost measurement setup and efficient analytical modeling. The active method is based on the principle that deformed light patterns due to a surface shape, lead to the reconstruction problem solution without prior information of a target object. In practical perspectives, since we are interested in objects of any size, several projectors and cameras are required to solve the reconstruction problem. The MPV system is a very useful approach to 3D reconstruction work using structured light systems. Faugeras [4] and Hartley [5] are widely referenced in multiple view geometry area, in addition to many others, but the contributions are focused on solving the correspondences between multiple viewpoints. In MPV system, the following are required to be considered:

(1) Orientations / locations of viewpoints and light sources

(2) The number of viewpoints and light sources

(3) The number of structured light patterns (sampling density)

To solve (3), there has been extensive research on sampling for signal reconstruction ([6], [7], [8], [9]). This paper discusses only (1) and (2). If a target object is illuminated by as many projectors as needed, and is captured by as many cameras as needed, the reconstruction results will be extremely accurate. But this experimental environment is not efficient when considering practical perspectives. Assuming a work space is known *a priori*, the number of projectors is appropriately used so that the target object is illuminated. Cameras capture deformed light patterns on the object, and with prior knowledge of the original light patterns, the reconstruction is performed based on establishing a relationship between original and deformed light patterns. Locations and orientations of projectors should be taken into account to avoid false projection, and are closely related to the specific geometric information of a surface shape. This problem has been also addressed in [10], and they have shown that normal vectors of a surface are closely related to the determination of sensor planning. To achieve a successful illumination, this paper shows that curvatures can determine the locations and the orientations of projectors. When the object is illuminated by multiple light sources, overlapped light patterns are captured, and the reconstruction process based on the captured light patterns may be repeated in these areas. We assume that the object is only illuminated by projected light sources, and do not consider ambient light and light divergence in this paper. In this environment, overlapped light patterns have different intensity values from the areas illuminated by a single light source (Fig. 3). Detecting the overlapped areas avoids redundancy and repeated reconstruction process. The location, the orientation and the number of cameras are related to the distance between neighboring cameras. The minimum

Manuscript received January 4, 2014; revised May 7, 2014.

doi: 10.12720/joig.2.1.77-82

number of cameras to capture the target object is required to increase system efficiency. This paper deals with a theoretical proof of projection and viewpoint planning, and the projection density. The rest of this paper is organized as follows. In Section 3, the proposed algorithms are briefly described before we detail them. Section 4 describes a relationship between the locations / orientations of projectors and the geometric information of an object. In Section 5, we derive a multiple view reconstruction system, using an input and an output, with redundant information elimination (Section 6). Once the reconstruction system is completed, the optimal number of viewpoints is derived in Section 7. To substantiate the proposed algorithms, experimental results using real objects are depicted in Section 8 prior to our conclusion.

II. Algorithm Description

The proposed algorithm in this paper proceeds in these steps.

(1) Project circular patterns with multiple projectors.

(2) Extract 3D coordinates from each viewpoint (i.e. camera, 2D image plane, etc.)

(3) Synthesize the reconstruction results into a single 3D model using coordinate transformation.

In the first step, orientations and locations of projectors are proposed. In the second step, the geometric reconstruction, which has been published in [11], is performed. In addition, the system efficiency with respect to the optimal number of viewpoints, is presented. In the third step, when merging a set of incomplete reconstruction results in a single model, the areas containing overlapped circular patterns are carefully analyzed to eliminate redundant information.

III. Input (Projection) Constraints

Projection constraints for *MPV* system consist of locations and orientations of structured light sources (e.g. projectors). Using some prior information, such as approximated range and geometry of a target object, projection constraints is determined, and this will avoid false projection. This section details the projection constraints of locations and orientations of structured light sources. To completely acquire deformed light patterns in a 2D image plane, all of the patterns from the light sources should be projected onto the object. Otherwise, such missing parts will not be acquired in a 2D image plane, and the 3D measurement results will not be reliable. To achieve a successful projection avoiding missing parts, two simple solutions may be considered. The first one which is very trivial, is simply adding more projectors, and is not discussed here. The second is efficiently planning the position of projectors. Let $S_3 \subset \mathbb{R}^3$ and $P \subset S_3$ be a domain of a 3D object and a 3D point on it, respectively. At any point P, two tangent vectors and curvatures are defined.

