INVITED PAPER

Multidimensional Optical Sensing and Imaging System (MOSIS): From Macroscales to Microscales

This review paper describes a passive multidimensional optical sensing and imaging system (MOSIS), which can be used for 3-D visualization, seeing through obscurations, material inspection, 3-D endoscopy, and 3-D object recognition from microscales to long-range imaging. The system utilizes time and space multiplexing, polarimetric, temporal, photon flux, and multispectral information to reconstruct multidimensionally integrated scenes.

By Bahram Javidi, *Fellow IEEE*, Xin Shen, Adam S. Markman, Pedro Latorre-Carmona, Adolfo Martínez-Uso, José Martinez Sotoca, Filiberto Pla, Manuel Martínez-Corral, Genaro Saavedra, Yi-Pai Huang, and Adrian Stern, *Member IEEE*

ABSTRACT | Multidimensional optical imaging systems for information processing and visualization technologies have numerous applications in fields such as manufacturing, medical sciences, entertainment, robotics, surveillance, and defense. Among different 3-D imaging methods, integral imaging is a promising multiperspective sensing and display technique. Compared with other 3-D imaging techniques, integral imaging can capture a scene using an incoherent light source and generate real 3-D images for observation without any special viewing devices. This review paper describes passive multidimensional imaging systems combined with different integral imaging configurations. One example is the integral-imaging-based multidimensional optical sensing and imaging system (MOSIS), which can be used for 3-D visualization, seeing through obscurations, material inspection, and object recognition from microscales to long-range imaging. This system utilizes many degrees of freedom such as time and space multiplexing, depth information, polarimetric, temporal, photon flux and multispectral information based on integral imaging to record and reconstruct the multidimensionally integrated scene. Image fusion may be used to integrate the multidimensional images obtained by polarimetric sensors, multispectral cameras, and various multiplexing techniques. The multidimensional images contain substantially more information compared with 2-D images or conventional 3-D images. In addition, we present recent progress and applications of 3-D integral imaging including human gesture recognition in the time domain, depth estimation, mid-wave-infrared photon counting, 3-D polarimetric imaging for object shape and material identification, dynamic integral imaging implemented with liquid-crystal devices, and 3-D endoscopy for healthcare applications.

Manuscript received August 19, 2016; revised October 18, 2016; accepted December 16, 2016. Date of publication March 2, 2017; date of current version April 20, 2017. The work of B. Javidi was supported in part by the National Science Foundation (NSF) under Grant NSF/IIS-1422179, in part by the Defence Advanced Research Projects Agency (DARPA), and in part by the U.S. Army under Contract W911NF-13-1-0485. The work of P. Latorre Carmona, A. Martínez-Uso, J. M. Sotoca, and F. Pla was supported by the Spanish Ministry of Economy under Project ESP2013-48458-C4-3-P, and by MICINN under Project MTM2013-48371-C2-2-PDGI, by Generalitat Valenciana under Project PROMETEO-II/2014/062, and by Universitat Jaume I through Project P11B2014-09. The work of M. Martínez-Corral and G. Saavedra was supported by the Spanish Ministry of Economy and Competitiveness under Grant DPI2015-66458-C2-1R, and by the Generalitat Valenciana, Spain, under Project PROMETEOII/2014/072.

B. Javidi, X. Shen, and **A. S. Markman** are with the Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT 06269 USA (e-mail: bahram.javidi@uconn.edu; xin.shen@uconn.edu; adam.markman@uconn.edu).
P. Latorre-Carmona, **A. Martínez-Uso, J. M. Sotoca,** and **F. Pla** are with the Institute of New Imaging Technologies, Universidad Jaume I, 12071 Castellon de la Plana, Spain (e-mail: latorre@uji.es; auso@uji.es; sotoca@uji.es; pla@uji.es).
M. Martínez-Corral and **G. Saavedra** are with the Department of Optics, University of Valencia, 46100 Burjassot, Spain (e-mail: manuel.martinez@uv.es; genaro.saavedra@uv.es).
Y.-P. Huang is with the Department of Photonics and the Institute of Electro-Optical Engineering/Display Institute, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: boundshuang@mail.nctu.edu.tw).
A. Stern is with the Electro-Optical Engineering Unit, Ben-Gurion University of the Negev, Beer Sheva 81405, Israel (e-mail: stern@bgu.ac.il).

Digital Object Identifier: 10.1109/JPROC.2017.2654318

KEYWORDS | 3-D endoscopy; 3-D human activity recognition; 3-D imaging; dynamic integral imaging; long-range integral imaging; material analysis; multidimensional object recognition; multispectral imaging; photon counting; polarimetric imaging

I. INTRODUCTION

There have been significant technological advancements in sensors, devices, materials, algorithms, and computational hardware. Therefore, sensing and visualization capabilities applied to real world objects have improved extensively. In recent decades, 3-D imaging technology has received interest from many research groups. Instead of conventional 2-D sensing techniques, which record the intensity of the scene, passive 3-D imaging also includes depth and directional information. Many techniques for 3-D imaging have been proposed, such as holography and interferometry [1], [2], two-view based stereoscopy [3], [4], and multi-view techniques for autostereoscopic 3-D imaging [5], [6], to cite a few.

Integral imaging [7] is an autostereoscopic 3-D sensing and imaging technique, which provides true 3-D images with full parallax and quasi-continuous viewing angles [8]. In addition, integral imaging can work well for long-range objects [9]. In contrast, some other 3-D sensing techniques, such as the time-of-flight camera [10] or structured light techniques [11], [12], may not work well for long-range objects. Integral imaging is a promising technique that has been used in various fields, such as 3-D sensing [13], 3-D displays [14]–[17], holographic display [18], 3-D imaging of objects in turbid water [19], 3-D tracking [20] and 3-D target detection and recognition [21], [22], photon counting 3-D sensing and visualization [23]–[25], 3-D microscopy [26]–[31] and endoscopy for microscale 3-D imaging and display [32], [33], head tracking 3-D display [34], 3-D augmented reality [35]–[38], to cite a few.

Originally developed for space-based imaging [39], multispectral imaging captures the information corresponding to specific wavelengths of light. The spectrum for an imaging system can be extended from the visible range to the near-infrared (NIR) range, mid-wave-infrared (MIR) range, or long-wave-infrared (LWIR) range. Applications of multispectral imaging range from remote sensing [40]–[42] to medical imaging [43], to name a few.

One of the fundamental properties of light is its state of polarization [44], [45]. From this information, we may obtain optical and physical properties of materials using noninvasive optical probes [46], [47]. This information can be helpful for material inspection and classification in manufacturing, remote sensing and security applications [48]–[50]. The polarization state of light allows the sensor to capture information about an object's surface material, such as birefringence, photoelastic effect, etc. When this information is combined with other sensor data, the overall effectiveness of a multidimensional imaging system, such as the integral-imaging-based multidimensional optical sensing and imaging system (MOSIS) [51], is enhanced. In MOSIS, polarimetric characteristics from a real-world scene are extracted from a polarimetric imaging system.

Integrating features from multidimensional and multimodal imaging, that is, 3-D imaging, multispectral imaging and polarization imaging, etc., provides unique information about a scene. In this paper, we present an overview of some recent work on multidimensional sensing and integrated visualization with 3-D integral imaging technology. In addition, new work on using MOSIS 2.0 for 3-D object shape, material inspection, and recognition, such that similar objects with different materials can be discriminated, is presented. To the best of our knowledge, this is the first time that all of these degrees of freedom are integrated in a passive 3-D integral imaging system. This paper is organized as follows: the original concept of MOSIS [51] is first reviewed in Section II, followed by the principle and recent progress of the integral imaging technique in Section III. Section IV presents the development of the 3-D polarimetric integral imaging sensing and visualization. Integral imaging techniques in the infrared domain are presented in Section V. Three-dimensional human gesture recognition using integral imaging videos is discussed in Section VI, and recent progress of MOSIS 2.0 is given in Section VII. Section VIII presents a brief overview of MOSIS in microscales for medical applications and dynamic integral imaging systems with time multiplexing implemented with liquid crystal devices. Conclusions are given in Section IX. Progress in these topics has grown substantially in the recent years, and therefore it is difficult to give a complete overview of all the reported work. Thus, we apologize if some relevant work has been omitted in this review.

II. MULTIDIMENSIONAL OPTICAL SENSING AND IMAGING SYSTEMS (MOSIS)

In this section, the integral-imaging-based MOSIS is reviewed. MOSIS is an extension of the conventional 3-D integral imaging technique to incorporate multimodality into image sensing and reconstruction. Additional information obtained by the system can be further used for object recognition, material inspection and integrated 3-D visualization, etc., which can significantly enhance the amount of information extracted from a scene.

The concept of MOSIS [51] is to use different degrees of freedom from photons of a scene, such as polarization, angular information, spectral, time variation, etc., to reveal new information of the scene. It is a more advanced imaging sensor and visualization system compared with a conventional integral imaging system. Although some experiments can be done with different imaging setups, MOSIS can increase the amount of information extracted from the scene due to the multimodal and multidimensional measurements. As shown in Fig. 1(a), MOSIS can record a scene with separate sensors corresponding to various optical properties. In one modality, by moving a lenslet array with the moving array lenslet technique (MALT) [52] within a period of the lenslet, time multiplexed integral imaging pickup with an increased sampling rate can be obtained to improve 3-D visualization.

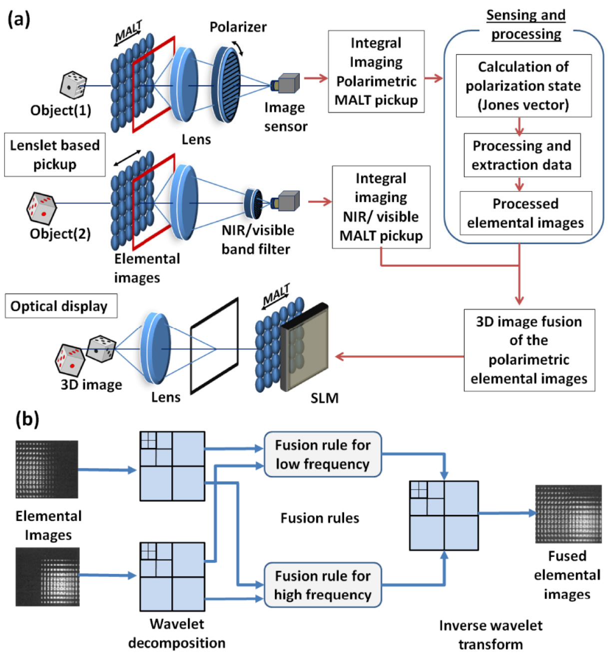

Fig. 1. ***(a) Overview of MOSIS. The proposed system fuses the polarimetric information, multispectral sensing, and multidimensional visualization with integral imaging. MOSIS may use the moving array lenslet technique (MALT) to improve resolution. (b) Elemental images fusion with wavelet decomposition in MOSIS [51].***

In addition, the image sensor can capture multispectral imaging and polarized information with specific filters and optical components. For polarimetric 3-D sensing, an object is first illuminated using linearly polarized light. The light reflected from the object's surface passes through the lenslet array, an imaging lens, and a rotating linear analyzer. The light then forms an array of elemental images, which is recorded by an image sensor. The reflected light's polarimetric information is determined through the Jones vector by using the rotation linear polarizer analyzer [46]. To optically visualize the polarimetric object, the polarization-selected elemental images are displayed in a spatial light modulator (SLM) with two quarter-wave plates. The multi-wavelength information in the visible and infrared range, including NIR, can be captured using a specific light source and a series of band-pass filters added in front of the image sensors.

The multidimensional data needs to be integrated for visualization. MOSIS may use wavelet decomposition to fuse the elemental images. The elemental images are decomposed into various channels guided by their local frequency content [53]–[55]. Fig. 1(b) depicts an example of the image fusion process with a three-level wavelet decomposition. A 2-D wavelet decomposition measures intensity fluctuations in the elemental images, along the horizontal, vertical and diagonal directions using wavelet filters. This is achieved by applying a low-pass or a high-pass filter from a wavelet family to an image's rows. After filtering, the image columns are down sampled by a factor of 2 such that only the even indexed columns are kept. A low-pass or a high-pass filter, from the wavelet family previously used, is then applied to the columns of the previously filtered images. This is then followed by down sampling the rows by a factor of 2 such that only the even indexed rows are kept. Three of the resulting images are the i^{th} level decomposition corresponding to the image's frequency information in the horizontal, vertical, or diagonal direction. The fourth image can then be inputted into the wavelet decomposition process producing another set of $(i+1)^{th}$ decomposition. For image fusion, the j^{th} decomposition level of the image and another image can be combined using image fusion rules, such as a weighted sum of the two levels. After fusion, an inverse wavelet transform is applied to obtain the fused elemental images.

III. PRINCIPLE AND RECENT PROGRESS OF INTEGRAL IMAGING

The original concept of integral imaging was proposed by Lippmann in 1908 [7] and called integral photography. The principle of this technique is to record a 3-D scene from multiple perspectives by using a lenslet array and a 2-D recording medium, such as film [56]–[58], since optoelectronic image sensors were not available at the time. Thanks to the rapid technological improvement in optoelectronic sensors, materials and devices, such as CCD and CMOS cameras, LC display screens, and the commercialization of computers, integral imaging has been revived in the recent decades [59]–[63]. There are two procedures in a typical integral imaging system for 3-D information acquisition and visualization, known as the pickup and reconstruction stages, respectively.

A. Pickup Stage of Integral Imaging

1) Lenslet Based Pickup Stage: Fig. 2(a) shows the characteristics of the integral imaging pickup stage. A lenslet array is

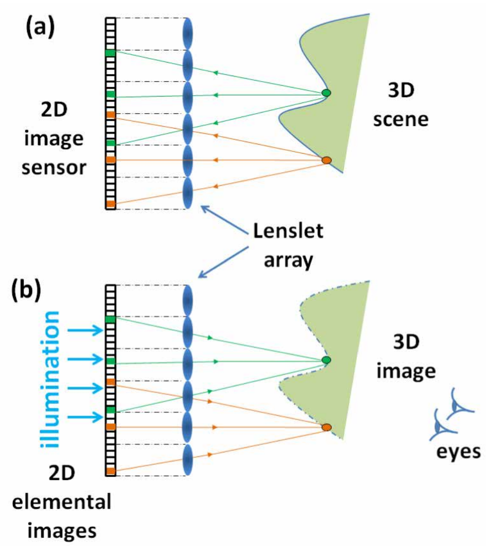

Fig. 2. ***Principle of integral imaging. (a) Pickup stage, and (b) reconstruction stage. Each object point in the pickup stage goes to a different pixel position in the 2-D sensor. During the 3-D reconstruction, those contributions make it possible for 3-D visualization of the object.***

placed in front of a 2-D image sensor. Light scattered by the 3-D scene surface passes through each lenslet and is then recorded by the sensor. Compared to the single lens imaging system, integral imaging obtains multiple 2-D images (named as elemental images) of the 3-D scene corresponding to each lenslet with different perspectives. Moreover, the image sensor, known as an elemental image array, captures both intensity and directional information of the light rays emitted by the scene.

The resolution of the captured elemental images may be limited by the configuration of the lenslet array and the pixel size of the sensor. The moving array lenslet technique was proposed in [52] to improve the resolution of the elemental images. There are many computational super resolution methods, but the moving array lenslet technique naturally increases the number of samples of the optical field which is available to improve the spatial sampling. By moving the lenslet array in the integral imaging pickup stage, the upper resolution limitation given by the Nyquist sampling theorem can be overcome. The parallax barriers (the dashed lines in Fig. 2) are needed on the image forming side of the lenslet array. Each of the captured elemental images corresponds to a specific lenslet and should only record the light information passing through it. If an elemental image records the light from the adjacent lenslet, the crosstalk phenomenon will happen on the elemental image and the 3-D display quality may be substantially degraded [64].

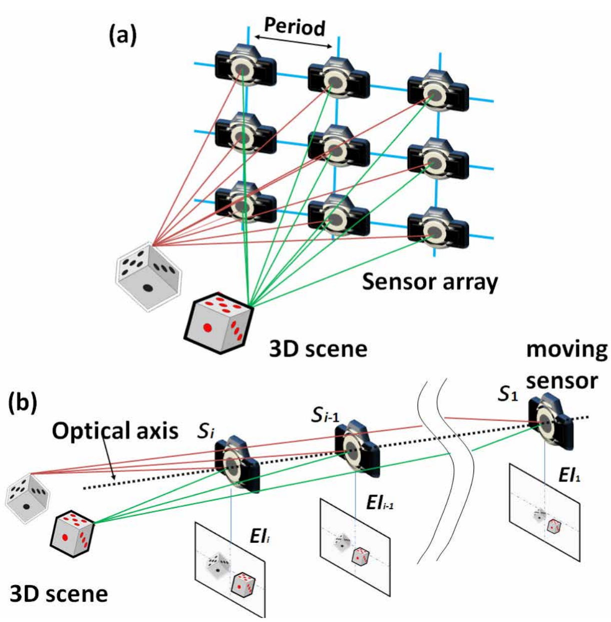

Fig. 3. ***(a) Example of synthetic aperture integral imaging (SAII) with a camera array in the pickup stage of integral imaging. A single camera on a moving platform may implement synthetic aperture integral imaging. (b) Example of a different passive 3-D imaging known as the axially distributed sensing (ADS) method, with a camera moving along its optical axis. S_i are the index of the camera positions and EI_i are the corresponding captured elemental images.***

2) Synthetic Aperture Integral Imaging Pickup Stage: Elemental images with high resolution, large field of view and extended depth-of-field can be achieved by using the synthetic aperture integral imaging technique [65] with the configuration of an array of imaging sensors or a moving image sensor array (an image sensor with a lens translated on a 2-D plane). A CCD or CMOS sensor records the scene with high resolution images. Furthermore, since the image sensor lens parameters (e.g., focal length and aperture, etc.) are controllable, synthetic aperture integral imaging provides flexibility for specific 3-D sensing requirements, which makes it more practical than the lenslet based integral imaging pickup technique. Synthetic aperture integral imaging may be implemented using a single camera on a moving platform or a camera array. Fig. 3(a) shows an example of a synthetic aperture integral imaging pickup stage by using a camera array. The period between adjacent image sensors, the number of sensors on the horizontal and vertical directions, and the sensor parameters can be adjusted in contrast to the conventional lenslet array. Synthetic aperture integral imaging allows the integral imaging pickup stage to increase the parallax of the captured images.

