# New Paradigms for Active and Passive 3D Remote Object Sensing, Visualization, and Recognition

P. F. McManamon[b], B. Javidi [a], E. A. Watson[b], M. DaneshPanah [a], R. T. Schulein[a]
[a]Dept. of Electrical and Computer Eng., University of Connecticut, Storrs, CT USA 06269
Email: Bahram.Javidi@uconn.edu
[b]U.S. Air Force Research Lab, Sensors Directorate, Wright Patterson AFB, OH USA 45433
Email: Paul.McManamon@wpafb.af.mil

## ABSTRACT

Three dimensional imaging is a powerful tool for object detection, identification, and classification. 3D imaging allows removal of partial obscurations in front of the imaged object. Traditional 3D image sensing has been Laser Radar (LADAR) based. Active imaging has benefits; however, its disadvantages are costs, detector array complexity, power, weight, and size. In this keynote address paper, we present an overview of 3D sensing approaches based on passive sensing using commercially available detector technology. 3D passive sensing will provide many benefits, including advantages at shorter ranges. For small, inexpensive UAVs, it is likely that 3D passive imaging will be preferable to active 3D imaging.

**Keywords:** Automatic target recognition, Three dimensional imaging, Passive sensing, Laser radar

## 1 INTRODUCTION

Performance based sensing involves setting knowledge based objectives to determine what sensors, processors, and personnel are required to obtain desired data sets. In the domains of automated and aided target recognition, of utmost concern is how to capture, classify, and recognize real-world objects. For years, target recognition systems have been based around 2D intensity imaging that determine classification by 2D object intensity (e.g. shape and size, texture), polarization signature or features in the frequency domain [1-8]. While 2D imaging recognition systems offer fast processing speed, they are limited in recognition capabilities by the limitations of 2D imaging [9, 10]. Movement of the object resulting in differing image magnification, changes in ambient lighting conditions, or the introduction of new occluding objects into the imaging path impede on the abilities of a 2D recognition system to accurately sense and classify an object. To overcome these difficulties, advanced techniques are used that add to system complexity and introduces new forms of error to the deterministic recognition algorithms. Three-dimensional (3D) recognition systems offer a number of advantages over 2D image recognition systems [7,8,11-20]. Because object shape and location are known in three dimensions, 3D systems can deal much better with moving objects that disrupt 2D systems. An additional advantage of 3D systems is the ability to segment the object of interest from the background, obscurants, and clutter [21-30]. For 2D imaging, segmentation is much more difficult to perform, is not as accurate, and in certain situations is impossible. Another benefit is the increased performance of target recognition algorithms. Not only can the algorithms operate with images that have less background and clutter, but the additional dimension provides a geometric increase in pixel count producing significantly better performance (generating sharper correlation peaks for example).

Traditionally, 3D shape data has been acquired using an active 3D sensing technology called laser radar in which range is estimated by measuring the time between when a laser pulse is launched and when a receiver detects the optical energy scattered by an object. Angle / angle or cross-range characteristics of the object can be measured either by scanning the laser beam (if the receiver has only a single or limited number of detectors) or by flood illuminating a large area and detecting the scattered return on a 2D focal plane array. A laser radar uses controlled illumination, which means that the sensor can operate at times of day and in locations that other EO sensing modes may not be able to operate. Another characteristic is the ability to turn on the receiver only when the scattered return from an object of interest is expected to arrive, a technique called range gating. Thus one can neglect the backscatter from intervening obscurants and

Automatic Target Recognition XVIII, edited by Firooz A. Sadjadi, Abhijit Mahalanobis
Proc. of SPIE Vol. 6967, 69670E, (2008) · 0277-786X/08/$18 · doi: 10.1117/12.778343

develop a higher contrast image through scattering media. The major downsides of laser radar are the costs involving money, weight, space, and power.

Because of the costs of laser radar, there has been increasing interest in passive 3D sensing and imaging systems among researchers in recent years [8, 21, 24, 31-37]. One of the most promising passive 3D technologies is Integral Imaging (II), where multiple 2D intensity imagers capture 3D image data from multiple perspectives [22, 31, 33, 38,]. Because multiple perspectives are used, optical directional information is obtained in addition to intensity data. The original 3D scene can be recreated optically by back propagating each captured 2D perspective image through the original pickup optics into a display space. By doing this, light cones that originally emanated from a specific location in object space will overlap at the same location in the display space, thus forming a 3D image. This process can be computationally simulated [22-24, 27, 29, 30-32, 36, 37, 40] so that the 2D perspective images are propagated through a virtual optical system to different reconstruction locations. An object will appear in focus when the reconstruction location is the same as the object's original position and the object will appear out of focus at other reconstruction locations. In fact, the further the origin of light rays from the reconstruction plane, the more defocused they appear after reconstruction. Computational reconstruction of elemental images has proved to be promising in a variety of applications including 3D object recognition [7,8, 11, 12, 14, 16-18], occlusion removal [24, 25, 28, 29], multiple viewing point generation [25].