$$\overrightarrow{T_P} = \left(\overrightarrow{T_{p1}}, \overrightarrow{T_{p2}}\right),\ \kappa_p = \left\| \frac{d\overrightarrow{T_P}}{ds} \right\| \tag{1}$$

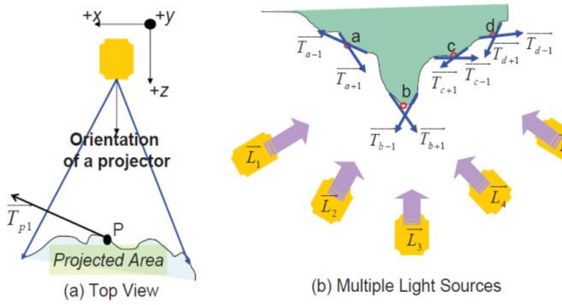

Figure 1. (a). Projected area. (b). False projection. When the orientation of a structured light pattern is same as the tangent direction of an object, the pattern is not projected onto the area of an object. This case is referred to as a False projection.

Once the location setup is completed, the orientation of a projector affects the quality of 3D measurement. Denoting $\vec{L}$ by an orientation of a projector, it should satisfy the following (also see Fig. 1):

$$\vec{L} \neq k\overrightarrow{T_{p1}},\ \vec{L} \neq k\overrightarrow{T_{p2}},\ \left\langle \overrightarrow{N_p}, \vec{L} \right\rangle < 0 \tag{2}$$

where $k \in \mathbb{R}$ and N_p are an arbitrary constant number and a normal vector at the point $P \in \mathbb{R}^3$, respectively.

Each camera has its own characteristics called intrinsic and extrinsic parameters being considered crucial in camera calibration applications ([12] and [13]). Lens distortions, one of the important factors as well, are assumed to be negligible or appropriately corrected. To achieve a successful multiple projection system, every tangent direction is to be measured at every point on the object, but this is not efficient in practice. From geometrical perspective, the curvature represents a characteristic of an object quantitatively. High curvature areas of a surface (e.g. face) may contain important characteristics, and could provide sufficient information. The optimal number of projectors is directly dependent on to the number of areas which have high curvatures. To achieve an efficient projection system, either single or multiple, only the necessary number of curvatures or corresponding tangent vectors is acquired. The projection rule for any projector is to avoid the direction along the tangents of dominant points, so as to not lose geometric information about the object.

IV. Multi-View Reconstruction System

Denoting by x an observed scene overlaid with circular light patterns, each camera may captures the area overlaid with overlapped light patterns. In case of two-projectors-viewpoints, Fig. 3 presents a simple example. The reconstruction process consists of the following:

$$\begin{aligned} &\mathrm{y} = \{y_1, y_2, \ldots\ldots, y_j, \ldots\ldots, y_D\} \\ &\mathrm{x} = \{x_1, x_2, \ldots\ldots, x_j, \ldots\ldots, x_D\} \\ &y_j = M^j x_j,\ y_j \in \mathbb{R}^3,\ x_j \in \mathbb{R}^3 \end{aligned} \tag{3}$$

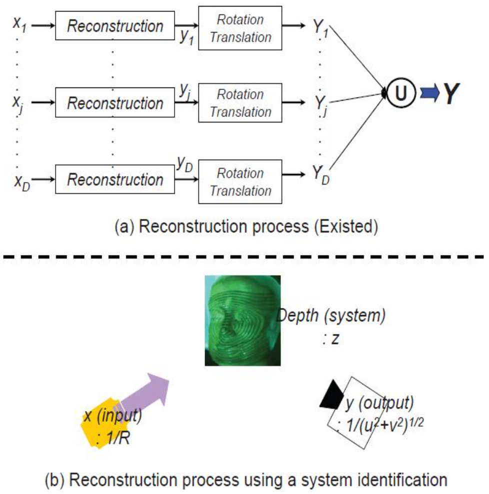

Figure 2. Reconstruction using MPV is combining all the reconstruction results followed by coordination and redundancy elimination.