3) Axially Distributed Sensing and Flexible Sensing: Recently, 3-D sensing techniques based on synthetic aperture integral imaging were modified for the case that the image sensor may not be distributed in a planar and regular grid configuration. A multi-perspective 3-D imaging architecture named the axially distributed sensing (ADS) method is presented in [66]. For the 3-D sensing process, various perspectives of the scene are acquired by either moving the sensor along a common optical axis, or the object of interest is translated parallel to the optical axis. This method can be used for 3-D information computational extraction and reconstruction, since its acquisition capability is not uniform over the field of view. To simplify the configuration, elemental images based on the axially distributed sensing method are obtained by translating a single camera longitudinally along its optical axis as shown in Fig. 3(b).

In [67], a new integral imaging methodology for randomly distributed sensors was proposed assuming no rotation amongst the sensors; however, they may be at different x-, y- and z-coordinates relative to a reference camera position. Similar arrangements can be implemented with ADS.

B. Reconstruction Stage of Integral Imaging

1) Lenslet-Based Optical Display: Fig. 2(b) depicts the concept of the integral imaging optical reconstruction stage. By displaying the acquired elemental images on a display device (LCD), light from the display device retraces through the lenslets and projects the elemental images onto the focal plane of the lenslet array. The overlap between all the projected elemental images converges in the 3-D space to form a real 3-D image. Since the observer's perspective is opposite to the lenslet array, the convex and concave portions of

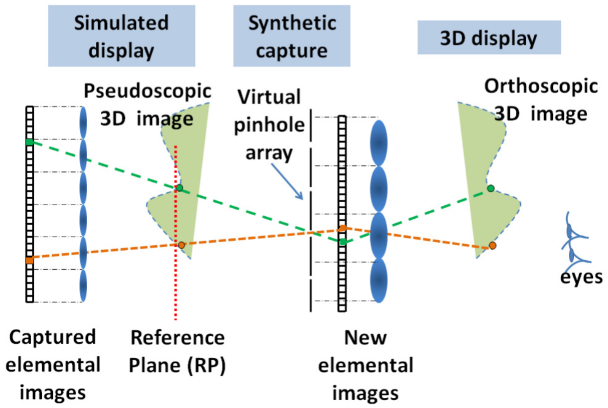

Fig. 4. ***Concept of the smart pseudoscopic-to-orthoscopic conversion method.***

the 3-D image appear reversed for viewers as a pseudoscopic 3-D image.

In order to convert a pseudoscopic (depth inverted) 3-D image to an orthoscopic (correct depth) 3-D image, one solution is to rotate each elemental image by 180° along its center. The 3-D image will form behind the lenslet as a virtual image [68]. A more general digital method named smart pseudoscopic-to-orthoscopic conversion is presented in [69], [70]. As shown in Fig. 4, smart pseudoscopic-to-orthoscopic conversion first performs a simulated display for the captured elemental images on a specific reference plane, then a new set of elemental images is generated by synthetic capture through a virtual pinhole array. Smart pseudoscopic-to-orthoscopic conversion allows for pseudoscopic to orthoscopic transformation of the 3-D image capable of adjusting the display parameters, which makes it a robust approach with various applications [34], [71], and [72].

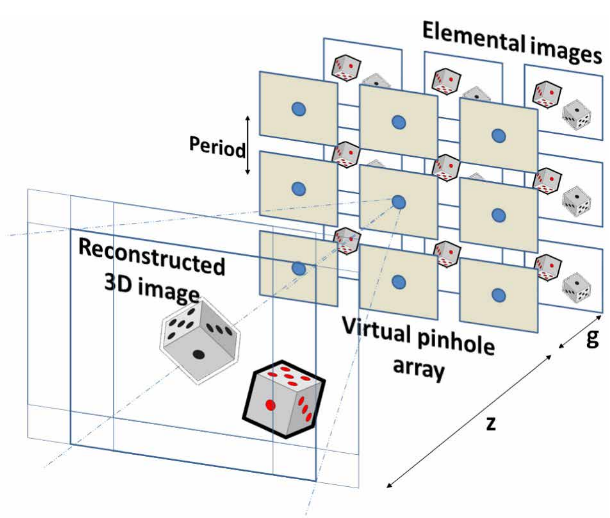

Fig. 5. ***Computational volumetric reconstruction of integral imaging.***

2) Computational Volumetric Reconstruction: Three-dimensional integral imaging visualization can be accomplished by computational volumetric reconstruction [73]–[75]. Since reconstruction is the inverse process of the pickup stage, volume pixels can be reconstructed at arbitrary distances from the display plane by computationally simulating optical reconstruction based on ray optics. As illustrated in Fig. 5, the captured 2-D elemental images are inversely mapped using a computationally synthesized virtual pinhole array and superimposed into the object space. For a specific reconstruction plane (z), the computationally reconstructed image $R(x, y; z)$ can be expressed as

$$R(x,y;z) = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N} EI^{i,j}\left(x + \frac{c_x^{i,j}}{r_z}, y + \frac{c_y^{i,j}}{r_z}\right) \quad (1)$$

where M, N are the number of elemental images in the x and y coordinates, $EI^{i,j}$ is the intensity of the elemental image in the i^{th} column and j^{th} row, $(c_x^{i,j}, c_y^{i,j})$ represents the position of the $(i, j)^{\text{th}}$ image sensor, and $r_z = z/g$ is the magnification factor. The 3-D image is represented by a collection of all the reconstructed planes within the depth range (Z_{range}). Note that for the computational reconstruction, we have not considered the effects of diffraction. If we do so, it will deteriorate the reconstruction. For optical reconstruction, the pinhole array would deteriorate the reconstruction due to diffraction effects.

In certain 3-D pickup geometries, the accurate sensor position and rotation may be difficult to measure if the sensors are on a moving or flexible surface, or if they are randomly distributed [76]. A camera pose estimation algorithm to estimate a camera's position without rotation was combined with an integral imaging reconstruction method in [77] and [78]. By using two known sensors' positions and rotations, the position and rotation of the rest of the sensors can be estimated using the two-view geometry theory and the camera projective model. The estimation method can be used to improve the quality of the 3-D reconstruction if measurement errors exist.

3) Three-Dimensional Profilometric Reconstruction: Three-dimensional information can be visualized as a 3-D profile of the scene. In [79], a method is proposed to estimate the depth information of a scene using a minimum variance (Min-Var) criterion. Considering a spectral radiation pattern function in the 3-D scene and relating it to various perspective elemental images, the depth of a convex surface under Lambertian illumination can be statistically inferred. Let us consider that the radiation intensity propagation in direction (θ, φ) and with wavelength, λ, is represented by the spectral radiation pattern function, defined as $L(\theta, \varphi, \lambda)$, which corresponds to a certain point $(x, y; z)$ in the 3-D space. Suppose that a set of elemental images is captured within an $M \times N$ planar grid. The variance of the spectral radiation pattern function is

$$D(x,y;z) = \frac{1}{3}\sum_{w=1}^{3}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[L(\theta^{i,j},\varphi^{i,j},\lambda^{w}) - \overline{L}(\theta,\varphi,\lambda^{w})\right]^{2}_{(x,y;z)} \quad (2)$$

where $\overline{L}$ is the mean value of the spectral radiation pattern function over all of the directions (perspectives), and w represents the color channel of the digital elemental images.The variance along the depth range (Z_{range}) of the 3-D scene will reach a minimum value when the point is located on an object surface. Depth information can be computed by searching the minimum variance of $D(x, y)$ throughout Z_{range}:

$$\hat{z}(x, y) = \underset{z \in Z_{range}}{\arg\min} D(x,y;z). \quad (3)$$

Combining the depth information and the 2-D elemental images, a 3-D profile of the scene can be reconstructed.

4) Depth Estimation Using a Photoconsistency-Based Criterion: Recently, a depth estimation method through a photoconsistency criterion based on a voting strategy has been presented in [80]. The proposed approach (hereafter called Max-Voting method) is based on a soft-voting procedure that takes into account the level of agreement (similarity) among the different camera views, using a similar strategy to those presented in [81]and [82]. The main idea for the voting process is that when an object is in focus at a certain depth level z, the pixels of each camera corresponding to that object should have a close color or intensity value among them, i.e., they should accomplish a so-called photoconsistency criterion. Although the concept of the Max-Voting method is similar to the minimum variance (Min-Var) criterion [see (3)], the proposed method takes into account the local information around each pixel, i.e., a weight is given depending on color or grey scale values of the pixels in the neighborhood of another one. However, the Min-Var method does not take this into account.

Consider an integral imaging reconstruction process. At a certain depth range $z \in Z_{range}$, the pixel at the position (i, j) of the image I and its square surrounding window W are defined as $W_{ij} = \{I(i + x, j + y) : -\tau \le x, y \le \tau\}$, where τ defines the window size. Suppose a squared camera array, where $||C||$ is the number of cameras whose central camera is $R \in C$ and I is the depth reconstruction at depth z. For each pixel (i, j) and its neighboring pixels (x, y) within the window W_{ij} (i.e., $\forall (x, y) \in W_{ij}$), we proposed in [80] a criterion based on a voting procedure where each camera votes in favor of the pixel (i, j) at depth level z depending on the similarity of the pixel intensities of each camera as compared to camera R. A threshold value (THR) is also assigned that denotes whether this similarity is good enough.

Similarity is measured using the Euclidean distance d between the a*b* values (from the L*a*b* color space) for each pixel. For the voting strategy, each camera's vote is weighted depending on the distance d, which is equal to 1 when the distance is zero, and decreasing exponentially until 0 when d is greater than the threshold (THR).

We can mathematically model the camera array elemental images $E(p_1, p_2, p_3)$, where p_1 and p_2 are the pixel coordinates and p_3 is the camera number. Thus, centered on the pixel position (i, j), for each neighborhood pixel $(x, y) \in W_{ij}$ and $\forall C^k \in C$, the distance d_{ij} is defined as the Euclidean distance among the pixel (i, j) from camera R and the pixels (x, y) from each camera C^k:

$$d_{ij}(x,y) = \sqrt{\sum_{k}^{||C||}\left(E(x,y,C^k) - E(i,j,R)\right)^2}. \quad (4)$$

The camera R never changes, distance d_{ij} is obtained for the pixel (i, j) at each position of the window W_{ij} and summed up as follows:

$$V(i,j,z) = \frac{\sum_{(x,y)}^{W_{ij}} e^{\frac{-[d_{ij}(x,y)]^2}{THR}}}{O_{ij}}. \quad (5)$$

The voting value is also weighted by O_{ij} to consider only the cameras that "see" the pixel (i,j), because some parts of the scene in R do not appear in other cameras. Thus a correct weight should only include those cameras that really contribute during the process.

Several experiments were conducted on synthetic images generated in 3ds Max software to computationally create two 3-D scenes where we can put a camera array and synthetically generate elemental images. It is generated in this way because it allows us to have the depth ground truth for the objects in the scene. We can then use the root mean square error (RMSE) to evaluate the depth estimation methods.

The experimental setup conditions for each one of the synthetic scenes can be found in Table 1. The second and third columns show the camera square array configuration and the depth range from Z_{min} to Z_{max} with a step size of Z_{step}. The fourth and fifth columns give the physical size of the camera sensor (c_x, c_y) in each direction ("x" and "y"), and the period of the cameras (p). The units for columns 3–5 are centimeters for the bathroom scene and millimeters for the Beethoven scene. The focal length of the camera is f = 50 mm.

Fig. 6(a) shows the elemental image corresponding to camera R for the Bathroom and the Beethoven synthetic images. The first column in Fig. 6(b) shows the depth map obtained using 3ds Max. The synthetic images show the indoor spaces (Bathroom) and a foreground image of a Beethoven bust. The second and third columns of Fig. 6(b) show the depth estimation results of the Min-Var

Table 1 Experimental Setup Features for Synthetic Images Created in 3dsMax [80]

Image name	Camera configuration	Z_{min}:Z_{step}:Z_{max}	(c_x, c_y)	p
Bathroom	7 × 7	220:10:830	(36, 36)	5
Beethoven	7 × 7	139:1:341	(36, 36)	5

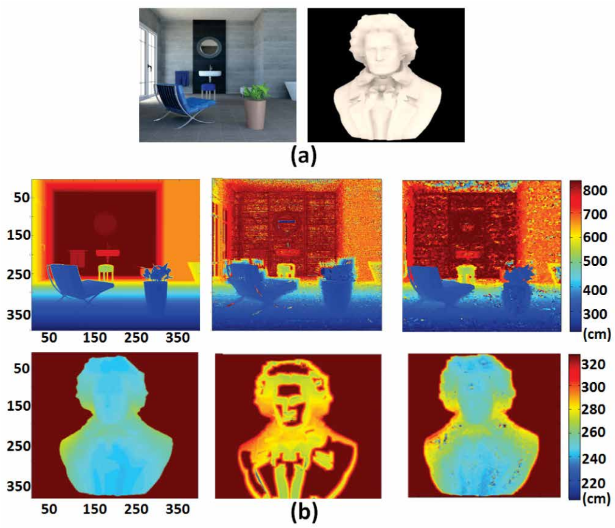

Fig. 6. ***(a) Synthetic images. Bathroom (left) and Beethoven (right) images. (b) Obtained depth maps. From left to right columns, ground-truth of the depth map, Min-Var method, and Max-Voting method [80].***

Table 2 RMSE Results for Bathroom Image While Window Size Increases. From Left to Right, Window Size Is Increasing From 3 × 3 to 13 × 13 [80]

Window size	3×3	5×5	7×7	9×9	11×11	13×13
Bathroom (RMSE)	77.98	64.45	59.09	56.32	54.89	54.02

Table 3 Quantitative Results on Synthetic Images (I). From Left to Right, in Blocks, Images, Results for Min-Var Approach and Results for the Max-Voting Approach [80]

Scene name	Min-Var Large Error	Max-Voting Large error
Bathroom (cm)	12.60%	**8.32%**
Beethoven (mm)	49.01%	**6.29%**

Table 4 Quantitative Results on Synthetic Images (II). Second and Third Blocks Show RMSE Values Obtained on Each Image (RMSE Column) and RMSE Obtained if Pixels With High Errors Are Not Taken into Account (RMSE* column) [80]

Scene name	Min-Var		Max-Voting	
	RMSE	RMSE*	RMSE	RMSE*
Bathroom (cm)	85.14	28.10	**64.45**	**27.16**
Beethoven* (mm)	81.85	45.66	**43.63**	**34.25**

and Max-Voting methods for the Bathroom and Beethoven images for a 5 × 5 window size and $THR = 1$.

Fig. 7 shows the results for the Bathroom image, where different window sizes have been applied, for $THR = 1$. From left to right, we show the generated depth map by the Max-Voting algorithm considering the following window sizes: 3 × 3, 7 × 7 and 13 × 13. We can see how the increase in the window size makes the results smoother; however, some details are lost.

The RMSE figure of merit has been chosen to evaluate the depth estimation results. Table 2 shows that the error is progressively lower when a bigger window size is used.

Tables 3 and 4 show the depth estimation error results obtained using the Min-Var method and the Max-Voting methods. Table 3 shows how errors in the scene are distributed. In particular, it shows the number of pixels (in percentage) whose errors are substantially large. The threshold value for considering large errors has been set to 100 cm for the Bathroom image and 50 mm for the Beethoven image.

Table 4 shows the RMSE values (expressed in centimeters or millimeters depending on the image) and the RMSE obtained if those pixels with high errors are not taken into account, showing that most of the RMSE error made by the algorithms is concentrated on a few pixels. Table 4 also shows that the real performance of the methods substantially improves if these pixels are not taken into account.

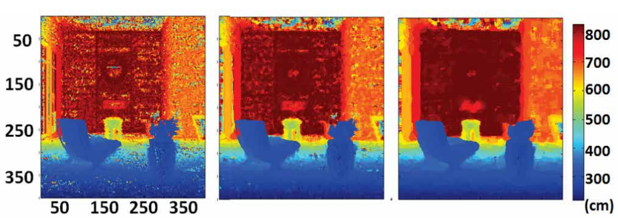

Fig. 7. ***Effect of the window size on the depth estimation for the Bathroom scene. From left to right, window sizes of 3 × 3, 7 × 7 and 13 × 13 [80].***

IV. THREE-DIMENSIONAL POLARIMETRIC INTEGRAL IMAGING

A passive polarimetric integral imaging technique has been used for 3-D polarization measurement, and optical or computational 3-D visualization [46], [47], [83], and [84]. In this section, we present the results obtained by polarimetric 3-D sensing and visualization systems under various conditions.

A. Linear Illumination Condition and Optical 3-D Integral Imaging Display

The reflected light from a scene illuminated using linearly polarized light can be recorded as an elemental image array using a linear polarizer and a lenslet array (Fig. 1). By placing a rotating linear polarizer between the acquisition system and the lenslet array, the Jones vector of the polarized light reflected from the object surface can be determined for the measurement of the polarization state of light. The elliptically polarized light can be modeled in terms of the Jones vector as

$$E = \begin{bmatrix} \cos\theta \exp(i\delta) \\ \sin\theta \end{bmatrix} \tag{6}$$

where θ represents the rotation between the principal axes of the polarization vector in relation to the horizontal axis and δ is the phase retardation between the orthogonal polarimetric components [46].