This overview keynote address paper is organized as follows. In section 2, scanning laser radar and flash laser radar are described as two active imaging modalities of LADAR. In section 3, concepts of passive 3D Integral Imaging (II) are presented. Optical pickup and computational reconstruction are discussed. Section 4 is devoted to application of passive 3D II for Automatic Target Recognition (ATR) using optimum non-linear distortion tolerant filters. Section 5 discusses the novel application of 3D passive photon counting Integral Imaging (II) to ATR, Section 6 presents sensing and recognition experimental results for both active and passive systems. We conclude the paper in section 7 with a discussion of the benefits of the active and passive approaches under different circumstances.

# 2 LASER RADAR BASED IMAGING

## 2.1 Scanning Laser Radar

This type of system utilizes a simple receiver (a single detector, or a limited number of detectors, with electronics that are relatively easy to implement) and a high pulse repetition rate laser (PRF ~ 10's of KHz) with low energy per pulse. While these characteristics may make those components of the laser radar easier to fabricate, the complexity comes in the optical aperture, scanning mechanism, image stabilization, and post-detection processing. Agile beam steering is required to scan the laser and the field of view of the detector, while considerable post detection processing is needed to compensate for motion while the image is gathered over significant fractions of a second. Precise image stabilization, to a fraction of one angular resolution pixel, is required or it will not be possible to put the image together. A 10 KHz PRF laser would require over 1 second to capture a 128X128 image. As a result, platform and object motion will result in smearing of the image that must be undone in post detection processing. While such image artifacts can reduce image quality in the cross-range dimensions, range estimation can be high quality because high bandwidth range estimation electronics for a single detector are easier to fabricate than for an array of detectors. For an array of detectors, high bandwidth electronics must either be located behind each pixel or located around the focal plane in some manner, thus reducing fill factor. In addition, the lower pulse energies required for a high PRF laser imply that shorter pulse widths can be used, resulting in less range blur. A benefit of high rep rate lasers used for scanning systems is that laser diodes tend to be quasi continuous wave (CW) sources, so the laser diode cost will be lower than for a low duty cycle laser radar system.

We can evaluate the utility of scanning laser radar for platforms at various altitudes. Interestingly, the signal to noise ratio does not necessarily change much for the different platforms. The desired sampling rate and resolution at the target will determine the laser spot size and number of measurements across the object. Hence, it is assumed that the optical system on different platforms will change to produce constant resolution (which implies the optics diameter changes) and constant sampling rate (which implies the focal length changes). The result is the f/# of the receiver optics remains constant and hence so does the gathered image irradiance. There is a dependence on range associated with losses in the atmosphere, but for clear air these losses are not significant if the wavelength of the laser radar is chosen properly. Obscurants such as fog, clouds, dust, etc, can of course introduce large range dependence. Typical pulse energies required in clear air can be on the order of 0.1 - 1 mJ for thermal noise limited receivers (no cooling) with reasonable optical gain such as provided by an avalanche photodiode.

One parameter that does change with platform is the angular scanning or pointing accuracy. To provide undistorted imagery the laser must be pointed within a small fraction of the sampling distance. If a ground sampling distance of one foot is desired, then pointing accuracies on the order of a few mrad are required for ranges of hundreds of meters, while accuracies on the order of 10 microrad are required for 10's of km ranges. Such pointing accuracies are expensive to achieve.

### 2.2 Low Rep Rate Flash Laser Radar

When used with an array of linear mode APDs a flash imaging system is characterized by low PRF lasers (10's of Hz) and high peak pulse energies. In single pulse flash imaging, images are acquired over a few nanoseconds so platform and object motion have little effect. Large pulse energies are required for low rep rate flash imaging to flood-illuminate the object to be measured (actually an area larger than the object to overcome pointing errors). There is an increased risk of damage to the optical components in the transmitter and an increased cost of quasi CW diodes to pump the high pulse energy lasers. Eye safety is also a concern.