where $M^j = M^j_{in} M^j_{ex}$ represents the relationship of observed points in 2D and its projected real world coordinates in 3D, and this relationship has been derived in [11] (Fig. 2-(a)). In addition, it is composed of intrinsic M^j_{in} and extrinsic parameters M^j_{ex} [4]. We now efficiently restate the reconstruction problem as a system identification. If an input and an output are respectively projected circular patterns and deformed ones, the 3D reconstruction result is our system (Fig. 2-(b)). For instance, different target objects are corresponding to different systems. Assuming parallel light projection and the constraints of the patterns, the input $(x, y, 0)$, the output (u, v, f) and the system (x, y, f) can be represented as the follows :

$$\begin{aligned} &\text{Input (original circles)} : x^2 + y^2 = R^2 \\ &\text{Output (deformed circles)} : u^2 + v^2 \\ &\text{Depth} : z = \frac{fR}{\sqrt{u^2+v^2}}, \end{aligned} \tag{4}$$

where f is a the *focal length*, and the depth is composed of R and $\sqrt{u^2+v^2}$. Let the depth z be a system, it is also written by *Output* / *Input*, where the *Input* is $1/R$ and the *Output* is $1/\sqrt{u^2+v^2}$. Thus, the reconstruction problem is restated as a *system identification* problem. Different systems are equivalent to different target objects or depths (Section. 8). To this end, the union set of every reconstruction, $\bigcup_{j=1}^{D} Y_j$ achieves a complete reconstruction. To complete a multiple-view reconstruction, we categorize our approach into four steps which will be discussed in the following sections: 1) eliminating redundant information, induced from the areas overlaid with overlapped projected light patterns, 2) a reconstruction from each viewpoint whose location satisfies the minimum number of viewpoints (which contributes the third step), and 4) merging all reconstruction results into a single image representation.

V. Eliminating Redundant Information

Based on the projection constraints, an object is perfectly illuminated without any occlusion. *MPV* system generates redundant information-overlapped light patterns on the object-which can decrease a reconstruction efficiency. Denoting y_1 and y_2 by reconstructed 3D coordinates from x_1 (left view) and x_2 (right view), respectively (Fig. 3),

$$y_1 = M^1 x_1,\ y_2 = M^2 x_2 \tag{5}$$

Once the coordinate system of y_1 and y_2 are aligned, the complete reconstruction is hence represented as

$$Y_1 = y_1,\ Y_2 = \Im y_1 \tag{6}$$

$$Y = Y_1 \cup Y_2 = Y_1 + Y_2 - (Y_1 \cap Y_2) \tag{7}$$

where Y_1 and Y_2 are reconstructed results from the right and left cameras, respectively, $\Im$ represents a coordinate transformation, and $Y_1 \cap Y_2$ represents the redundant information. The approach to detection of overlapped light patterns on the surface is measuring the intensity variation of the object (Fig. 3).

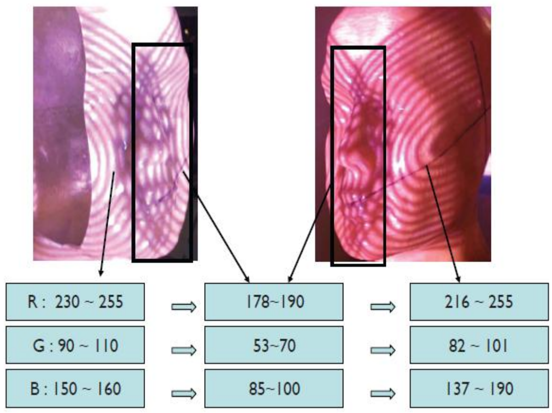

Figure 3. Under the assumption that the object is illuminated only by the projected light patterns the areas overlaid with overlapped light patterns of different intensities from other areas. From the intensity variations, overlapped patterns are detected.

Let's consider a set of D viewpoints. Denoting by I_j and I_{j+1} a set of intensity values of the objects overlaid with the circular patterns contributed by the j^{th} and the $j+1^{th}$ light sources, respectively, we can write

$$I_j \in x_j,\ I_{j+1} \in x_{j+1} \text{ or } I_j \in y_j,\ I_{j+1} \in y_{j+1}$$

If the observed intensities have common values, $I_j \cap I_{j+1} \neq \varnothing$, there exists redundancy and corresponding geometric information is discarded from either of I_j or I_{j+1}. A coordinate transformation is then carried out for a complete reconstruction.