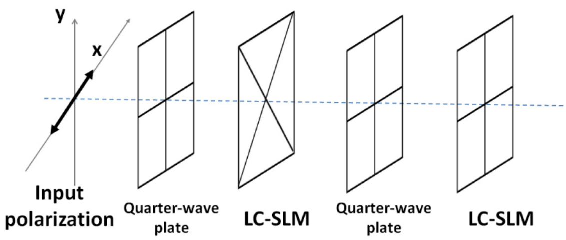

Fig. 8. ***Diagram of the optical system to obtain an arbitrary state of elliptical polarization from the elemental images. Lines in the quarter-wave plate and liquid-crystal spatial light modulator (LC-SLM) denote the principal axes [46].***

Once the elemental images are captured, and the polarization state is obtained, the 3-D scene with a particular polarization distribution can be optically reconstructed. An optical system, as shown in Fig. 8, can generate arbitrary states of polarization for each elemental image [85]. The elemental images can be displayed using spatial light modulators (SLMs) with two quarter-wave plates. The SLMs can be any type of LC-modes which can switch between 0–$\lambda/4$. The sample used in this experiment is TN-mode due to its large aperture ratio. In the future, FFS mode can be implemented to further enlarge the viewing angle. Moreover, the 3-D objects can be visualized with the polarimetric information. In this case, the mathematical expression for the Jones vector is [46]

$$M = \begin{bmatrix} \exp(i\delta) & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\frac{\varphi}{2} & i\sin\frac{\varphi}{2} \\ i\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix} \times \begin{pmatrix} -i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{bmatrix} \exp(i\delta)\cos\frac{\varphi}{2} \\ -\sin\frac{\varphi}{2} \end{bmatrix} \tag{7}$$

where φ is the angle along the direction of polarization, and δ represents the phase retardation between the two orthogonal components [46], [85], δ and φ denote the amount of phase retardation between the orthogonal axes and the rotation angle of the polarization direction, respectively. The variable rotation angle φ of the principal axes at each pixel can be realized by combining two quarter-wave plates and a phase-only liquid crystal SLM.

B. Natural Illumination Condition and Computational 3-D Integral Imaging Reconstruction

A 3-D polarimetric computational integral imaging system has been presented in [83]. This system can measure the polarimetric information of the 3-D scene with natural illumination using the Stokes parameters. The Stokes vectors [44] can be defined as follows:

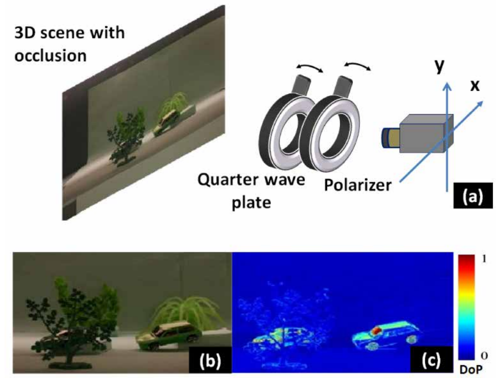

Fig. 9. ***(a) Schematic of the polarimetric 3-D pickup system based on synthetic aperture integral imaging. (b) Example of captured elemental images. (c) Polarimetric elemental image with the degree of polarization corresponding to (b).***

$$\begin{cases} S_0 = E_{0x}^2 + E_{0y}^2 \\ S_1 = E_{0x}^2 - E_{0y}^2 \\ S_2 = 2E_{0x}E_{0y}\cos\delta \\ S_3 = 2E_{0x}E_{0y}\sin\delta \end{cases} \tag{8}$$

where E_{0x} and E_{0y} are the instantaneous amplitudes of the orthogonal components of the electric field, and δ is the instantaneous phase factor of the plane wave. The Stokes parameters of interest are denoted as $S_i, i = 0,\ldots,3$. The Stokes parameters enable us to describe the degree of polarization (DoP) for any state of polarization

$$DoP = \frac{I_{pol}}{I_{tot}} = \frac{\left(\sqrt{S_1^2 + S_2^2 + S_3^2}\right)}{S_0}, \quad 0 \le DoP \le 1 \tag{9}$$

where I_{pol} is the sum of the polarized intensity of the light beam, and I_{tot} is the total intensity of the light beam. When DoP is 1, the measured light is completely polarized, and when DoP is 0, the light is unpolarized. The degree of linear polarization (DoLP) and the degree of circular polarization (DoCP) can be expressed as $DoLP = \sqrt{S_1^2 + S_2^2}/S_0$ and $DoCP = \sqrt{S_3^2}/S_0$, respectively.

The 3-D polarimetric sensing system based on the moving sensor array synthetic aperture integral imaging [65] technique is shown in Fig. 9(a). A linear polarizer and a quarter-wave plate are combined and placed in front of a digital camera for polarimetric imaging [86], [87]. Using (8), the Stokes parameters are measured as

$$\begin{cases} S_0 = I^{0°,0} + I^{90°,0} \\ S_1 = I^{0°,0} - I^{90°,0} \\ S_2 = I^{45°,0} + I^{135°,0} \\ S_3 = I^{45°,\pi/2} - I^{135°,\pi/2} \end{cases} \tag{10}$$

where I is the intensity of captured polarimetric images, $I^{\alpha°,0}$ represents the linear rotating polarizer with an angle of α degrees in relation to the x axis. $I^{\alpha°,\pi/2}$ indicates that a quarter-wave plate is combined with the polarizer. When light passes through the quarter-wave plate, one-quarter of a wave phase shift between the orthogonal components is introduced. In the measurement, the wave plate is fixed with its fast axis along the x axis and the transmission axis of the linear polarizer is rotated with α°. A total of six sets of polarimetric elemental images is needed for each sensor position.

In Fig. 9(a), an experiment is depicted where two cars were placed at a distance of 530 mm from the sensor and a moving camera was used to record images. A car was occluded by a tree approximately 450 mm from the sensor while two trees were located approximately 720 mm away. A total of 6 × 6 elemental images were taken at different positions for 3-D sensing. Fig. 9(b) gives an example of the captured elemental image. The polarimetric image with the measured degree of polarization is illustrated in Fig. 9(c). The degree of polarization of the light reflected from the surface of the occluded car is larger than the degree of polarization of the occlusion and the background [83].

The 3-D integral imaging computational reconstruction algorithm was modified to combine the polarization information and the original pixel information of the captured elemental images. Three-dimensional reconstruction can be implemented with a threshold (THR) applied to the degree of polarization images. Only pixels whose respective degree of polarization is higher than the threshold will contribute to the reconstruction. The reconstructed 3-D images contain both the depth information of the scene and the polarization state of the light reflected from the object surface. The experimental results of the 3-D polarimetric reconstruction are given in Fig. 10. The 3-D images obtained by the conventional integral imaging computational reconstruction method are shown in Fig. 10(a). The objects in the 3-D scene are in focus at their respective depth positions. Fig. 10(b) illustrates the 3-D polarimetric reconstructed images at the respective depth with a threshold of $THR = 0.2$. With this threshold, only the surface (cars) with higher degree of polarization reflected light is reconstructed, and the occlusion and background can be avoided because of the relatively lower degree of polarization information. The 3-D computational polarimetric integral imaging can be used for material inspection and detection which will be discussed in Section VII-A.

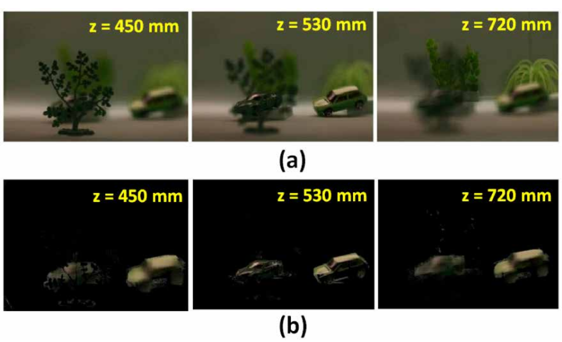

Fig. 10. ***Integral imaging depth reconstruction results at 450 mm, 530 mm and 720 mm. (a) 3-D reconstructions using conventional integral imaging. (b) 3-D reconstructions using the polarization state of each pixel. Degree of polarization threshold is 0.2 [83].***

C. Three-Dimensional Polarimetric Integral Imaging in Photon Starved Conditions

In [84], a method for polarimetric 3-D integral imaging in photon starved conditions was proposed. As the photon counting images captured under low light illumination conditions are very sparse, the Stokes parameters and the degree of polarization are difficult to measure with conventional methods. By using the maximum likelihood estimation method, polarimetric 3-D integral images are generated [88]. In order to obtain high quality polarimetric reconstructed images, a total variation denoising filter is implemented to efficiently remove the noise from the image and preserve the signal corresponding to the scene [89], [90]. The extracted polarimetric features can be used in pattern recognition algorithms.

As discussed in Section IV-B, a quarter-wave plate and a linear polarizer can be combined and placed ahead of the sensor for polarimetric imaging. 3-D polarimetric elemental images can be obtained by capturing the polarimetric distributions $i_{k,l}^{0°,0}$, $i_{k,l}^{90°,0}$, $i_{k,l}^{45°,0}$, $i_{k,l}^{135°,0}$, $i_{k,l}^{45°,\pi/2}$ and $i_{k,l}^{135°,\pi/2}$ at each camera position (k, l). For the case of an integral imaging acquisition process in photon starved conditions, the photon counting model should be used. The detected photons in the captured images can be simulated using the Poisson distribution function [91]

$$P(m;x,y) = \frac{\left[n_{k,l}^{\alpha,\beta}(x,y)\right]^{m} \exp\left[-n_{k,l}^{\alpha,\beta}(x,y)\right]}{m!} \tag{11}$$

where (x, y) is the pixel index, m represents the number of the photons that have been detected. α is the degree of the linear rotating polarizer in relation to the x axis, and β represents the quarter-wave plate. If $\beta = \pi/2$, a quarter-wave plate is placed in front of the linear polarizer, otherwise ($\beta = 0$) only the linear polarizer is used for the acquisition. $n_{k,l}^{\alpha,\beta}(x,y)$ is the normalized irradiance [84]

$$n_{k,l}^{\alpha,\beta}(x,y) = \frac{N_p\, i_{k,l}^{\alpha,\beta}(x,y)}{\sum_{x,y} i_{k,l}^{\alpha,\beta}(x,y)} \tag{12}$$

where N_p is the number of photon counts predetermined in the scene. As discussed in [92], using the maximum likelihood estimation for integral imaging, the 3-D reconstruction can be obtained by averaging the normalized photon counting irradiance ($\hat{i}_{k,l}^{\alpha,\beta}$) from the captured polarimetric elemental images ($i_{k,l}^{\alpha,\beta}$). Using (12) and (1), the photon counting 3-D polarimetric reconstructed image can be expressed as

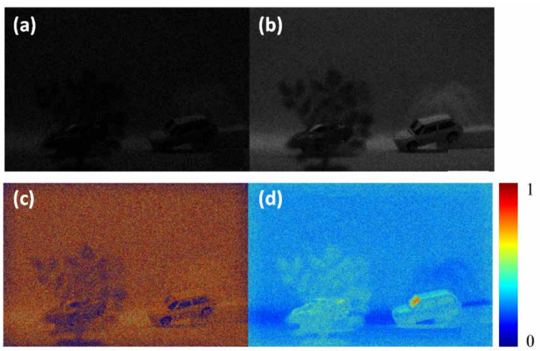

Fig. 11. ***Three-dimensional integral imaging reconstruction results for photon counting conditions using maximum likelihood estimation at z = 530 mm: (a) 0.01 photons per pixel, and (b) 0.05 photons per pixel. Degree of polarization for integral imaging with photon counting: (c) 0.01 photons per pixel, and (d) 0.05 photons per pixel [84].***

$$\hat{I}^{\alpha,\beta}(x,y;z) \propto \sum_{k=1}^{M}\sum_{l=1}^{N} i_{k,l}^{\alpha,\beta}\left(x+\frac{c_x^{k,l}}{r_z}, y+\frac{c_y^{k,l}}{r_z}\right). \quad (13)$$

The Stokes parameters and degree of polarization can be calculated using (10) and (9), respectively.

Experiments were performed with the same setup described in Section IV-B. Photon counting imaging was computationally applied to the elemental images shown in Fig. 9(b). The model of the recording device used for generating the polarimetric photon counting elemental images is a binary photon counting camera and the elemental images were statistically transformed by the Poisson distribution. As a result, we were able to determine the number of photons per pixel. We arbitrarily chose 0.01 and 0.05 photons per pixel, as images with such few photons have limited information. The reconstructed results using maximum likelihood estimation are noisy and have a low dynamic range. Fig. 11(a) and (b) shows the 3-D reconstructed results using the maximum likelihood estimation. The elemental images used for the 3-D reconstruction contained few photons. In the 3-D reconstruction process, noise will be dominant under the photon starved conditions due to the low SNR value, and noise may degrade the quality of the measurement of the recorded polarization. Fig. 11(c) and (d) illustrates the degree of polarization measured by the maximum likelihood estimation using (10) and (13). As shown in Fig. 11(c), when the number of photons is relatively low, (0.01 photons/pixel), the whole scene was measured with high polarized characteristics. For the case where the number of photons per pixel increases to 0.05, as shown in Fig. 11(d), the degree of polarization result improves. However, the background areas, as shown in Fig. 11(d), still have high degree of polarization values compared to the results with Fig. 9(c).

The mean structural similarity index measure (MSSIM) [93] was implemented to quantitatively compare two reconstructed images using the 3-D integral imaging

Table 5 MSSIM Comparison Results Using Maximum Likelihood Estimation and Degree of Polarization in Fig. 11 [84]

	0.01 photons / pixel	0.05 photons / pixel
3D reconstructed image	0.029	0.054
Degree of Polarization	0.001	0.002

computational reconstruction method. In order to compare two images (X and Y), MSSIM considers a set of M 8×8 pixels subimages $\{x_j\}$ and $\{y_j\}$ obtained from X and Y. The local structural similarity index measure (SSIM) between the respective subimages x_j and y_j is [84]

$$SSIM(x_j,y_j) = \frac{(2\mu_{xj}\mu_{yj}+c_{1j})(2\sigma_{xyj}+c_{2j})}{\left(\mu_{xj}^2+\mu_{yj}^2+c_{1j}\right)\left(\sigma_{xj}^2+\sigma_{yj}^2+c_{2j}\right)} \quad (14)$$

where μ_{xj} and μ_{yj} are the averages of the subimages x_j and y_j, respectively. σ_{xj}, σ_{yj} and σ_{xyj} are the variances of x_j, y_j and the covariance, respectively. c_{1j} and c_{2j} are two tuning parameters which correspond to the square of the dynamic range (D), i.e., $c_{1j} = (k_{1j}\,\mathrm{D})^2$ and $c_{2j} = (k_{2j}\,\mathrm{D})^2$. The dynamic range of an image depends on the maximum and minimum pixel values, the k_1 and k_2 values used were 0.01 and 0.03 respectively. Finally, the MSSIM is obtained by averaging the SSIM over all windows:

$$MSSIM(X,Y) = \frac{1}{M}\sum_{j=1}^{M} SSIM(x_j,y_j). \quad (15)$$

Since this index is normalized, MSSIM $= 1$ only if the two images (X, Y) are the same. The MSSIM values for the 3-D reconstruction and the corresponding degree of polarization results in Fig. 11 are shown in Table 5. The

Table 6 MSSIM Comparison Results Using Maximum Likelihood Reconstruction and Total Variation Minimization Denoising in Fig. 12 [84]

	0.01 photons / pixel	0.05 photons / pixel
3D reconstructed image	0.727	0.909
Degree of Polarization	0.325	0.666

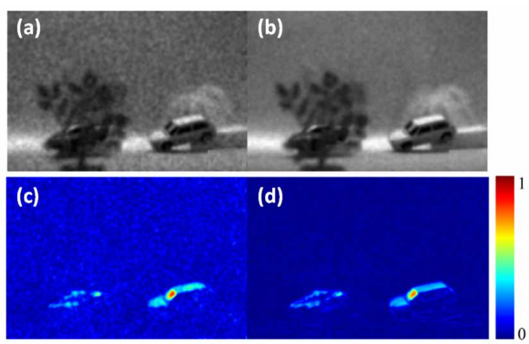

Fig. 12. ***Three-dimensional integral imaging reconstruction under photon starved conditions at z = 530 mm. Total variation minimization is applied on the reconstructed image with (a) 0.01 photons per pixel, and (b) 0.05 photons per pixel. Total variation minimization is then applied to the polarimetric images at z = 530 mm with (c) 0.01 photons per pixel, and (d) 0.05 photons per pixel [84].***

reference reconstructed image and degree of polarization are in Fig. 9(b) and (c), respectively. The MSSIM values for both reconstructed image and the degree of polarization are very small.

In order to improve the 3-D reconstruction results of integral imaging with photon counting, total variation denoising filters are used, since these filters are able to remove noise from the image without affecting the signal [90]. In the experiment, the authors used the Chamolle approach [94] implemented in the scikit-image library [95]. Using the nomenclature in [89], the total variation denoising strategy can be mathematically written as [84]

$$\min_{\mu}\left[\int|\mu_0-\mu|^2\,dxdy+\gamma\int\sqrt{\left(\frac{\partial\mu}{\partial x}\right)^2+\left(\frac{\partial\mu}{\partial y}\right)^2}\,dxdy\right] \tag{16}$$

where γ is a regularization parameter, μ_0 is the noisy image, and μ is the reconstructed image.