There are a few techniques that can be used to implement a flash ladar. One method is to use high bandwidth electronics behind or in close proximity to each detector in the array. This has resulted in the need to trade the number of range bins that can be recorded against the physical size of the detector pixel. Larger physical format focal plane arrays require larger focal length optics to provide the required sampling rate at the ranges of interest. Longer focal lengths are more difficult and expensive to implement. A second approach to low rep rate flash imaging is to use an external polarization rotator in front of the detector area. The orientation of polarization is then used to encode range and a slow response detector array can be used.[i]

An advantage of all flash imaging techniques is reduced processing load due to the entire image being gathered within a few nanoseconds. Another significant advantage for longer range applications is reduced image stabilization requirements. For flash imaging you only need to angle stabilize to a fraction of the flash illumination beam divergence. For a 128 x 128 detector area this means you only need to stabilize to 128 times the angular accuracy you need for a scanning 3D laser radar. Pointing to only a fraction of the illuminated spot diameter is required, or about 1 mrad at 10 km range. This can have a significant positive cost impact.

As with the scanning ladar system, SNR does not vary much as a function of platform or range because the optics must change to keep a constant illuminated spot size on the ground in transmission and constant f/# on receiver to maintain sampling and resolution. The pulse energies, however, are much higher because the entire scene must be illuminated on each pulse. Hence, a 10m x 10m illuminated scene with .3 m x .3 m resolution would require about a thousand times more energy per pulse than the example given above for the scanned laser radar. This increases the cost of the laser diodes required to produce such a high energy pulse. A long storage time solid state medium mitigates the diode cost issue.

### 2.3 C High Rep Rate Flash Imaging Using Geiger Mode APDs

A Gieger mode 3D imaging system will have low pulse energies and high rep rate, similar to a scanning system. The detector array in this case is sensitive enough to detect a single incident photon. Pulse energies from the laser are deliberately kept low so that the probability of detecting a single photon per laser pulse per pixel is less than unity. This prevents saturation of the detector from intervening obscurants. The image is then built up on a statistical basis using many pulses over the same area to compile a range distribution. A typical rep rate will be tens of kilohertz. Pointing accuracy and required stabilization is similar to low rep rate flash imaging. Laser requirements are similar to scanning systems. Processing may be complex because many samples must be used to estimate the location of the surface of the object. This technique usually produces range images with limited gray scale.

## 3 PASSIVE THREE DIMENSIONAL SENSING

### 3.1 3D Passive Integral Imaging (II)

Among various techniques which can quantitatively measure one or more of the psychological depth cues, one major thrust is Integral Imaging (II) (a.k.a integral photography) which is based on the original work of Lippmann with lenticular sheets [41, 42] and is classified under multi-perspective 3D imaging systems. Integral Imaging provides autostereoscopic images by recording intensity and direction of light rays (i.e. the light field [31]) in the form of a set of elemental images from slightly different perspectives. This technique is a promising method compared to other techniques due to its continuous viewing angle, full parallax and full color display without the need for coherent sources of illumination and its relative simplicity of implementation.

Conventional II systems use a microlens array to capture light rays emanating from 3D objects in such a way that the light rays that pass through each pickup microlens are recorded on a 2D image sensor [22, 30, 35, 36, 40]. The captured 2D images are referred to as elemental images. The elemental images are 2D images, flipped in both the x and y direction, each with a different perspective of a 3D scene. To reconstruct the 3D scene optically from the captured 2D elemental images, the rays are reversely propagated from the elemental images through a display microlens array that is similar to the pickup microlens array. Developments in 3D II display [21, 33] include use of gradient index lens arrays to handle the orthoscopic to pseudoscopic conversion, also resolution improvement methods including use of moving lenslet technique (MALT) [36, 40] and electronically synthesized moving Fresnel lenslets. However, the optical reconstruction approach suffers from low resolution, low sampling rate, quality degradation due to light diffraction, limited dynamic range and overall visual quality due to limitation of electro-optical projection devices.

In order to overcome image quality degradation introduced by optical devices used in the optical II reconstruction process, and also to obtain arbitrary perspective within the total viewing angle, computational II reconstruction techniques have been proposed [22-32, 36]. Computational II simulates the optical reconstruction process by back propagating the elemental images through a virtual optical system similar to the original pickup system. Objects will appear in focus at a reconstruction plane equal to the original object distance and will appear out of focus at other reconstruction planes. This technique allows us to reconstruct 3D voxel values at any arbitrary distance from the display microlens array and create 3D data sets that may be manipulated or analyzed. Because computational reconstruction operates in the electronic domain rather than optical, optical reconstruction image degradations are eliminated. However, manipulation of integral image data does cause some computational burden.