VI. OUTPUT (VIEWPOINT) CONSTRAINTS

The dominant computational cost affecting the reconstruction system efficiency may be categorized as follows:

(1) Acquisition of the current view (2D) : t_1
(2) Reconstruction results from 2D to 3D : t_2
(3) Fusion and image registration : t_3
(4) Performing projectors planning (i.e. locating and determining light sources and density of light patterns, respectively) : t_4
5) Performing viewpoints panning (i.e. determining the locations and number of cameras) : t_5

Let $M_j N_j$ be the size of data of the j^{th} viewpoint, then total complexity can be represented as

$$C = \left(\sum_{j=1}^{D} M_j N_j \right) (t_1 + t_2) + t_3 + t_4 + t_5 \qquad (8)$$

To improve our system efficiency (i.e. to minimize C), we reduce D, the number of viewpoints. This approach is equivalent to maximizing the distance between neighboring viewpoints.

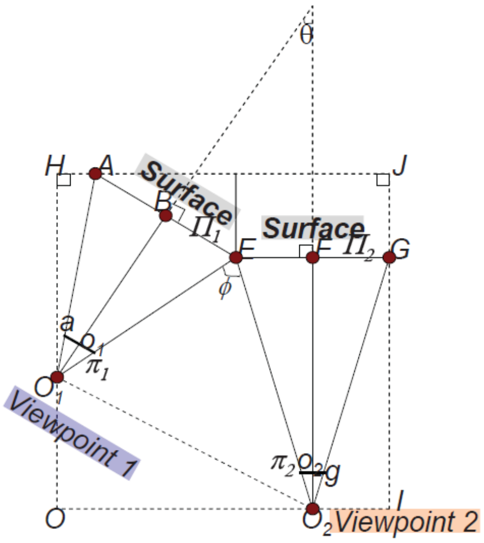

Figure 4. Viewpoint constraint is a position of cameras to completely cover whole object. To achieve an efficient position planning (i.e. minimize the number of viewpoints), maximum distance between neighboring cameras are estimated.

In Fig. 4, minimizing C is equivalent to maximizing $\overline{OO_2}$, $\overline{O_1O_2}$, or maximizing a distance between neighboring cameras. We assume that the cameras are calibrated to pinhole models, and lens distortions are negligible or already corrected. In addition, we assume that we have prior information such as intrinsic and extrinsic parameters, approximate range of a target surface ($\overline{AE}$ and $\overline{EG}$ in Fig. 4), and the distance between the optical center (o_1, o_2) and arbitrary reference planes (B , F). Denoting $\overline{O_1O_2}$, $\overline{OO_2}$ and $\overline{OO_1}$ by d , d' and d'' respectively, we can write

$$d = \sqrt{D_1^2 + D_2^2 - 2D_1D_2\cos\phi} \qquad (9)$$

$$d'^2 = d^2 - d''^2 \qquad (10)$$

where C_1, C_2, D_1 and D_2 are $\overline{BE}$, $\overline{EF}$, $\overline{O_1E}$ and $\overline{O_2E}$, respectively. If $d'^2 \geq d^2 - d''^2$, then the viewpoints are so far from each other that the object cannot be fully covered by them. With no prior information of approximate ranges, C_1 and C_2 or d_1 and d_2 , the viewpoint problem is still solvable by using the *Least-Square* problem formulation.

VII. EXPERIMENTAL SYSTEM

The system setup includes 2 projectors, 2 cameras and an object. Each projector connected to a laptop computer generates circular patterns, and is located approximately 1 meter far from the object. Regular cameras are used to capture the object overlaid with the projected circular patterns. These are also approximately located 1 meter away from the object. We used the following projectors and cameras; *1024* × 768 resolution *COMPAQ MP1600*, *1024* × 768 resolution *LCD EPSON POWERLITE 76C*, and *1280* × *720* pixel resolution Canon camera. The camera calibration is performed using a checkerboard to estimate internal characteristics. Fig. 5 shows an experiment using 2 projectors each of which projects 7 circular patterns. Two cameras each of which has a different perspective in capturing the projected light patterns. The estimated focal length here is about *8000* pixels (actually in a range of *7844* ~ *8452* pixels) and an image center is $(u_0, v_0) = (174, 68)$. Our experiment used *3* circular patterns whose radii are *1*, *2*, and *4* inch respectively, and the object is illuminated, with the deformed patterns, is viewed from 2 different positions (Fig. 6). Once 3D real coordinates (x, y, z) are obtained from each viewpoint (Fig. 7), redundant information (overlapped patterns) is eliminated (Fig. 8).