The photon counting 3-D reconstructed image processed by the total variation method is shown in Fig. 12(a), and the corresponding degree of polarization is illustrated in Fig. 12(b). Using the total variation denoising, the corresponding degree of polarization was obtained from the denoised version of the photon counting 3-D polarimetric reconstructed images $\hat{I}^{\alpha,\beta}$ [see (13)]. The degree of polarization results shown in Fig. 12 (c) and (d) is similar to the reference degree of polarization image [Fig. 9(c)]. The MSSIM comparison results for the maximum likelihood reconstruction and total variation minimization are presented in Table 6. Note that for 0.05 photons/pixel, the MSSIM value was approximately 1 for the reconstructed 3-D image while the MSSIM value for the degree of polarization image was about 0.67.

V. THREE-DIMENSIONAL INTEGRAL IMAGING IN THE INFRARED DOMAIN

Multispectral imaging allows for the acquisition of images in a series of spectral bands. Nowadays, image acquisition capabilities have extended from the visible spectrum to the near-infrared (NIR), mid-wave-infrared (MWIR) [9] and long-wave-infrared (LWIR) ranges. Moreover, multispectral imaging has been used in fields such as medical imaging and remote sensing. In this section, we present 3-D integral imaging acquisition and visualization methods in the infrared domain.

A. Long-Distance Mid-Wave-Infrared Integral Imaging

We have demonstrated that integral imaging can work well for long-range objects. This section describes an overview of the work about synthetic aperture integral imaging 3-D acquisition and reconstruction of scenes in short range (indoor scenes) and long-range distances of up to 2 km using sensors that operate in the visible, MWIR and LWIR ranges [9], [24], and [102].

1) High-Level Illumination Conditions: An Aura imaging system (working in the MWIR range) was used to acquire a group of 10 elemental images of an airfield base. The elemental images were acquired with a horizontal-only movement of the camera. The horizontal pick up range was 7 m, and the acquisition positions were periodically spread over this range for the corresponding number of camera acquisitions. The camera has a 120 mm lens and pixel size of 19.5 μm. Each elemental image has a resolution of 1024 × 1024 pixels. Fig. 13(a) shows an example of an elemental image of this airfield scene.

A technique based on a 3-D template matching approach for robust depth estimation in the mid-wave-infrared range was developed. The template data was selected from one elemental image and the function that was optimized is

$$z' = \arg\min_Z\left\{\sum_x\sum_y[T(x,y)-R(x,y;z)]^2\right\} \tag{17}$$

where $T(x, y)$ is the template and $R(x, y; z)$ is the reconstructed scene for each depth. Fig. 14 shows a diagram of the template matching strategy for depth estimation. Results using this search algorithm for four targets at known ranges are shown in Table 7. Fig. 13(b) shows the resulting range map for the scene. We can see that the overall range

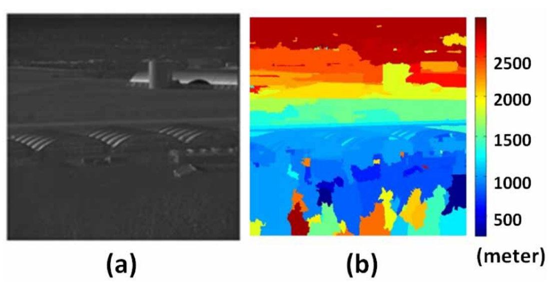

Fig. 13. ***(a) Captured elemental image. (b) Range map using automatic segmentation [9].***

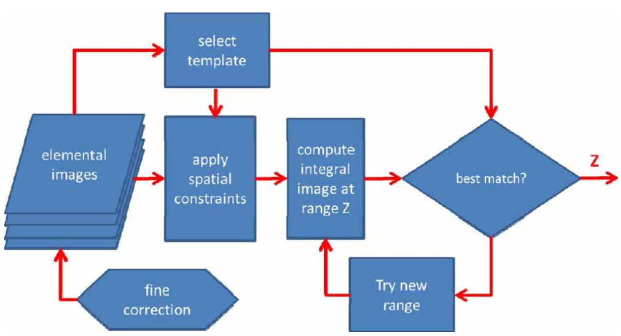

Fig. 14. ***Overview of the range search algorithm [9].***

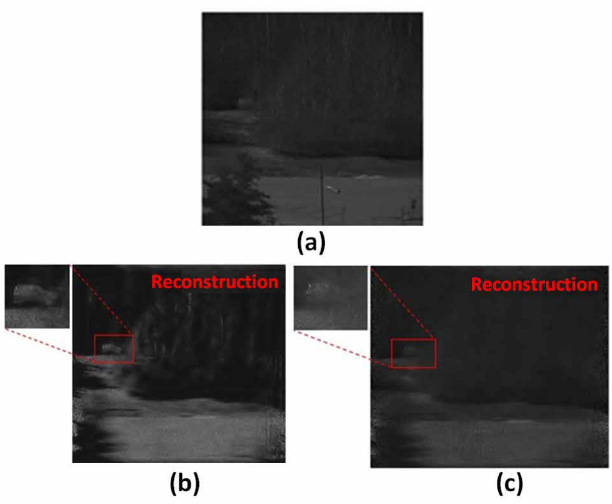

Fig. 15. ***(a) An example of an elemental image used with a road, a series of trees and a car behind one of them. Reconstruction results for z = 237 m when $N_p = 3 \times 10^5$ photon counting photons per elemental image exist, (b) without ($N_{dc} = 0$) and (c) with ($N_{dc} = 10^6$) photons corresponding to dark current noise, respectively [24].***

estimation results are correct, but the method finds some problems on trees that are closer to the camera array.

2) Photon Counting Illumination Conditions for Long-Range Objects: Another group of elemental images consisting of a scene with a road, a series of trees and a car behind them, as shown in Fig. 15(a), was captured and a photon counting simulation process was applied on the elemental images [24]. Our working assumption is that infrared detectors are applied in a regime for which the measured photon counts are well approximated by Poisson distribution [96].

The likelihood of the irradiance estimation from the photon counting elemental images can be modeled as [92]

$$L\left[I_p^{z_0} \mid C_{kl}(p + \Delta p_{kl})\right] = \prod_{k=1}^{K}\prod_{l=1}^{L}\Pr\left(C_{kl}(p + \Delta p_{kl}) \middle| I_p^{z_0}\right) \tag{18}$$

where $p \equiv (x, y)$ is an object pixel position for an elemental image, and Δp_{kl} describes its shift on each elemental image. $C_{kl}(p + \Delta p_{kl})$ is the value of pixel $p + \Delta p_{kl}$ for elemental image of indices kl in photon counting conditions.

Table 7 Estimated Range Results Using Method Shown in Fig. 14 [9]

Measured (m)	Estimated (m)	Δ (m)
666	710	-44
969	1015	-46
1429	1443	-14
2241	2065	176

The maximum likelihood irradiance estimation is given by [24]

$$\tilde{I}_p^{z_0} = C_N \cdot \sum_k \sum_l C_{kl}(p + \Delta p_{kl}). \tag{19}$$

Therefore, the result of the reconstruction for a specific depth also gives an estimation of its corresponding irradiance.

Three noise levels (layers) were added to each one of the elemental images $N_{dc} = \{10^4, 10^5, 10^6\}$ simulating the existence of dark current (dc) noise. Some HgCdTe mid-wave-infrared detectors may also have these dark current noise levels (see [25, Fig. 8]).

Fig. 15(b) and (c) shows the 3-D reconstruction for z = 237m when the number of photon counts for each elemental image is $N_p = 3 \times 10^5$, and where the dark current noise level is $N_{dc} = 0$ and $N_{dc} = 10^6$ photons, respectively. The visualization quality of the depth reconstructed images was highlighted using an image denoising technique based on the wavelet shrinkage approach [97]. The threshold value was fixed at $T = 4$. The peak signal to noise ratio (PSNR) was used as a reference for the quality assessment of the photon counting reconstructed scenes. The PSNR is defined as

$$PSNR = 10 \cdot \log_{10}\left[\frac{I_{Max}^2}{MSE(I,\tilde{I})}\right] \tag{20}$$

where $MSE(I,\tilde{I})$ is the mean square error which provides an estimation of the average error per reconstructed pixel of the 3-D scene, and I_{Max}^2 is the square of the maximum value of the original reconstructed scene.

Fig. 16 shows the PSNR value as the number of photon counts per elemental image, N_p increases, for a reconstruction distance of 237 m (the depth where the car behind the trees is in focus), for the following cases: $N_{dc} = \{0, 10^4, 10^5, 10^6\}$. For the case where no dark current noise exists $PNSR \propto \log(N_p)$ as shown in [92]. This theoretical dependence is also shown in Fig. 16.

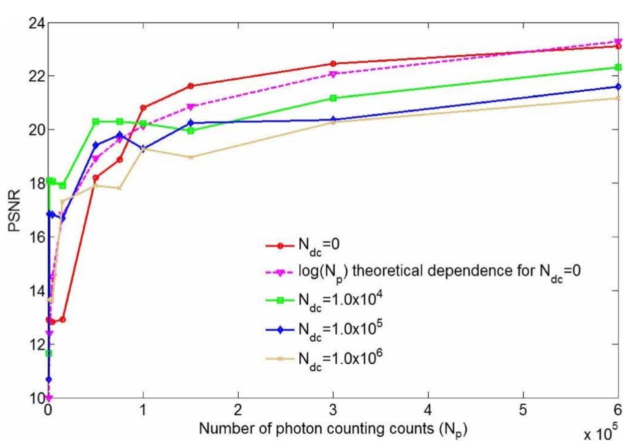

Fig. 16. ***PSNR versus N_p for the trees and car scene reconstructed at Z = 237 m, for N_{dc} = {0, 10^4, 10^5, 10^6} [24].***

We conclude that the RMSE error decreases as N_p increases and therefore the PSNR value increases. On the other hand, we see that the PSNR and N_{dc} noise are inversely proportional.

B. Three-Dimensional Image Reconstruction Using the Alternating Extragradient Method on Mid-Wave-Infrared Photon Counting Images

In this section, we discuss the application of an image restoration method called "Alternating Extragradient Method". This method was recently proposed in [98]. To the best of the authors' knowledge, it is the first time that this method is applied on MWIR photon counting images.

An image with approximately N_p number of photons can be simulated if we consider its normalized irradiance version I_i (such that $\sum_{i=1}^{N_T} I_i$, where i is the pixel number, and N_T the total number of pixels of an image), and assume a Poisson random number with mean parameter $N_p \cdot I_i$. In this framework the Poisson distribution can be written as $Pr(C_i \mid I_i) = \frac{(I_i)^{C_i} \cdot e^{-I_i}}{C_i!}; C_i = 0, 1, 2, \ldots$ where C_i means C photons at pixel i. On the other hand, let us consider the image formation as a linear process.

We can define $C \in \mathrm{R}^{N_T}$ as the detected data, and C_i as the value at each pixel under the assumption that the variable follows a Poisson distribution with expected value $(Hx + b)_i$ and where, on the one hand $x \in \mathrm{R}^{N_T}$ is the scene we aim at recovering, $H \in \mathrm{R}^{m \times N_T}$ models the optical system and $b \in \mathrm{R}^m$ is a positive offset term. We can model this restoration problem as an optimization of the type [99] $\min_{x \geq \eta} f(x) \equiv f_0(x) + \beta \cdot f_1(x)$, where $f_0(x)$ measures data similarity, $f_1(x)$ is a regularizer-type functional and b is an offset. The restored image should have positive values, and therefore $x \geq \eta$, where $\eta \in \mathrm{R}^{N_T}$, $\eta \geq 0$.

It can be shown that this problem has a primal-dual (or saddle-point) equivalent formulation: $\min_{x \in X} \max_{y \in Y} F(x, y)$ where X and Y are two feasible sets with restrictions such that $D = X \times Y$ is a closed and convex domain and F is a smooth convex-concave function [98]. The Kullback-Leibler distance can be used in this case and expressed as follows [98]:

$$\min_{x \in X} \max_{y \in Y} F(x, y) \equiv \sum_i \left\{ C_i \ln \frac{C_i}{(Hx + b)_i} + (Hx + b)_i - C_i \right\} + \beta \cdot y^T \cdot z(x) \tag{21}$$

where $X = \{x \in \mathrm{R}^{N_T} : x \geq \eta\}$ and $z(x)$ and Y are given by: $z(x) = Ax$, $A = (A_1, A_2, \ldots, A_{N_T})^T$ and $Y = \{y \in \mathrm{R}^{2n} : y_{2i-1}^2 + y_{2i}^2 \leq 1; i = 1, 2, \ldots, N_T\}$. $A_k \in \mathrm{R}^{2 \times N_T}$ is a matrix with only two nonzero entries on each row, equal to -1 and $+1$. The Alternating Extragradient Method uses the following three iteration formulas:

$$\begin{cases} \bar{y}^{(k)} = O_Y\left[y^{(k)} + \gamma_k \nabla_y F\left(x^{(k)}, y^{(k)}\right)\right] \\ x^{(k+1)} = O_X\left[x^{(k)} - \gamma_k \nabla_x F\left(x^{(k)}, \bar{y}^{(k)}\right)\right] \\ y^{(k+1)} = O_Y\left[y^{(k)} + \gamma_k \nabla_y F\left(x^{(k+1)}, y^{(k)}\right)\right] \end{cases} \tag{22}$$

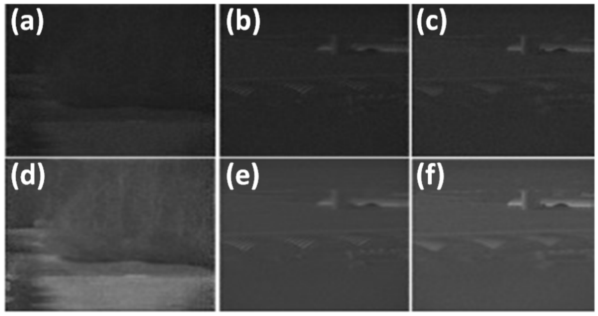

Fig. 17. ***Reconstruction results for two different scenes, considering $N_p = 3.0 \times 10^5$ photons, and where $N_{dc} = 10^6$ photons have also been added (1.24 photons/pixel in total). (a) Depth reconstruction for the scene with a car and trees occluding it, for z = 237 m. (b) Depth reconstruction for the scene of an airfield base for z = 960 m. (c) Depth reconstruction for the same scene of the airfield base for z = 2.2 km. (d)–(f) Reconstruction results for the cases (a)–(c) when using the alternating extragradient method [98] for $\beta = \{0.25, 0.14, 0.14\}$, respectively.***

where O_X and O_Y denote the orthogonal projection operators onto the sets X and Y, and $\gamma > 0$ a constant. We refer the reader to [98] for further algorithmic details.

The elemental images used in this section are the same as those used in Section V-A. The corresponding photon counting elemental images were generated and an additional noise level of $N_{dc} = 10^6$ photons was added to each one of them for both scenes.

Fig. 17 illustrates the results for the integral imaging reconstruction case when $N_{dc} = 10^6$ dark current photons are added to the $N_p = 3.0 \times 10^5$ photons for the previous case, and therefore, a total amount of 1.24 photons/pixel are present in each elemental image. Fig. 17(a) shows the depth reconstruction for the scene with a car and trees occluding the car, for z = 237 m. Fig. 17(b) shows the depth reconstruction for the scene of an airfield base for z = 960 m. Fig. 17(c) shows the depth reconstruction for the base for z = 2.2 km. Fig. 17(d)–(f) shows the reconstruction results for the cases (a)–(c) when using the alternating extragradient method [98] for $\beta = \{0.25, 0.14, 0.14\}$, respectively. Finally, we should stress that the alternating extragradient method outperformed the maximum likelihood method in the whole photon counting domain tested (from 0 to 10^6 photons per elemental image). As shown in Fig. 18, the results obtained by the alternating extragradient method (AEM) were substantially better than those of maximum likelihood (ML) method for the whole photon counting level range considered, both for the case where no dark current was considered as for the case when $N_{dc} = 10^6$ photons were added.

C. Three-Dimensional Imaging in Long-Wave-Infrared Spectrum

Images in the long-wave-infrared (LWIR) range acquire self-radiation of an object rather than the light reflected

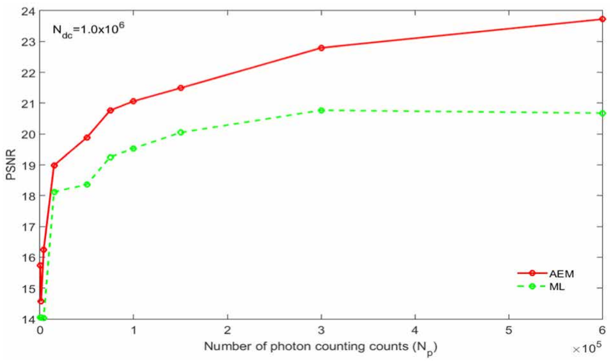

Fig. 18. ***PSNR versus N_p for the results obtained by using the alternating extragradient method (AEM) method and the maximum likelihood (ML) method.***

from an object's surface thus eliminating illumination issues [100]. This makes LWIR imaging especially useful in night time settings. The LWIR sensors capture information between the wavelength range of approximately 8 to 15 μm [100] and was originally used for military applications including surveillance and night vision. This technology has found applications in other fields such as diagnosing inflammations in the legs and hoofs of horses, fungal infections in wheat, and finding the heat loss from air vents and windows [101]. It is worth noting that the resolution of an LWIR image is poorer compared to an image in the visible spectrum due to the longer light wavelength as described by the Abbe diffraction limit [100].

We have implemented passive 3-D imaging using synthetic aperture integral imaging with LWIR imaging for outdoor applications [102]. Three-dimensional image reconstruction can remove occlusions in front of an object assuming there is sufficient distance between the object and the occlusion [103]. To capture the elemental images, the LWIR camera used was a TAMARISK 320 with a resolution of 320 × 240 pixels and pixel size of 17 μm with a field of view (FOV) of 27°. Moreover, the 3-D experiment used a 7 × 3 camera array with a period of 30 mm in a night time setting. The output of the camera is an analog signal which is converted to a digital signal yielding a 640 × 480 pixels image. Fig. 19(a) depicts a sample 2-D elemental image, which contains a person located 14.5 m away occluded by branches. Fig. 19(b) depicts the 3-D reconstruction at $z = 14.5$ m which was able to remove the branches in front of the person.