Conventional II systems use an array of small lenses mounted on a planar surface (lenslet array) to capture elemental images on a single opto-electronic sensor. There has been effort to increase depth of field of each lenslet [35]. Each lens creates a unique perspective view of the scene at its image plane. As long as elemental images do not overlap in the image plane of the lenslet, one can capture all elemental images on a CCD at once. Figure 1 depicts a system setup to capture an occluded 3D scene using lenslet arrays. This technique has the merits of simplicity and speed. However, for objects close to the lenslet array, the elemental images may overlap in the image plane, which requires one to use additional optics to project the separated elemental images on a sensor. Also, the small aperture of the lenslets creates low resolution or abberated elemental images. In addition, the pixels of the imaging device has to be divided between all elemental images which leads to low number of pixels per elemental image.

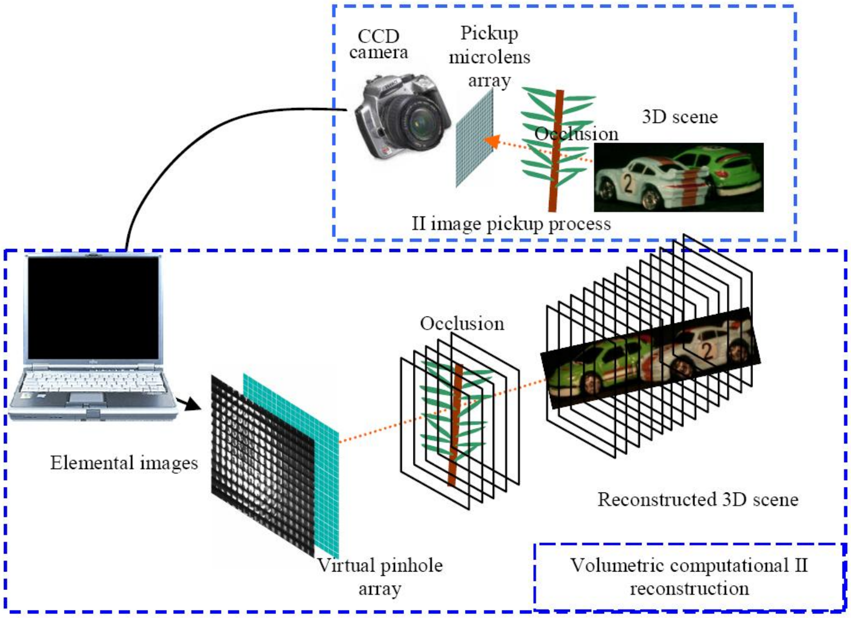


Fig. 1. Experimental setup for 3D object capturing and computational reconstruction with II system. The full 3D volume image at various distances can be reconstructed separately.

Synthetic Aperture Integral Imaging (SAII) is a method that can potentially resolve some of the problems associated with conventional II [23, 25, 29]. In this technique, each perspective view is acquired separately by a 2D imaging sensor located at the respective location on the pickup aperture. Thus, this method involves mechanical translation of one or more imaging devices on a grid and capturing elemental images at certain positions. This way, high resolution perspective views are captured from a synthetic aperture in which the imaging device scans. Figure 2 shows a conceptual diagram of the SAII technique. Since capturing the perspective views requires multiple acquisitions, SAII in this form is not suitable for imaging dynamic objects in which the movements are faster than the time required for a complete aperture scan. However, methods have been proposed to solve this problem by introducing an array of imaging devices (cameras) on a grid. In addition, since the elemental images can be captured with well corrected optics on a large optoelectronic sensor, the resolution and aberration of each elemental image can be enhanced dramatically comparing to lenslet based II. [25, 26]

### 3.2 Computational reconstruction of elemental images

Once the elemental images from different perspectives were picked up, the collected visual information can be computationally reconstructed to recreate the 3D scene. Several methods have been investigated for computational reconstruction of II data. In the Fourier domain, digital refocusing has been proposed [30] by applying Fourier slice theorem in 4D light fields. This technique is relatively fast with complexity of $O(n^2\log n)$, $n$ being the total number of image pixels. In the spatial domain, a fast, ray tracing based reconstruction from the observers point of view is proposed in [22] with complexity of $O(m)$, $m$ being the number of elemental images. Although fast and simple, this method yields low resolution reconstructions. Yet another spatial domain reconstruction method is based on series of 2D image back projections [25, 37]. This method offers a much better reconstruction resolution comparing to at the expense of an algorithm with complexity of $O(n)$, since typically $n>>m$.