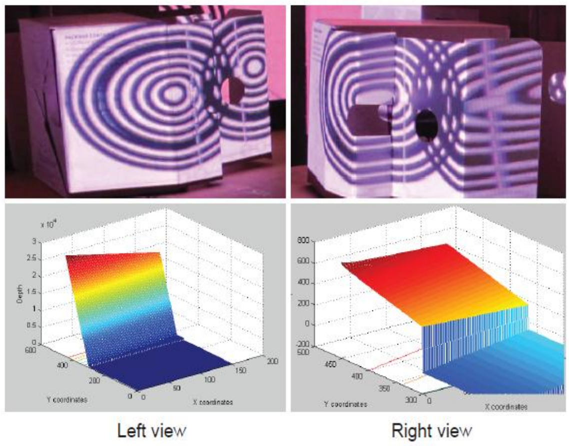

Figure 5. A box model illuminated by 2 projectors each of which generates 7 circular patterns, and 2 cameras capture the patterned image from the left and right perspective of view. The figures in the second row are reconstructed resulted from each viewpoint.

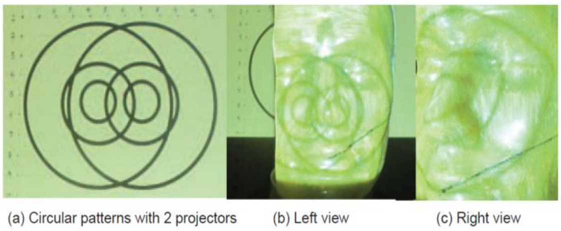

Figure 6. 3 Circular patterns are used in our experiment. Radii of these are 1, 2 and 4 inches respectively.

By rotating and translating the reconstruction result based on the coordinate system of the other one, reconstruction is completed (Fig. 9). Another experimental result is shown in Fig. 10.

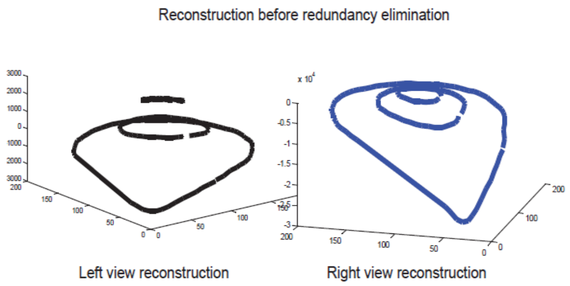

Figure 7. Reconstruction is performed from the different viewpoints.

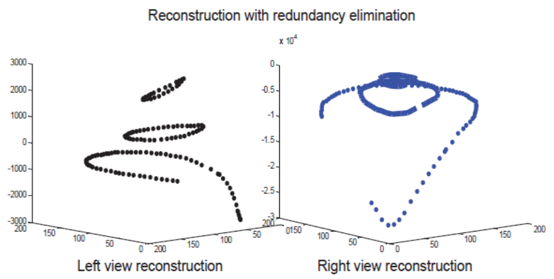

Figure 8. Redundant information is eliminated based on the intensity variations.

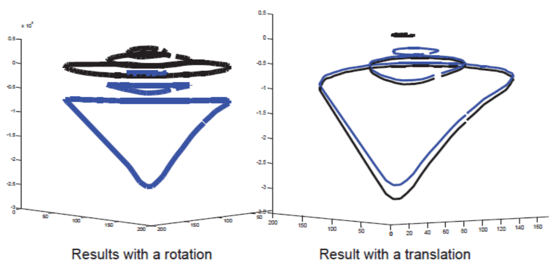

Figure 9. Coordinate transformation completes the reconstruction.

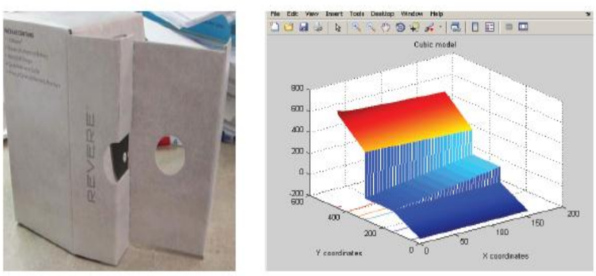

Figure 10. Reconstruction of a box model in Fig. 5.