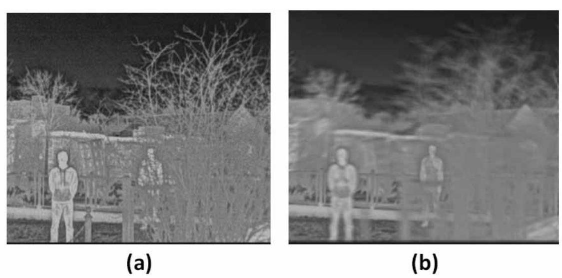

Fig. 19. ***3-D scene captured using LWIR imaging. (a) Person located 14.5 m away is occluded by branches while (b) 3-D reconstructed image at $z = 14.5$ m removes the occlusion in front of the person [102].***

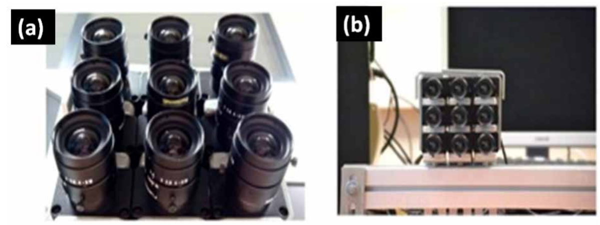

Fig. 20. ***A 3 × 3 camera array used for acquisition of the videos aiming at human activity recognition [104].***

VI. THREE-DIMENSIONAL INTEGRAL IMAGING WITH TIME DOMAIN FOR 3-D GESTURE RECOGNITION

For dynamic objects and targets, information also varies in the time domain. Three-dimensional sensing, processing and visualization in the time domain are also discussed in this section. In this section, we present a 3-D video system by using the integral imaging technique for 3-D human gesture and activity recognition [104].

In order to acquire a series of human actions/gestures, a group of nine cameras in a 3 × 3 array configuration was considered (Fig. 20). This is a synthetic aperture integral imaging system working in the resolution priority integral imaging mode [105]. In particular, nine Stingray F080B/C cameras (with a resolution of 1024 × 768 pixels) were synchronized through a 1394 bus, acquiring videos at 15 frames per second.

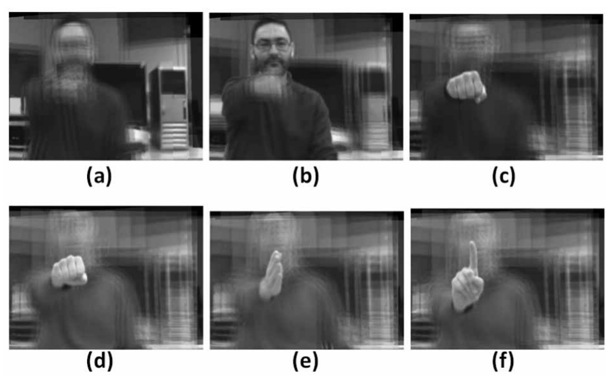

Fig. 21. ***3-D gesture recognition experiments. Images show the reconstruction capability of the system, for the same frame, and for a specific person and gesture: (a) background, (b) head, and (c) fist. Depth reconstruction focusing at the hand's gesture: (d) open, (e) left, and (f) deny [104].***

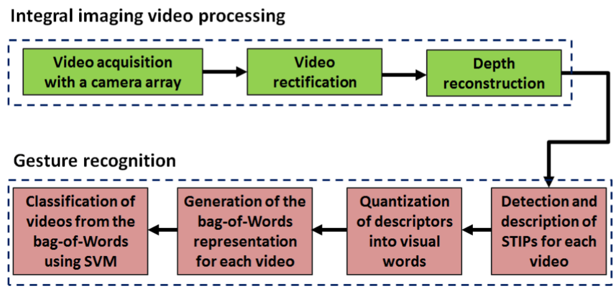

Fig. 22. ***Building blocks of the proposed 3-D human gesture recognition system [104].***

Impact of the errors associated to the position and orientation of the cameras can be diminished if the acquired videos are rectified [106], [107].

Fig. 21(a)–(c) shows the reconstruction capability of the system for the same person and gesture, for three different depths, and the same frame. Fig. 21(d)–(f) shows the reconstructed scene at the depth where the hand (doing the gesture) is approximately in focus. These videos (obtained after the reconstruction process) can be used for 3-D gesture recognition in time domain.

The procedure for gesture recognition as shown in Fig. 22, follows the so-called "bag of visual words" approach [108].

Fig. 23 shows the mean accuracy versus the number of words used. Integral imaging accuracy is higher than monocular imaging for all the number of words used, except for $K = 50$. The Histogram of Optical Flow (HOF) and the Histogram of Oriented Gradients (HOG) are histograms formed from the optical flow or the oriented gradient information in the reconstructed image. Both techniques have shown a great potential when used as feature vectors in order to apply a "bag-of-words" approach for action recognition.

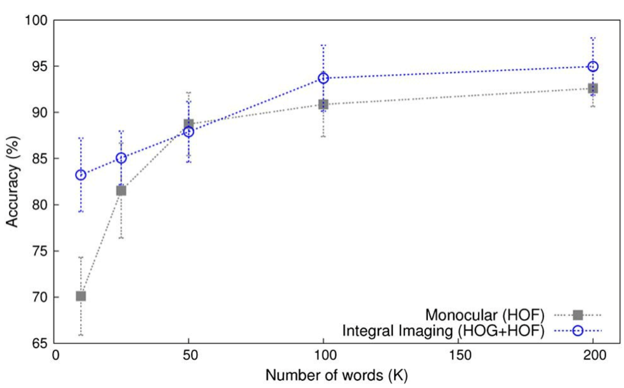

Fig. 23. ***Three-dimensional gesture recognition classification results using the best descriptor in each case [monocular (2-D) and 3-D integral imaging] [104].***

VII. MULTIDIMENSIONAL OPTICAL SENSING AND IMAGING SYSTEM (MOSIS) 2.0

In this section, the recent progress of integral-imaging-based MOSIS 2.0 is presented. MOSIS 2.0 is an improvement of the original concept of MOSIS [51], for object recognition, material inspection, integrated 3-D visualization, etc., which can significantly improve image understanding.

A. Multidimensional Optical Sensing and Imaging System 2.0: Visualization, Target Recognition, and Material Inspection

We present some recent progress on the MOSIS 2.0 for target recognition, material inspection and integrated visualization from a scene.

MOSIS 2.0 is the successor to MOSIS [51]. The degrees of freedom of MOSIS 2.0 include visible and IR bands, including near-infrared (NIR) spectral bands, state of polarization of light reflected from object surface, and depth and directional information of the scene. MOSIS 2.0 uses synthetic aperture integral imaging [65] for 3-D sensing of a complex scene which may include objects with heavy occlusion. Computationally reconstructed images provide in focus information of the objects on the respective depth planes with mitigated occlusion. The 3-D object recognition can be performed on the reconstructed scene. In the experiments, we have used histogram of oriented gradients (HOG) [109] for feature extraction and a support vector machine (SVM) [110]–[112] as a classifier. In MOSIS 2.0 with polarimetric imaging, the degree of polarization of the light reflected from the 3-D scene is calculated using the Stokes parameters. Depth and degree of polarization information are integrated during 3-D reconstruction. The polarimetric characteristics of the reflected light from the object's surface are used for material inspection. By implementing the segmentation algorithm within the multispectral bands, materials with specific spectral properties are extracted. Multidimensionally integrated visualization reveals more information of the scene for improved image understanding and information analysis. Fig. 24 illustrates the diagram of the proposed MOSIS 2.0.

1) Multidimensional Optical Sensing With MOSIS 2.0: In MOSIS 2.0, we have implemented multidimensional image sensing with synthetic aperture integral imaging. Compared with the lenslet based pickup, synthetic aperture integral imaging captures a 3-D scene with a camera array or a moving camera. Therefore, the viewing resolution and field of view of the system can be improved. The acquisition structure and parameters are flexible as well. Multispectral sensing can be done using a CMOS or CCD camera with specific filters. To consider the visible spectrum only, a near-infrared (NIR) cutoff filter needs to be fixed in front of the sensor. To capture the NIR spectrum image, a color spectral cutoff

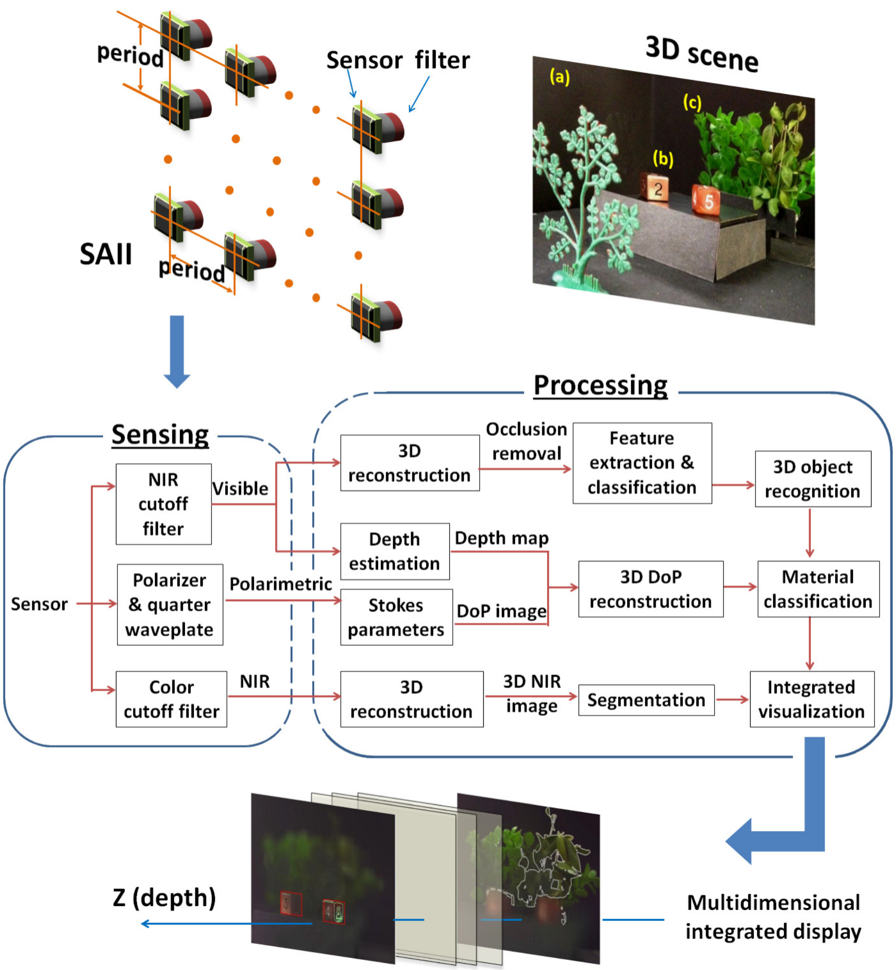

Fig. 24. *Diagram of the proposed MOSIS 2.0.*

filter is used to block the visible band. Likewise, a variety of IR bands may be used.

2) Three-Dimensional Object Recognition with MOSIS 2.0: As discussed in Section III, the depth estimation method shown in [79] can be used to create a depth map related to a particular sensor perspective. By using integral imaging computational reconstruction [73], [75], occlusion in front of the target can be mitigated in the 3-D reconstructed images.

Applying the histogram of oriented gradients (HOG) [109], the reconstructed image is used for 3-D object recognition. The pixel gradient vector in a 3-D image is extracted as

$$\begin{cases} |\nabla f| = \sqrt{(I_x^z)^2 + (I_y^z)^2} \\ \alpha = \arctan\left(I_y^z / I_x^z\right) \end{cases} \tag{23}$$

where I_x^z and I_y^z are the pixel gradients along the x and y directions of the 3-D image reconstructed at depth of z. The oriented gradients vectors and histogram are computed to quantize and compress the feature descriptor.

The extracted HOG features are then fed into a support vector machine (SVM) for classification between true class (target of interest) and false class (others) by finding an optimized separating hyper plane. The hyper plane can be expressed by a discriminant function $g(x) = w^T x + b$. The best classification result should have a maximum margin, so that the hyper plane boundary width can be maximized. The optimization problem is

$$\min \frac{1}{2} ||w||^2,$$
$$s.t.\ y_i(w^T x_i + b) \geq 1, i = 1, 2 \ldots N \tag{24}$$

where w is a coefficients vector, b is a constant, x_i are the training vectors, y_i are the labels of the corresponding data points and N is the number of data points of the training set. For the case that noisy data points exist, slack variables are added to allow for misclassification [110]. To solve the nonlinearly separable problem, the original space can be transformed into a higher dimensional feature space to make the feature set separable. With such nonlinear mapping, the discriminant function becomes $g(\mathrm{x}) = \mathrm{w}^T \varphi(\mathrm{x}) + b$, and (24) is modified as

$$\min_{w,b,\xi} \frac{1}{2} \|\mathrm{w}\|^2 + C \sum_{i=1}^{N} \xi_i ,$$
$$s.t.\ y_i \left(w^T \varphi(x_i) + b \right) \geq 1 - \xi_i \tag{25}$$

where ξ_i are positive slack variables, $\varphi(.)$ is the nonlinear mapping of x_i to a higher dimensional space, and C is a penalty parameter to control over-fitting.

The decision function for a training vector x_j is a sign function

$$sign\left(\sum_{i=1}^{N} y_i a_i K(x_i, x_j) + b \right) \tag{26}$$

where a_i are the Lagrange multipliers found by optimization. $K(x_i, x_j) \equiv \varphi(x_i)^T \varphi(x_j)$ is the kernel function.

The 3-D reconstructed images are used for 3-D object recognition. The occlusion is mitigated in the 3-D reconstructed images, which is why the 3-D object recognition process is effective. If we apply the object recognition process to the 2-D elemental images directly, the features of the objects cannot be extracted because of the occlusion, and the SVM does not work either.

3) Three-dimensional Material Inspection with MOSIS 2.0: We perform material inspection on the 3-D reconstructed image with polarimetric properties. If there is occlusion in front of an object, conventional computational reconstruction introduced in Section II can visualize objects by mitigating the occlusion; however, the pixel overlap and averaging process may degrade the polarization characteristics of the light reflected from the object's surface. To preserve this polarimetric information, we combine the estimated depth information with the degree of polarization property for 3-D polarimetric reconstruction. The modified reconstruction approach is expressed as

$$R_{DoP}(x, y; z) = \frac{1}{K} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[EI^{i,j}_{DoP} \left(x + \frac{c_x^{i,j}}{r_z}, y + \frac{c_y^{i,j}}{r_z} \right) \times \xi^{i,j}(x, y) \right] \tag{27}$$

where M and N are the number of elemental images in the x and y directions, $c_x^{i,j}$, $c_y^{i,j}$ are the positions of the sensor in the x and y directions, respectively, and r_z is the magnification factor for the 3-D reconstruction at depth position (z). $EI^{i,j}_{DoP}(\,.\,)$ indicates the degree of polarization image calculated from the Stokes parameters. $\xi^{i,j}(x, y) \in \{0, 1\}$ is a binary variable defined as

$$\xi^{i,j}(x, y) = \begin{cases} 1, & \text{if } Dep^{i,j}\left(x + \frac{c_x^{i,j}}{r_z}, y + \frac{c_y^{i,j}}{r_z} \right) > THR \\ 0, & \text{otherwise,} \end{cases} \tag{28}$$

where $Dep^{i,j}(.)$ is the depth map obtained for the $(i, j)^{th}$ elemental image from the visible spectrum by using the depth estimation method [(2) and (3)] [113]. A known reconstruction depth threshold (THR) is introduced to separate the depth position between the occlusion and the target for each 3-D point. K is the total number of elemental images used in reconstruction, i.e., $K = \sum_{i,j} \xi^{i,j}(x, y)$.

We assume the occlusion is a convex and Lambertian surface. By setting the depth threshold (THR), the degree of polarization components measured from the light reflected by the occlusion can be removed for the 3-D reconstruction. The reconstructed image provides accurate polarization characteristics corresponding to the object surface. As the majority of object surfaces can be classified based on their electrical properties (such as metal and dielectric), and the polarization property varies for materials between the metallic and nonmetallic surfaces [114], 3-D polarimetric imaging may be helpful for material inspection and classification, industrial inspection and target segmentation, etc. [115].

Besides the polarimetric characteristic, some materials can be identified from their various spectral reflection signatures. By implementing the multispectral integral imaging method, specific materials such as vegetation can be identified due to having a high NIR spectrum and a relatively low visible band reflectance. The k-means clustering algorithm [116], [117] is used in MOSIS 2.0 for target segmentation by minimization of the cluster sum of squared Euclidean distances. The minimization problem is

$$\min_{C} \sum_{i=1}^{k} \sum_{x_j \in C_i} D^2(x_j, \mu_i) \tag{29}$$

where k is the number of clusters (classes), C_i represents the set of data points that belong to the i^{th} cluster, μ_i is the i^{th} cluster centroid, and $D^2(x_j, \mu_i)$ is the squared Euclidean distance between x_j and μ_i. Edge detection algorithms can be further applied to outline the detected objects [118], [119]. With MOSIS 2.0, 3-D object recognition, including material properties inspection, can be performed simultaneously. The multidimensionally integrated visualization of a scene may reveal more information to improve imaging understanding and information extraction.

B. Experimental Results for MOSIS 2.0

This section describes the experiments we have performed for the proposed MOSIS 2.0. A color board level camera (EO-2013BL) [120] was fixed on a translation stage for multidimensional sensing with synthetic aperture integral imaging. The 3-D scene includes: 1) A pair of dice with similar size and color, but different materials placed at 370 mm from the camera; 2) an occlusion set in front of the dice at 280 mm; 3) a background containing camouflage (plastic) foliage, and real (vegetation) foliage at 510 mm. Fig. 24 illustrates the sensing system and the 3-D scene used in the experiments.