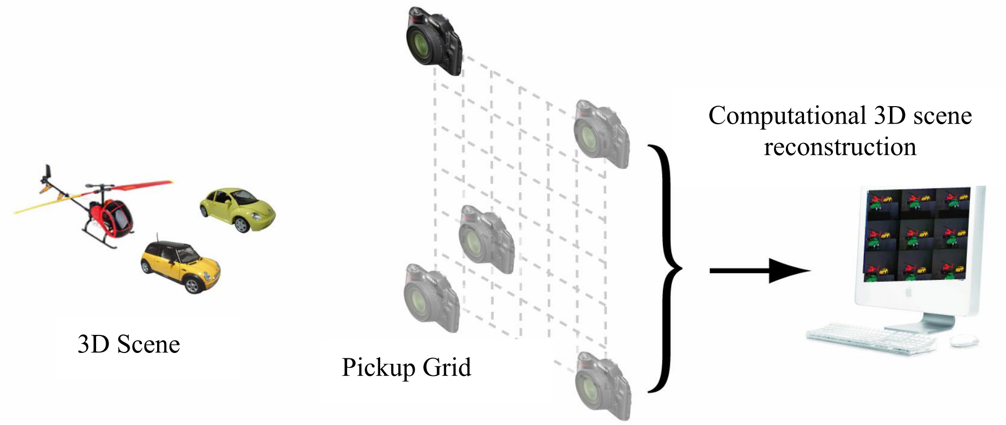


Fig. 2. Pickup process of three dimensional synthetic aperture integral imaging. The camera at each location captures the scene form a unique perspective, which is later used for 3D computational reconstruction.

For 3D computational reconstruction in spatial domain, one approach is to simulate the pinhole array from which the elemental images were taken and perform the inverse propagation for rays. Since each elemental image conveys a unique perspective of the scene, the directional information is also recorded in an integral image as well as the 2D intensity information. The depth information in particular can be extracted from the relative shift of an object between the elemental images. Therefore, a 2D scene can be reconstructed at a particular distance by properly propagating the rays back from their respective pickup locations. The collection of 2D scenes reconstructed at all distances then gives the full 3D scene.

A computationally efficient algorithm for II reconstruction is described in [37]. Each elemental image can be described as $O_{kl}(x,y)$, where $k$ and $l$ denote the position indices of the elemental image in the pickup grid [see Fig. 3]. The magnification factor $M$, is given by $z_0/g$, where $z_0$ is the desired reconstruction distance and $g$ denotes the effective focal length of the sensor. In a synthetic aperture mode [see Fig. 2], the reconstruction by ray back propagation is possible by flipping elemental images and shifting them according to $z_0$, and averaging the overlapping pixels. The following expression describes the reconstruction process:

$$I(x,y,z_0)=\sum_{k=1}^{K}\sum_{l=1}^{L}O_{kl}\left(-x+\left(1+\frac{1}{M}\right)S_xk,-y+\left(1+\frac{1}{M}\right)S_yl\right)/R^2(x,y) \qquad (1)$$

in which $S_x$ and $S_y$ denote the separation of sensors in $x$ and $y$ directions at the pickup plane respectively, $K$ and $L$ denote the number of elemental images acquired in the $x$ and $y$ directions; also $R$ compensates for intensity variation due to different distances from the object plane to elemental image $O_{\mathrm{kl}}$ on the sensor and is given by $R^2(x,y)=(z_0+g)^2+[(Mx-S_xk)^2+(My-S_yl)^2]\times(1+M^{-1})^2$ [see Fig. 3 and [37]].

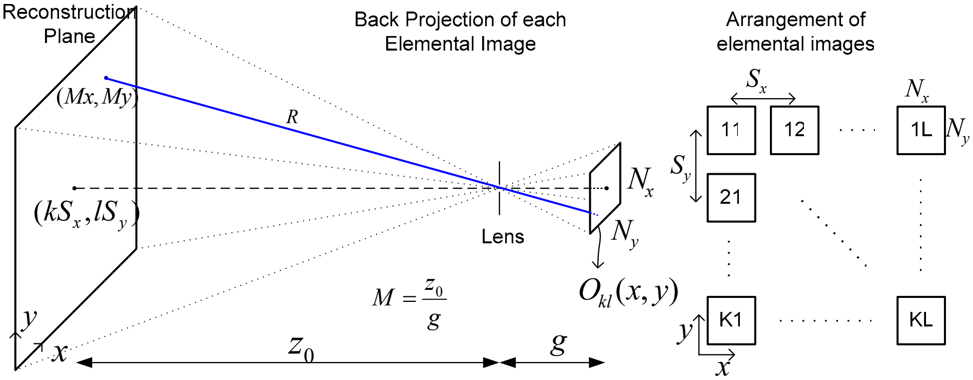


Fig 3. Schematic of II reconstruction process (left), Arrangement of elemental images (right)