VIII. Conclusion

This paper proposes an approach to determining the parameters of an MPV system which can be used for robotics, automatic vision system and others. This paper shows that the optimal number of projectors is determined by geometrically dominant areas of an object. Once the projection constraints are determined, overlapped light patterns are considered and mitigated. After a coordinate transformation, reconstruction results from all of the viewpoints are merged into a single image representation. Since the experiments are conducted using real objects, the ground truth models are not provided, the reconstruction results cannot be compared to the original object. The future research includes sampling density determination given the density of previous light sources. Other illuminations such as ambient light and diverged light patterns should also be considered in practice and would be noise sources [14], [15]. Another technical issue is a viewpoint system including errors from internal characteristics of camera. For example, lens distortion can also affect reconstruction results when shape classification between similar objects is required.

References

[1] U. R. Dhond and J. K. Aggarwal, "Structure from Stereo-A Review," *IEEE Transaction on Systems, Man, and Cybernetics*, vol. 19, 1989.

[2] J. Batlle, E. Mouaddib, and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: A survey," *Pattern Recognition*, vol. 31, no. 7, pp. 963–982, 1998.

[3] J. Geng, "Structure from stereo-A review," *Journal of Advances in Optics and Photonics*, vol. 3, pp. 128–160, 2011.

[4] O. Faugeras and Q. T. Luong, *The Geometry of Multiple Images*, 2001.

[5] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd Ed. Cambridge University Press, 2004.

[6] D. Lee and H. Krim, "A sampling theorem for a 2D surface," *SSVM*, pp. 556–567, 2011.

[7] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review, in *Proc. the IEEE*, vol. 65, 1977.

[8] J. R. Higgins, "A sampling theorem for irregularly spaced sample points (Corresp)," *IEEE Transactions on Information Theory*, vol. 22, 1976.

[9] Y. Tsaig and D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, 2006.

[10] S. Y. Chen, T. F. Li, J. W. Zhang, and W. L. Wang, *Active Sensor Planning for Multiview Vision Tasks*, Springer, 2008.

[11] D. Lee and H. Krim, "3D surface reconstruction using structured circular light patterns," *ACIVS 2010*, vol. 2, pp. 279–289.

[12] G. J. Zhang, Z. Liu, J. H. Sun, and Z. Z. Wei, "Novel calibration method for a multi-sensor visual measurement system based on structured light," *Optical Engineering*, vol. 49, 2010.

[13] R. S. Lu and Y. F. Li, "A global calibration method for large-scale multi sensor visual measurement systems," *Sensor and Actuators*, vol. 116, pp. 384–393, 2004.

[14] W. Zhou, C. Kambhamettu, and R. Kambhamettu, "Estimation of illuminant direction and intensity of multiple light sources," Department of Computer Science, Drexel University, Prior, pp. 206–220, 2002.

[15] L. T. Maloney and B. A. Wandell, "Color constancy: A method for recovering surface spectral reflectance," *JOSA A*, vol. 3, no. 1, pp. 29-33, 1986.

Deokwoo Lee received the degrees in Electrical Engineering (B.S from Kyungpook National University, South Korea, M.S and Ph.D from North Carolina State University, 2007, 2008, 2012, respectively). He was a Postdoctoral Research Associate in School of Medicine at Washington University in St. Louis, MO, USA. He subsequently joined Camera R&D Lab, Division of Mobile Communications, Samsung electronics, Suwon, South Korea in 2013 as a Senior Research Engineer. His research interests include image processing, computer vision, pattern recognition and others related to applications for mobile camera, automatic vision systems, medical imaging, etc.

Hamid Krim received the degrees in Electrical Engineering (B.S and M.S from University of Washington, WA, and Ph.D from Northeastern University, MA). He was a Technical Staff Member with AT&T Bell Laboratories, where he was with the area of telephony and digital communication systems and subsystems. In 1991, he became an NSF Postdoctoral Scholar with the Foreign Centers of Excellence, LSS Supelec University of Orsay, Paris, France. He subsequently joined the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, as a Research Scientist performing and supervising research, and later as a faculty member with the Department of Electrical and Computer Engineering at North Carolina State University, Raleigh, NC, USA in 1998. He is an original contributor and currently an affiliate of the center for Imaging Science sponsored by the Army. His current research interests include statistical signal processing and mathematical modeling with a keen emphasis on applications. He was a recipient of the NSF Career Young Investigator Award. He was on the editorial board of the IEEE Transactions on Signal Processing and regularly contributes to the society in a variety of ways. He was a member of the Society for Industrial and Applied Mathematics and Sigma Xi.