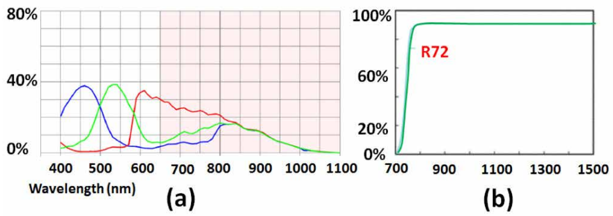

Fig. 25. ***(a) Quantum efficiency of the board level camera [120]. (b) Transmission curve of the NIR-pass filter (Hoya R72) used for NIR imaging.***

Fig. 25(a) shows the quantum efficiency of the CMOS camera. For multispectral imaging, a NIR cutoff filter was first added to the sensor with a cutoff wavelength of 650 nm, so that only the visible spectral range can pass through and be recorded. To capture the NIR band, the cutoff filter was replaced by a NIR-pass filter (Hoya R72) which blocks the visible band. The transmission curve of this filter is shown in Fig. 25(b).

A classical polarization estimation method [44], as discussed in Section IV-B, was implemented to measure the Stokes parameters in our experiments. The camera array for the synthetic aperture integral imaging pickup process is implemented using a moving single camera and includes a total of 36 (6×6) lateral positions, with a camera period of 5 mm. The resolution of each elemental image is 1200 (H) $\times$ 1600 (V) pixels and the camera focal length is 8 mm. Fig. 26(a)–(d) illustrates the multidimensional elemental images corresponding to the visible spectrum, NIR spectrum, measured Stokes parameters of the polarimetric characteristics, and the depth, respectively. The multidimensional elemental images provide different perspectives of the scene.

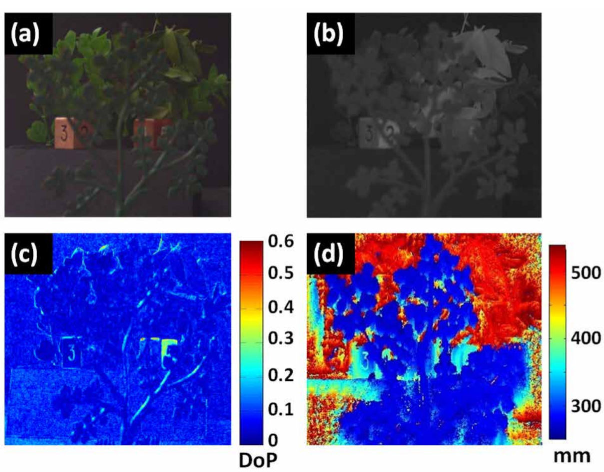

Fig. 26. ***MOSIS 2.0 for 3-D object shape and material inspection experiments in the presence of occlusion. Captured and computed multidimensional elemental images. (a) Visible spectrum and (b) NIR spectrum. (c) Degree of polarization (DoP) computed by the Stokes parameters and (d) depth map by the estimation method.***

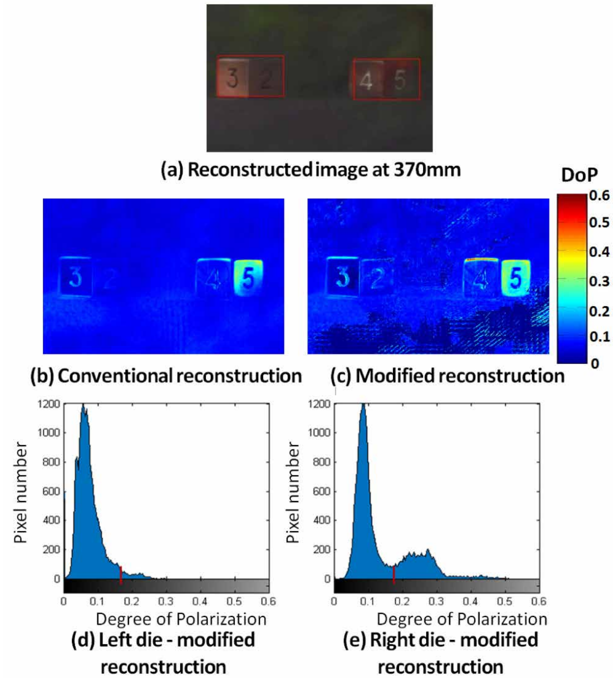

Fig. 27. ***MOSIS 2.0 for 3-D object shape and material inspection experiments in the presence of occlusion. (a) 3-D reconstructed images at the object planes of 370 mm with target recognition. Degree of polarization reconstructed images at 370 mm (b) by the conventional reconstruction, and (c) by the modified reconstruction method. Distribution of degree of polarization for the target surface reflected lights. (d) Left die by the modified reconstruction in (c), and (e) right die by the modified reconstruction in (c).***

Fig. 27(a) is the 3-D reconstructed image at 370 mm, where the dice are in focus. Compared with Fig. 26(a), the occlusion is significantly mitigated, and features related to the object surface can be extracted for object recognition. In the experiments, 26 true class (the surfaces of a die) and 48 false class (trees, other objects, background, etc.) images were used as the training data for the SVM model classification. The dice in the reconstructed image can be recognized corresponding to the highest two estimation probabilities from SVM. The red boxes visualized in Fig. 27(a) indicate the windows corresponding to the recognized targets.

Suppose prior information is given in the sense that the pair of dice is made of metallic and nonmetallic materials. However, it is difficult to identify the material in the visible wavelength range. With the reconstructed degree of polarization image, material inspection can be performed. Fig. 27(b) indicates the direct 3-D reconstructed result of the degree of polarization image at the target depth position (370 mm). The polarimetric characteristic around the right corner of the surface is degraded due to occlusion. By using the modified reconstruction approach [see (27)], the degree of polarization components from the occlusion is removed by combining the degree

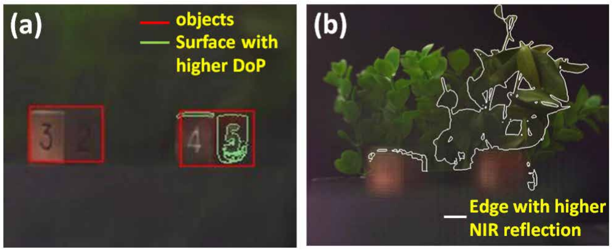

Fig. 28. ***MOSIS 2.0 for 3-D object shape and material inspection experiments in the presence of occlusion. Multidimensionally integrated visualization results. (a) 3-D reconstruction at 370 mm. (b) 3-D reconstruction at 510 mm.***

of polarization elemental images with depth information. In Fig. 27(c), the reconstructed degree of polarization image provides the accurate polarimetric characteristics on the target surface.

The degree of polarization histogram within the areas of the target surface is then extracted. The distributions of the degree of polarization from the reflected lights of the left and right detected die are shown in Fig. 27(d) and (e), respectively. An additional peak is centered at the degree of polarization value around 0.26 in Fig. 27(e). The material discrimination between the targets can be performed by thresholding with degree of polarization (DoP $= 0.18$). Results indicate the reflected light from the right target surface has a higher degree of polarization components. As the dielectric surface partially polarizes incident light upon specular reflection more strongly than the metal surface [121], we can conclude that the right die is the plastic one and the left die has a metal surface.

K-means clustering was performed on the NIR reconstructed 3-D images at 510 mm for the segmentation between vegetation and plastics in the background. The outline of the clusters was extracted by using the sobel edge [122] operator for both polarimetric and multispectral visualization. Multidimensional visualization with 3-D object recognition and material inspection can be integrated for enhancing image understanding and information extraction. In Fig. 28(a), the red boxes highlight the recognized objects at 370 mm, the green outline sketches the surfaces with higher degree of polarization, indicating the presence of a dielectric material surface. Fig. 28(b) is the visualization result at 510 mm. The real foliage is pulled out from the plastic ones using the NIR information.

VIII. DYNAMIC MOSIS IN MICROSCALES FOR MICROSCOPY AND MEDICAL APPLICATIONS

In this section, we present a brief overview of MOSIS for medical applications using dynamic integral imaging systems with time multiplexing implemented with liquid crystal devices.

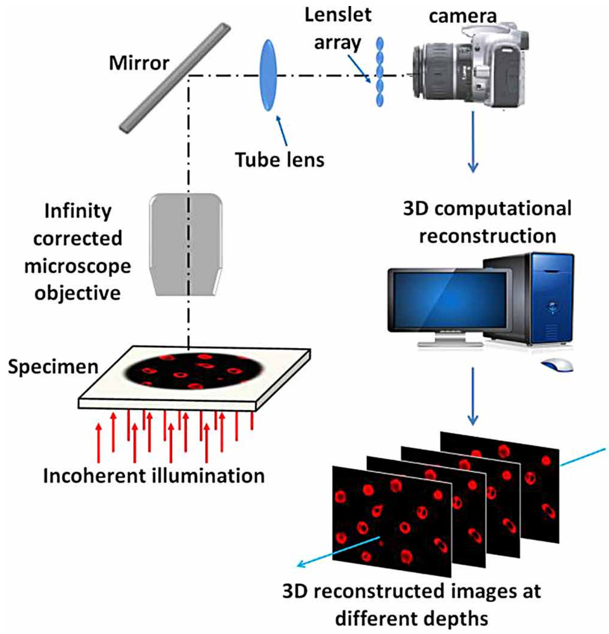

Fig. 29. ***3-D integral imaging microscopy for cell identification [28]. 3-D sensing process can also be performed using synthetic aperture integral imaging technique.***

Three-dimensional integral imaging has found applications in 3-D endoscopy and can be used for cell identification and classification. In 3-D endoscopy, a liquid crystal (LC) lens array or a LC lens is used to capture objects close to the imaging sensor. Furthermore, by using an electrically moveable LC lens array, a time multiplex technique called the moving array lenslet technique can be used to improve the 3-D image resolution.

A. Three-Dimensional Integral Imaging Microscopy for Cell Identification

Integral imaging technology can be employed for 3-D microscopy [26], [28]–[31]. The identification of biological microorganisms with 3-D integral imaging has been proposed in [28]. The schematic setup of integral imaging microscopy for cell identification [28] is shown in Fig. 29. Incoherent light passes through a 3-D specimen, and it is subsequently magnified by an infinity corrected microscope objective to form a real image. For the 3-D sensing process, a 2-D sensor records the object from various perspectives using the synthetic aperture integral imaging technique or a lenslet array. Three-dimensional integral imaging reconstruction can be performed by the computational reconstruction method. The 3-D reconstructed images contain depth and profile information of the micro object, which can be used for identification and classification by statistical pattern recognition techniques [28].

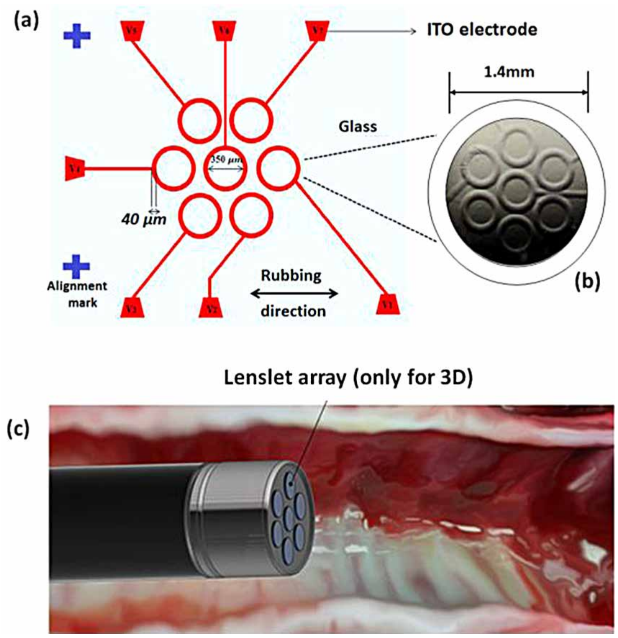

Fig. 30. ***(a) Diagram of the Indium tin oxide (ITO) electrode. (b) Manufactured hexagonal liquid crystal micro lens array and the magnified image of the hexagonal electrodes. (c) Example of a 3-D endoscope where the electrode liquid crystal lens in its hexagonal convex form is located just in front of the endoscope [32].***

B. Three-Dimensional Integral Imaging with Endoscopy

1) Three-Dimensional Endoscopy Using a Hexagonal Liquid Crystal Lens Array: The conventional stereo 3-D endoscopes consist of double image sensors with double lenses. This configuration may lead to a relatively large physical size (about 10 mm) and has the limitations of stereo endoscopic systems. A liquid crystal (LC) lens array has been developed for 3-D sensing with a single sensor based on the integral imaging technique [32], [123]. It can be applied for 3-D mode by simply mounting the LC lens array in front of the conventional 2-D endoscope, which has the same diameter as the 2-D endoscope lens (less than 1.4 mm). Details of the fabrication process of the LC lens array are discussed in [32]. Fig. 30(a) shows the pattern of the LC lens electrode with a hexagonal arrangement. An independent voltage level can be applied to each lens in the array. The hexagonal convex-ring electrodes and its magnified image are illustrated in Fig. 30(b). Fig. 30(c) depicts a 3-D endoscope with the embedded hexagonal convex-ring electrode LC lens which is placed in front of the endoscope [32]. This electrode lens produces a parabolic-type electric field distribution, and the focal length can be shortened to a value less than 2.5 cm. Thus, the 3-D endoscope can be used to acquire objects close to the LC lens for medical applications.

Fig. 31 illustrates experimental results of biological samples captured by the 3-D integral imaging endoscope. Fig. 31(a) and (b) shows the images with 2-D and 3-D topography. The images without and with focusing of the LC lenses are shown in Fig. 31(c) and (d), respectively.

2) Two-Dimensional/Three-Dimensional Adjustable Endoscopy and Axially Distributed Sensing Using Electrically Controlled Liquid Crystal Lens: A multi-functional liquid-crystal lens (MFLC-lens) is demonstrated for 2-D and 3-D switchable and focus tunable function without any mechanical

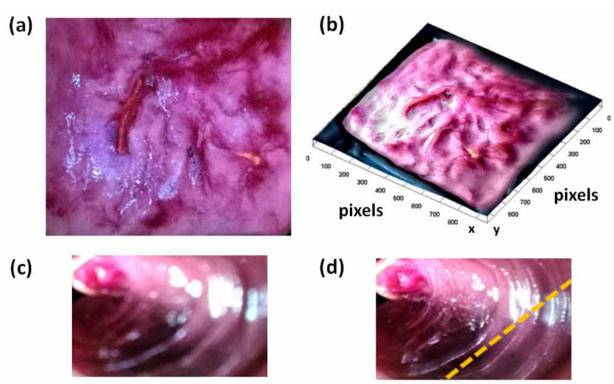

Fig. 31. ***Biological samples captured by the 3-D endoscope. (a) 2-D and (b) 3-D surface topography, and the (c) nonfocusing and (d) focusing image of biological sample captured by using the LC lenses.***

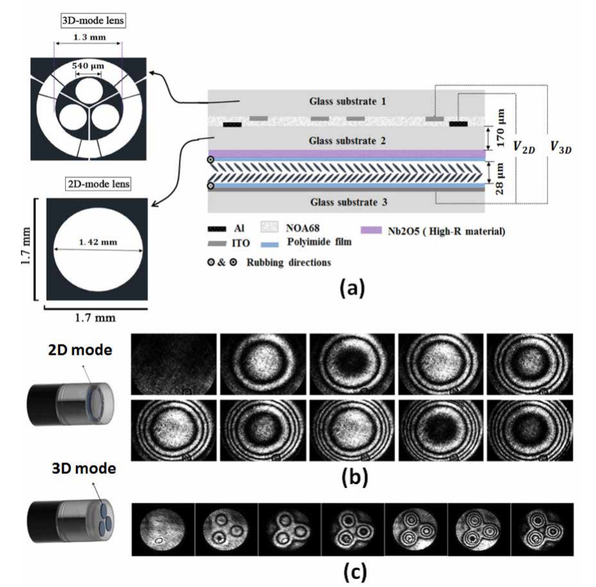

Fig. 32. ***(a) Top view of the electrode patterns and cross section of the multi-functional liquid-crystal lens (MFLC-LC) lens cell for a 2-D/3-D tunable endoscope. The experiment results of the interference pattern using (b) 2-D and (c) 3-D mode [123].***

movement. To achieve multiple focal length lens functions, a LC lens structure with dual-layer electrode coated by a high resistive transparent film was developed, as shown in Fig. 32 [123]. The diameter of the proposed MFLC-lens is only 1.42 mm with tunable focal length from infinity to 80 mm in 2-D mode, and to 20 mm in 3-D mode. It can be easily applied to micro-imaging systems, such as an endoscopic system, and objects in close proximity sensing for both 2-D and 3-D image capturing.

The axially distributed sensing [66], [124] configuration is also very practical for 3-D endoscopy, as long as multiple lenses are not used. In the axially distributed sensing method, a single sensor is translated along its optical axis or the objects are moved parallel to the optical axis of a single sensor, and the captured images have various perspectives and magnifications of the objects. Three-dimensional information can be computationally reconstructed based on ray back-projection. By employing a multi-focal lengths LC lens, we can apply axially distributed sensing on the 3-D endoscopy by changing the voltages.

C. Dynamic Integral Imaging Display with Electrically Moving Array Lenslet Technique Using Liquid Crystal Lens

Micro integral imaging can be combined with optical see-through head-mounted display (OST-HMD) for an augmented reality 3-D display system [17], [35], [36], and [105]. Micro integral imaging creates a 3-D image source for the head-mounted display (HMD) viewing optics. This configuration can reduce the accommodation-convergence mismatch and visual fatigue issue common in traditional augmented reality HMD. Thus, the proposed system can be used for biomedical surgery. We note that by using the integral imaging technology, the display's resolution will decrease.

To improve the resolution of the integral imaging display, an electrically movable LC lens array was developed for dynamic integral imaging system with the moving array lenslet technique (MALT) [52], [125]. Depending on the advantage of the two separated lenticular LC lens arrays with the multi-electrode structure, as shown in Fig. 33(a), the subpitch movements of the lenslets along a vertical, horizontal, or diagonal direction can be realized and electrically controlled via moving the driving voltage to the next electrodes, as shown in Fig. 33(b). The image when MALT was not used with the LC lens [Fig. 33(c)] is compared with the reconstructed image when MALT was used [Fig. 33(d)]. The results demonstrate that using the LC lens with MALT results in a smoother and continuous

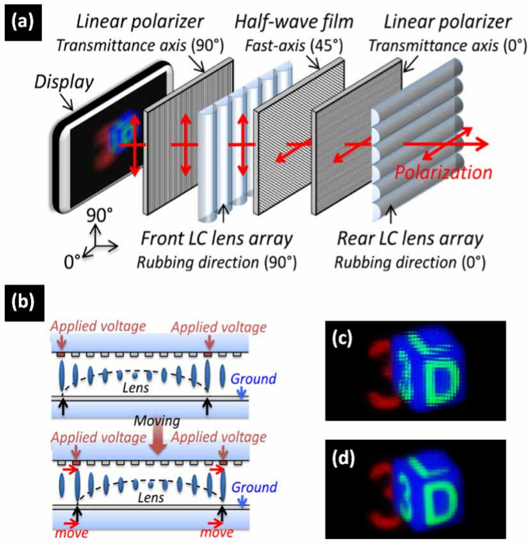

Fig. 33. ***(a) System diagram of the electrically movable LC lens array for dynamic integral imaging display, and (b) its driving method. Reconstructed images of the 3-D scene. (c) Without MALT, and (d) with the proposed LC lens MALT [16].***

image. Moreover, multifacet effect is eliminated resulting in improved resolution and image quality [16].

IX. CONCLUSION

In this paper, we have presented a literature overview of the passive integral-imaging-based MOSIS with applications from macroscales to microscales of objects. In addition, we have presented new results on using MOSIS 2.0 for 3-D object shape, material inspection and recognition such that objects with similar shapes but different materials can be discriminated. Multidimensional information, which may include time-space multiplexing, visible and infrared spectrum, polarization of light, time domain variations, and photon flux changes, etc., can be measured, extracted, and visualized. Image processing algorithms are needed for postprocessing, image fusion and multidimensional visualization. Human activity and gesture recognition in 3-D is discussed along with dynamic integral imaging systems using tunable liquid crystal lenses. Moreover, applications to cell identification and 3-D endoscopy have been presented. Obtaining multidimensional information from the images and scenes increases the information content to significantly improve information extracted from the scene and objects. ■

REFERENCES

[1] S. Benton and M. Bove, *Holographic Imaging*. Hoboken, NJ, USA: Wiley, 2008.

[2] C. Slinger, C. Cameron, and M. Stanley, "Computer-generated holography as a generic display technology," *Computer*, vol. 38, no. 8, pp. 46–53, Aug. 2005.

[3] C. Wheatstone, "Contributions to the physiology of vision—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philos. Trans.*

Roy. Soc. London A, Math. Phys. Sci., vol. 128, pp. 371–394, Jan. 1838.

[4] H. Kang *et al.*, "3.1: A novel polarizer glassestype 3D displays with a patterned retarder," in *SID Symp. Dig. Tech*, 2010, vol. 41. no. 1, pp. 1–4.

[5] T. Inoue and H. Ohzu, "Accommodative responses to stereoscopic three-dimensional display," *Appl. Opt.*, vol. 36, no. 19, pp. 4509–4515, Jul. 1997.

[6] K. Mashitani and G. Hamagishi, "Autostereoscopic video display with a parallax barrier having oblique apertures," U.S. Patent 7317494 B2, Jan. 8, 2008.

[7] G. Lippmann, "Épreuves réversibles donnant la sensation du relief," *J. Phys. Theoretical Appl.*, vol. 7, no. 1, pp. 821–825, 1908.

[8] B. Javidi, O. Fumio, and J.-Y. Son, *Three-Dimensional Imaging, Visualization, and Display*. New York, NY, USA: Springer, 2009.

[9] D. LeMaster, B. Karch, and B. Javidi, "Mid-wave infrared 3D integral imaging at long range," *J. Display Technol.*, vol. 9, no. 7, pp. 545–551, Jul. 2013.

[10] A. Bhandari and R. Raskar, "Signal processing for time-of-flight imaging sensors: An introduction to inverse problems in computational 3-D imaging," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 45–58, Sep. 2016.

[11] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1. Jun. 2003, pp. I-195–I-202.

[12] J. Geng, "Structured-light 3D surface imaging: A tutorial," *Adv. Opt. Photon.*, vol. 3, no. 2, pp. 128–160, 2011.

[13] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Dept. Comput. Sci., Stanford University, Stanford, CA, USA, Tech. Rep. CSTR 2005-02, Apr. 2005.

[14] J. Arai *et al.*, "Integral three-dimensional television using a 33-megapixel imaging system," *J. Display Technol.*, vol. 6, no. 10, pp. 422–430, Oct. 2010.

[15] G. Wetzstein, L. Douglas, H. Matthew, and R. Ramesh, "Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting," *ACM Trans. Graph.*, vol. 31, no. 4, Jul. 2012, Art. no. 80.

[16] T.-H. Jen, X. Shen, G. Yao, Y.-P. Huang, H.-P. D. Shieh, and B. Javidi, "Dynamic integral imaging display with electrically moving array lenslet technique using liquid crystal lens," *Opt. Exp.*, vol. 23, no. 14, pp. 18415–18421, Jul. 2015.

[17] X. Shen, Y. J. Wang, H. S. Chen, X. Xiao, Y. H. Lin, and B. Javidi, "Extended depth-of-focus 3D micro integral imaging display using a bifocal liquid Crystal lens," *Opt. Lett.*, vol. 40, no. 4, pp. 538–541, Feb. 2015.

[18] X. Xiao *et al.*, "Three-dimensional holographic display using dense ray sampling and integral imaging capture," *J. Display Technol.*, vol. 10, no. 8, pp. 688–694, Aug. 2014.

[19] M. Cho and B. Javidi, "Three-dimensional visualization of objects in turbid water using integral imaging," *J. Display Technol.*, vol. 6, no. 10, pp. 544–547, Oct. 2010.

[20] Y. Zhao, X. Xiao, M. Cho, and B. Javidi, "Tracking of multiple objects in unknown background using Bayesian estimation in 3D space," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 28, no. 9, pp. 1935–1940, Sep. 2011.

[21] S.-H. Hong and B. Javidi, "Distortion-tolerant 3D recognition of occluded objects using computational integral imaging," *Opt. Exp.*, vol. 14, no. 25, pp. 12085–12095, Dec. 2006.

[22] Y. Frauel and B. Javidi, "Digital three-dimensional image correlation by use of computer-reconstructed integral imaging," *Appl. Opt.*, vol. 41, no. 26, pp. 5488–5496, Sep. 2002.

[23] I. Moon and B. Javidi, "Three-dimensional recognition of photon-starved events using computational integral imaging and statistical sampling," *Opt. Lett.*, vol. 34, no. 6, pp. 731–733, Mar. 2009.

[24] P. Latorre-Carmona, B. Javidi, and D. A. LeMaster, "Three-dimensional visualization of long range scenes by photon counting mid-wave infrared integral imaging," *J. Display Technol.*, vol. 11, no. 11, 908-912, Nov. 2015.

[25] J. W. Beletic *et al.*, "Teledyne imaging sensors: Infrared imaging technologies for astronomy and civil space," *Proc. SPIE*, vol. 7021, pp. 1–14, Jul. 2008.

[26] J.-S. Jang and B. Javidi, "Three-dimensional integral imaging of micro-objects," *Opt. Lett.*, vol. 29, no. 11, pp. 1230–1232, Jun. 2004.

[27] J.-S. Jang and B. Javidi, "Very large-scale integral imaging (VLSII) for 3-D display," *Opt. Eng.*, vol. 44, no. 1, p. 014001, 2005.

[28] B. Javidi, I. Moon, and S. Yeom, "Three-dimensional identification of biological microorganism using integral imaging," *Opt. Exp.*, vol. 14, no. 25, pp. 12096–12108, Dec. 2006.

[29] M. Levoy, Z. Zhang, and I. McDowall, "Recording and controlling the 4D light field in a microscope using microlens arrays," *J. Microscopy*, vol. 235, no. 2, pp. 144–162, Aug. 2009.

[30] A. Llavador, E. Sánchez-Ortiga, J. C. Barreiro, G. Saavedra, and M. Martínez-Corral, "Resolution enhancement in integral microscopy by physical interpolation," *Biomed. Opt. Exp.*, vol. 6, no. 8, pp. 2854–2863, 2015.

[31] A. Llavador, J. Sola-Picabea, G. Saavedra, B. Javidi, and M. Martínez-Corral, "Resolution improvements in integral microscopy with Fourier plane recording," *Opt. Exp.*, vol. 24, no. 18, pp. 20792–20798, 2016.

[32] A. Hassanfiroozi, Y.-P. Huang, B. Javidi, and H.-P. D. Shieh, "Hexagonal liquid crystal lens array for 3D endoscopy," *Opt. Exp.*, vol. 23, no. 2, pp. 971–981, 2015.

[33] Y. J. Wang, X. Shen, Y. H. Lin, and B. Javidi, "Extended depth-of-field 3D endoscopy with synthetic aperture integral imaging using an electrically tunable focal-length liquid-crystal lens," *Opt. Lett.*, vol. 40, no. 15, pp. 3564–3567, Aug. 2015.

[34] X. Shen, M. M. Corral, and B. Javidi, "Head tracking three-dimensional integral imaging display using smart pseudoscopic-to-orthoscopic conversion," *J. Display Technol.*, vol. 12, no. 6, pp. 542–548, Jun. 2016.

[35] H. Hua and B. Javidi, "A 3D integral imaging optical see-through head-mounted display," *Opt. Exp.*, vol. 22, no. 11, pp. 13484–13491, Jun. 2014.

[36] J. Wang, X. Xiao, H. Hua, and B. Javidi, "Augmented reality 3D displays with micro integral imaging," *J. Display Technol.*, vol. 11, no. 11, pp. 889–893, Nov. 2015.

[37] A. Markman, X. Shen, H. Hua, and B. Javidi, "Augmented reality three-dimensional object visualization and recognition with axially distributed sensing," *Opt. lett.*, vol. 41, no. 2, pp. 297–300, Jan. 2016.

[38] O. Bimber and R. Raskar, *Spatial Augmented Reality: Merging Real and Virtual Worlds*. Boca Raton, FL, USA: CRC Aug. 2005.

[39] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*. San Francisco, CA, USA: Academic, Aug. 2006.

[40] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1152, Jun. 1985.

[41] G. Moser and S. B. Serpico, "Automatic parameter optimization for support vector regression for land and sea surface temperature estimation from remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 909–921, Mar. 2009.

[42] M. I. Mishchenko, Y. S. Yatskiv, V. K. Rosenbush, and G. Videen, *Polarimetric Detection, Characterization and Remote Sensing*. New York, NY, USA: Springer, May 2011.

[43] I. Kuzmina *et al.*, "Towards noncontact skin melanoma selection by multispectral imaging analysis," *J. Biomed. Opt.*, vol. 16, no. 6, p. 0605021, 2011.

[44] E. Collett, *Polarized Light: Fundamentals and Applications*. New York, NY, USA: Marcel Dekker, Dec. 1992.

[45] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Elsevier, 1980.

[46] O. Matoba and B. Javidi, "Three-dimensional polarimetric integral imaging," *Opt. Lett.*, vol. 29, vol. 20, pp. 2375–2377, Oct. 2004.

[47] H. N. Fructuoso, M. Martinez-Corral, G. S. Tortosa, A. P. Marti, and B. Javidi, "Photoelastic analysis of partially occluded objects with an integral-imaging polariscope," *J. Display Technol.*, vol. 10, no. 4, pp. 255–262, Apr. 2014.

[48] H.-S. Chen and C. R. N. Rao, "Polarization of light on reflection by some natural surfaces," *J. Phys. D, Appl. Phys.*, vol. 1, no. 9, p. 1191, Sep. 1968.

[49] H. Chen and L. B. Wolff, "Polarization phase-based method for material classification in computer vision," *Int. J. Comput. Vis.*, vol. 28, pp. 73–83, Jun. 1998.

[50] M. I. Mishchenko, Y. S. Yatskiv, V. K. Rosenbush, and G. Videen, *Polarimetric Detection, Characterization and Remote Sensing*. New York, NY, USA Springer, May 2011.

[51] B. Javidi, S.-H. Hong, and O. Matoba, "Multidimensional optical sensor and imaging system," *Appl. Opt.*, vol. 45, no. 13, pp. 2986–2994, May 2006.

[52] J.-S. Jang and B. Javidi, "Improved viewing resolution of three-dimensional integral imaging by use of nonstationary micro-optics," *Opt. Lett.*, vol. 27, no. 5, pp. 324–326, Mar. 2002.

[53] J. S. Walker, *A Primer on Wavelets and Their Scientific Applications*. Boca Raton, FL, USA: CRC, 2008.

[54] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[55] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, no. 7, pp. 909–996, 1988.

[56] H. E. Ives, "Optical properties of a Lippmann lenticulated sheet," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 21, no. 3, pp. 171–176, 1931.

[57] T. Okoshi, "Three-dimensional displays," *Proc. IEEE*, vol. 68, no. 5, pp. 548–564, May 1980.

[58] N. Davies, M. McCormick, and L. Yang, "Three-dimensional imaging systems: A new development," *Appl. Opt.*, vol. 27, no. 21, pp. 4520–4528, 1988.

[59] A. Stern and B. Javidi, "Three-dimensional image sensing, visualization, and processing using integral imaging," *Proc. IEEE*, vol. 94, no. 3, pp. 591–607, Mar. 2006.

[60] R. Martinez-Cuenca, G. Saavedra, M. Martinez-Corral, and B. Javidi, "Progress in 3-D multiperspective display by integral imaging," *Proc. IEEE*, vol. 97, no. 6, pp. 1067–1077, Jun. 2009.

[61] J.-H. Park, K. Hong, and B. Lee, "Recent progress in three-dimensional information processing based on integral imaging," *Appl. Opt.*, vol. 48, no. 34, pp. H77–H94, Dec. 2009.

[62] X. Xiao, B. Javidi, M. Martinez-Corral, and A. Stern, "Advances in three-dimensional integral imaging: Sensing, display, and applications," *Appl. Opt.*, vol. 52, no. 4, pp. 546–560, Feb. 2013.

[63] P. Latorre-Carmona, F. Pla, A. Stern, I. Moon, and B. Javidi, "Three-dimensional imaging with multiple degrees of freedom using data fusion," *Proc. IEEE*, vol. 103, no. 9, pp. 1654–1671, Sep. 2015.

[64] F. Okano, J. Arai, K. Mitani, and M. Okui, "Real-time integral imaging based on extremely high resolution video system," *Proc. IEEE*, vol. 94, no. 3, pp. 490–501, Mar. 2006.

[65] J.-S. Jang and B. Javidi, "Three-dimensional synthetic aperture integral imaging," *Opt. Lett.*, vol. 27, no. 13, pp. 1144–1146, Jul. 2002.

[66] R. Schulein, M. DaneshPanah, and B. Javidi, "3D imaging with axially distributed sensing," *Opt. Lett.*, vol. 34, no. 13, pp. 2012–2014, Jul. 2009.

[67] M. DaneshPanah, B. Javidi, and E. A. Watson, "Three dimensional imaging with randomly distributed sensors," *Opt. Exp.*, vol. 16, no. 9, pp. 6368–6377, Apr. 2008.

[68] F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography," *Appl. Opt.*, vol. 36, no. 7, pp. 1598–1603, 1997.

[69] H. Navarro, R. Martínez-Cuenca, G. Saavedra, M. Martínez-Corral, and B. Javidi, "3D integral imaging display by smart pseudoscopic-to-orthoscopic conversion (SPOC)," *Opt. Exp.*, vol. 18, no. 25, pp. 25573–25583, 2010.

[70] M. Martínez-Corral, A. Dorado, H. Navarro, G. Saavedra, and B. Javidi, "Three-dimensional display by smart pseudoscopic-to-orthoscopic conversion with tunable focus," *Appl. Opt.*, vol. 53, no. 22, pp. E19–E25, Aug. 2014.

[71] X. Shen, X. Xiao, M. Martínez-Corral, and B. Javidi, "Format matching using multiple-planes pseudoscopic-to-orthoscopic conversion for 3D integral imaging display," *Proc. SPIE*, vol. 9495, p. 94950W, May 2015.

[72] X. Xiao, X. Shen, M. Martinez-Corral, and B. Javidi, "Multiple-planes pseudoscopic-to-orthoscopic conversion for 3D integral imaging display," *J. Display Technol.*, vol. 11, no. 11, pp. 921–926, Nov. 2015.

[73] H. Arimoto and B. Javidi, "Integral three-dimensional imaging with digital reconstruction," *Opt. Lett.*, vol. 26, no. 3, pp. 157–159, Feb. 2001.

[74] S.-H. Hong, J.-S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Exp.*, vol. 12, no. 3, pp. 483–491, Feb. 2004.

[75] V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. B. Kang, "Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2006, pp. 2331–2338.

[76] M. Daneshpanah and B. Javidi, "Three-dimensional imaging with detector arrays on arbitrarily shaped surfaces," *Opt. Lett.*, vol. 36, no. 5, pp. 600–602, Mar. 2011.

[77] X. Xiao, M. DaneshPanah, M. Cho, and B. Javidi, "3D integral imaging using sparse sensors with unknown positions," *J. Display Technol.*, vol. 6, no. 12, pp. 614–619, Dec. 2010.

[78] J. Wang, X. Xiao, and B. Javidi, "Three-dimensional integral imaging with flexible sensing," *Opt. Lett.*, vol. 39, no. 24, pp. 6855–6858, Dec. 2014.

[79] M. DaneshPanah and B. Javidi, "Profilometry and optical slicing by passive three-dimensional imaging," *Opt. Lett.*, vol. 34, no. 7, pp. 1105–1107, Apr. 2009.

[80] A. Martínez-Usó, P. Latorre-Carmona, F. Pla, J. M. Sotoca, and B. Javidi, "Depth estimation in integral imaging based on a maximum voting strategy," *J. Display Technol.*, vol. 12, no. 12, pp. 1715–1723, Dec. 2016.

[81] G. Vogiatzis, C. H. Esteban, P. H. S. Torr, and R. Cipolla, "Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2241–2246, Dec. 2007.

[82] A. Fitzgibbon, Y. Wexler, and A. Zisserman, "Image-based rendering using image-based priors," *Int. J. Comput. Vis.*, vol. 63, no. 2, pp. 141–151, 2005.

[83] X. Xiao, B. Javidi, G. Saavedra, M. Eismann, and M. Martinez-Corral, "Three-dimensional polarimetric computational integral imaging," *Opt. Exp.*, vol. 20, no. 14, pp. 15481–15488, Jul. 2012.

[84] A. Carnicer and B. Javidi, "Polarimetric 3D integral imaging in photon-starved conditions," *Opt. Exp.*, vol. 23, no. 5, pp. 6408–6417, Mar. 2015.

[85] J. A. Davis, D. E. McNamara, D. M. Cottrell, and T. Sonehara, "Two-dimensional polarization encoding with a phase-only liquid-crystal spatial light modulator," *Appl. Opt.*, vol. 39, no. 10, pp. 1549–1554, Apr. 2000.

[86] E. Wolf, *Introduction to the Theory of Coherence and Polarization of Light*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2007.

[87] B. Schaefer, E. Collett, R. Smyth, D. Barret, and B. Fraher, "Measuring the Stokes polarization parameters," *Amer. J. Phys.*, vol. 75, no. 2, pp. 163–168, Feb. 2007.

[88] D. Aloni, A. Stern, and B. Javidi, "Three-dimensional photon counting integral imaging reconstruction using penalized maximum likelihood expectation maximization," *Opt. Exp.*, vol. 19, no. 20, pp. 19681–19687, Sep. 2011.

[89] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[90] T. F. Chan, G. H. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 20, no. 6, pp. 1964–1977, 1999.

[91] J. W. Goodman, *Statistical Optics*. Hoboken, NJ, USA: Wiley, 1985.

[92] B. Tavakoli, B. Javidi, and E. Watson, "Three dimensional visualization by photon counting computational integral imaging," *Opt. Exp.*, vol. 16, no. 7, pp. 4426–4436, 2008.

[93] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[94] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1, pp. 89–97, 2004.

[95] S. van der Walt *et al.*, "Scikit-image: Image processing in Python," *PeerJ*, vol. 2, p. e453, Jan. 2014.

[96] E. L. Dereniak and G. L. Boreman, *Infrared Detectors and Systems*. New York, NY, USA: Wiley, 1996.

[97] N. Kingsbury, "The dual-tree complex wavelet transform: A new efficient tool for image restoration and enhancement," in *Proc. 9th Eur. Signal Process. Conf.*, 1998, pp. 319–322.

[98] S. Bonettini and V. Ruggiero, "An alternating extragradient method for total variation-based image restoration from Poisson data," *Inverse Problems*, vol. 27, no. 9, p. 095001, 2011.

[99] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[100] A. Daniels, *Field Guide to Infrared Systems*. Washington, DC, USA: SPIE Press, 2007.

[101] R. Gade and T. B. Moeslund, "Thermal cameras and applications: A survey," *Mach. Vision Appl.*, vol. 25, no. 1, pp. 245–262, 2014.

[102] S. Komatsu, A. Markman, and B. Javidi, "Long wave IR three-dimensional integral imaging and object recognition," *Appl. Opt.*, vol. 56, 2017.

[103] S.-H. Hong and B. Javidi, "Three-dimensional visualization of partially occluded objects using integral imaging," *J. Display Technol.*, vol. 1, no. 2, p. 354, Dec. 2005.

[104] V. J. Traver, P. Latorre-Carmona, E. Salvador-Balaguer, F. Pla, and B. Javidi, "Human gesture recognition using three-dimensional integral imaging," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 31, no. 10, pp. 2312–2320, 2014.

[105] J.-S. Jang, F. Jin, and B. Javidi, "Three-dimensional integral imaging with large depth of focus by use of real and virtual image fields," *Opt. Lett.*, vol. 28, no. 16, pp. 1421–1423, Aug. 2003.

[106] N. Ayache and C. Hansen, "Rectification of images for binocular and trinocular vision," in *Proc. 9th Int. Conf. Pattern Recognit.*, 1998, p. 1.

[107] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[108] H. Wang, M. M. Ullah, A. Kläser, I. Laptev, and C. Schmid, "Evaluation of local spatio-temporal features for action recognition," in *Proc. Brit. Mach. Vision Conf. (BMVA)*, 2009, pp. 124.1–124.11.

[109] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit. (CVPR)*, vol. 1. Jun. 2005, pp. 886–893.

[110] B. Schölkopf and A. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2002.

[111] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[112] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *J. Mach. Learn. Res.*, vol. 5, pp. 975–1005, Dec. 2004.

[113] X. Xiao, M. Daneshpanah, and B. Javidi, "Occlusion removal using depth mapping in three-dimensional integral imaging," *J. Display Technol.*, vol. 8, no. 8, pp. 483–490, Aug. 2012.

[114] L. B. Wolff, "Polarization-based material classification from specular reflection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 11, pp. 1059–1071, Nov. 1990.

[115] F. Goudail, P. Terrier, Y. Takakura, L. Bigué, F. Galland, and V. DeVlaminck, "Target detection with a liquid-crystal-based passive Stokes polarimeter," *Appl. Opt.*, vol. 43, no. 2, pp. 274–282, 2004.

[116] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, vol. 1, no. 14, pp. 281–297, Jun. 1967.

[117] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[118] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[119] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[120] [Online]. Available: http://www.edmundoptics.com/cameras/board-level-cameras/eo-usb-2-0-board-level-cameras/63871/

[121] S. Tominaga and T. Yamamoto, "Metal-dielectric object classification by polarization degree map," in *Proc. 19th Int. Conf. Pattern Recognit. (ICPR)*, Tampa, FL, USA: Dec. 2008, pp. 1–4.

[122] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, MA, USA: Addison-Wesley, 1992, pp. 414–428.

[123] A. Hassanfiroozi, Y.-P. Huang, B. Javidi, and H.-P. D. Shieh, "Dual layer electrode liquid crystal lens for 2D/3D tunable endoscopy imaging system," *Opt. Exp.*, vol. 24, no. 8, pp. 8527–8538, Apr. 2016.

[124] C.-W. Chen, M. Cho, Y.-P. Huang, and B. Javidi, "Three-dimensional imaging with axially distributed sensing using electronically controlled liquid crystal lens," *Opt. Lett.*, vol. 37, no. 19, pp. 4125–4127, 2012.

[125] S.-H. Hong and B. Javidi, "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," *Opt. Exp.*, vol. 12, no. 19, pp. 4579–4588, 2004.

ABOUT THE AUTHORS

Bahram Javidi (Fellow, IEEE) received the B.S. degree from George Washington University, and the M.S. and Ph.D. degrees from the Pennsylvania State University, all in electrical engineering.

He is the Board of Trustees Distinguished Professor at the University of Connecticut, Storrs, CT, USA. He has over 1000 publications, including nearly 450 peer reviewed journal articles, over 450 conference proceedings, including over 120 Plenary Addresses, Keynote Addresses, and invited conference papers. His papers have been cited 33 000 times according to the Google Scholar Citations (*h-index* = 85, *i10-index* = 537).

Prof. Javidi received the Quantum Electronics and Optics Prize for Applied Aspects by the European Physical Society in June 2015. He has been named Fellow of several scientific societies, including IEEE, OSA, and SPIE. In 2010, he was the recipient of The George Washington University's Distinguished Alumni Scholar Award, the university's highest honor for its alumni in all disciplines. In 2008, he received a Fellow award by the John Simon Guggenheim Foundation. He received the 2008 IEEE Donald G. Fink Prized Paper Award among all (over 150) IEEE Transactions, Journals, and Magazines. In 2007, The Alexander von Humboldt Foundation awarded Prof. Javidi with the Humboldt Prize for outstanding U.S. scientists. He received the Technology Achievement Award from the SPIE in 2008. In 2005, he received the Dennis Gabor Award in Diffractive Wave Technologies from SPIE. He was the recipient of the IEEE Photonics Distinguished Lecturer Award twice in 2003 to 2004 and 2004 to 2005. He was awarded the IEEE Best Journal Paper Award from IEEE Transactions on Vehicular Technology twice in 2002 and 2005. Early in his career, the National Science Foundation named him a Presidential Young Investigator and he received The Engineering Foundation and the IEEE Faculty Initiation Award. He was selected in 2003 as one of the nation's top 160 engineers between the ages of 30 and 45 by the National Academy of Engineering (NAE) to be an invited speaker at The Frontiers of Engineering Conference which was co-sponsored by The Alexander von Humboldt Foundation. He is an alumnus of the Frontiers of Engineering of The National Academy of Engineering since 2003. He has served on the Editorial Board of the Proceedings of the IEEE Journal (ranked #1 among all electrical engineering journals), the advisory board of the IEEE Photonics Journal, and he was on the Founding Board of Editors of IEEE/OSA Journal of Display Technology. He is a coauthor on nine Best Paper Awards.

Xin Shen received the B.S. degree in optical information science and technology from Xidian University, Xi'an, China, in 2010, and a dual M.S. degree, one M.S. degree in electronics and communication engineering from Xidian University, in 2013, and one M.S. degree of science in engineering from Doshisha University, Kyoto, Japan, in 2013. He is currently working toward the Ph.D. degree in electrical and computer engineering at the University of Connecticut, Storrs, CT, USA.

His research interests include optical sensing, imaging and 3D visualization.

Mr. Shen is the Vice President of the student chapter of the Optical Society of America (OSA) at the University of Connecticut. He is a member of SPIE and OSA and Recipient of the 2016 summer Graduate Predoctoral Fellowship, 2012 Harris Science School Foundation.

Adam S. Markman received the B.Sc. degree in biomedical engineering and M.Sc. degree in electrical and computer engineering from the University of Connecticut, Storrs, CT, USA, in 2011 and 2013, respectively. He is currently working towards the Ph.D. degree in electrical and computer engineering at the same university.

From 2014 to 2015, he worked as a Systems Engineer in the sonar field at Alion Science and Technology. He currently has ten journal publications and four conference presentations and posters. His research interests include 3-D imaging, optical security, low-light imaging, pattern recognition, statistical optics, and image processing.

Mr. Markman is currently Vice President of the student chapter of SPIE at the University of Connecticut and a student member of OSA. He is the Recipient of the 2017 GE Graduate Fellowship, 2017 SPIE Officer Travel Grant, 2016 to 2017 FEI Graduate Fellowship, 2016 Summer Predoctoral Fellowship, 2016 Doctoral Student Travel Award, and 2016 Sikorsky Fellowship.

Pedro Latorre-Carmona received the B.S. degree in physics from the University of Valencia, Spain, in 1999, and the Ph.D. degree in computer science from the Polytechnical University of Valencia, Spain, in 2005.

He is a Postdoctoral Researcher at the Departamento de Lenguajes y Sistemas Informaticos of the Universitat Jaume I, Castellon, Spain. His current research interests are 3-D image analysis, feature selection and extraction, pattern recognition, multispectral (including remote sensing) image processing, colorimetry and vision physics.

Adolfo Martínez-Uso received the B.Sc. and Ph.D. degrees in computer science from Universitat Jaume I, Castellón de la Plana, Spain.

He is currently with the Institute of New Imaging Technologies, Universitat Jaume I, as a Post-Doctoral Researcher in the computer vision group. His research interests are image processing, pattern recognition and data mining, particularly, 3-D image analysis, hyperspectral imaging, feature selection and image segmentation.

José Martínez Sotoca received the B.Sc. de-gree in physics from the Universidad Nacional de Educacio´n a Distancia, Madrid, Spain, in 1996 and the M.Sc. and Ph.D. degrees in physics from the University of Valencia, Valencia, Spain, in 1999 and 2001, respectively. His Ph.D. work was on surface reconstructions with structured light.

He is currently an Assistant Lecturer in the Departamento de Lenguajes y Sistemas Informáticos, Universitat Jaume I, Castellón de la Plana, Spain. He has collaborated on different projects, most of which are in the medical application of computer science. He has published more than 45 scientific papers in national and international conference proceedings, books, and journals. His research interests include pattern recognition and biomedical applications, including image pattern recognition, hyperspectral data, structured light, and feature extraction and selection.

Filiberto Pla received the B.Sc and the Ph.D degrees in physics from the Universitat de València, Spain, in 1989 and 1993, respectively.

He is now a Full Professor at the Departament de Llenguatges i Sistemes Informàtics of the University Jaume I, Castellón, Spain. He has been a Visiting Scientist at Silsoe Research Institute, the University of Surrey, the University of Bristol, U.K., CEMAGREF, France, the University of Genoa, Italy, Instituto Superior Técnico, Lisbon, Portugal, the Swiss Federal Institute of Technology ETH-Zurich, the Idiap Research Institute in Switzerland, and the Technical University of Delft, The Netherlands. At present, he is the Director of the Institute of New Imaging Technologies, at the University Jaume I. His current research interests are color and spectral image analysis, visual motion analysis, 3D image visualization and pattern recognition techniques applied to image processing.

Dr. Pla is member of the Spanish Association for Pattern Recognition and Image Analysis (AERFAI), which is a member of the International Association for Pattern Recognition (IAPR).

Manuel Martínez-Corral received the Ph.D. degree in physics (Best Thesis Award) from the University of Valencia, Valencia, Spain, in 1993.

Currently, he is a Full Professor of Optics at the University of Valencia and co-leads the 3D Imaging and Display Laboratory. His research interests include resolution procedures in 3-D scanning microscopy and 3-D imaging and display technologies. He has supervised 12 Ph.D. theses (three honored with the Best Thesis Award), published over 90 technical papers in major journals (which received more than 2000 citations), and delivered a number of invited and keynote presentations in international meetings.

Dr. Martínez-Corral has been a Fellow of the SPIE since 2010 and of the OSA since 2017. He is co-chair of the SPIE 3D Imaging, Visualization, and Display Conference. He has served as the Topical Editor of the IEEE/OSA Journal of Display Technology, and of the OSA journal of Applied Optics.

Genaro Saavedra received the M.Sc. and Ph.D. (*cum laude*) degrees in physics from the University of Valencia, Spain, in 1990 and 1996, respectively.

He is currently a Full Professor at the University of Valencia. Since 1999, he has been the co-leader of the 3D Imaging and Display Laboratory at that university. His research interests include optical diffraction, phase-space representation of scalar optical fields, and 3-D optical microscopy. He has supervised on these topics five Ph.D. theses (two of them honored with the Best Thesis Award). He has published about 60 technical papers on these topics in major journals. He has published in more than 60 conference proceedings, including more than 40 invited presentations. He filed seven patents on 3-D display and imaging techniques.

Yi-Pai Huang received the B.S. degree from National Cheng Kung University, in 1999, and the Ph.D. degree in electro-optical engineering from the National Chiao Tung University, Hsinchu, Taiwan.

In 2004, he was a Project Leader in the technology center of AU Optronics (AUO). He joined National Chiao Tung University in 2006 as an Assistant Professor and was promoted as a Full Professor in 2013. He is currently the Associate Vice-president of R&D Office at National Chiao Tung University and the Director of Advanced Display Research Center. His expertise includes 3-D display and interactive technologies, optics of AR/VR, display optics, and color science. In the above-mentioned research, he has so far published 60 international journal papers with more than 133 international conference papers, including 33 invited presentations. He has also obtained more than 98 issued patents and contributed three book chapters.

Dr. Huang has four times received the SID's Distinguished Paper Award (2001, 2004, 2009, 2015) and J-SID Best Paper of the Year Award in 2010. Furthermore, he is also currently the General Director of the 3D Interaction and Display Association (3DIDA), Taiwan, the Chairman of the Society of Information Display Taipei chapter (SID-Taipei), and Editorial Board Committee of the *Journal of 3D Research*.

Adrian Stern (Member, IEEE) received the B.Sc., M.Sc. (*cum laude*), and Ph.D. degrees from Ben-Gurion University of the Negev, Israel, in 1988, 1997, and 2003, respectively, all in electrical and computer engineering.

Currently, he is a Full Professor in the Electro-Optical Engineering Department, Ben-Gurion University, where he serves as department head. During the years 2002 through 2004 he was a Postdoctoral Fellow at the University of Connecticut. During 2007 and 2008 he served as a Senior Research and Algorithm Specialist for GE Molecular Imaging, Israel. In 2014 and 2015, during his sabbatical leave, he was a Visitor Scholar and Professor at Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. His current research interests include 3-D imaging, compressed imaging, computational imaging, phase-space optics, biomedical imaging, image restoration and reconstruction. He has published over 150 technical articles in leading peer reviewed journals and conference proceedings, more than a quarter of them being invited papers.

Dr. Stern is a Fellow of SPIE and member of OSA. He has served as an Associate Editor for several journals and is the Editor of *Optical Compressive Sensing and Imaging,* 2016.