Note that in computational reconstruction the adjacent flipped and shifted elemental images overlap such that for objects close to the reconstruction plane, the overlap of all elemental images would be aligned, whereas for objects located away from the reconstruction plane the overlap would be out of alignment resulting in a blurred reconstruction. Thus, with computational reconstruction one is able to get an in-focus image of an object at the correct reconstruction distance, while the rest of the scene appears out of focus.

# 4 AUTOMATIC TARGET RECOGNITION USING 3D PASSIVE SENSING

In a computational three-dimensional (3D) volumetric reconstruction integral imaging (II) system, volume pixels (voxels) of the scene are reconstructed plane by plane. At desired target planes where the target is in focus, the foreground occlusion appears completely washed out if there is enough spatial separation between the occlusion and the occluded object. Using volumetric computational II reconstruction, we are able to recognize distorted and/or occluded objects with correlation based recognition algorithms [11, 15]. A distortion tolerant optimum non-linear filter is developed based on minimize a linear combination of the output energy due to the input noise and the output energy due to the input scene under the filter constraint [7, 13, 15].

## 4.1 Distortion Tolerant Optimum Non-Linear Filters for ATR

In this section we briefly overview synthesis of a distortion tolerant optimum nonlinear filter for ATR [12, 13, 15]. When $r_{\mathrm{i}}(t)$ denotes one of the distorted reference targets where $i$ = 1, 2, …, $T$, $T$ is the size of reference target set, the input image $s(t)$ is:

$$s(t)=\sum_{i=1}^{T}v_ir_i(t-\tau_i)+n_b(t)\left[w(t)-\sum_{i=1}^{T}v_iw_{ri}(t-\tau_i)\right]+n_a(t)w(t), \qquad (2)$$

where $v_i$ is a binary random variable which takes a value of 0 or 1. $v_i$ indicates whether the target $r_i(t)$ is present in the scene or not. $p(v_i=1)=1/T$, $p(v_i=0)=1-1/T$. If $r_i(t)$ is one of the reference targets, $n_b(t)$ is the non-overlapping background noise with mean $m_b$, $n_a(t)$ is the overlapping additive noise with mean $m_a$, $w(t)$ is the window function for the entire input scene, $w_{ri}(t)$ is the window function for the reference target $r_i(t)$, $\tau_i$ is a uniformly distributed random location of the target in the input scene, whose probability density function is $f(\tau_i)=w(\tau_i)/d$ ($d$ is the area of the support region the input scene). $n_b(t)$ and $n_a(t)$ are assumed to be wide-sense stationary random processes and statistically independent of each other.

The filter is designed so that when the input to the filter is one of the reference targets, then the output of the filter in the Fourier domain expression becomes:

$$\sum_{k=0}^{M-1} H(k)^* R_i(k) = MC_i \quad , \tag{3}$$

where $H(k)$ and $R_i(k)$ are the discrete Fourier transforms of $h(t)$ (impulse response of the distortion tolerant filter) and $r_i(t)$, respectively. * denotes complex conjugate, $M$ is the number of sample points, and $C_i$ is a positive real desired constant. Equation (3) is the constraint imposed on the filter. To obtain noise robustness, we minimize the output energy due to the disjoint background noise and additive noise. We can gather both disjoint background and additive noise in one noise term and define $n(t) = n_b(t)\left\{w(t) - \sum_{i=1}^{T} v_i w_{ri}(t-\tau_i)\right\} + n_a(t)w(t)$. We minimize a linear combination of the output energy due to the input noise and the output energy due to the input scene under the filter constraint: The minimization is done through Lagrange multipliers $\lambda_{1i}$, $\lambda_{2i}$. For detailed derivation of the filter, we refer the reader to [15]. The following optimum nonlinear distortion tolerant filter $H(k)$ is computable:

$$\begin{aligned} H(k) = {} & \sum_{i=1}^{T} \left(\lambda_{1i} - j\lambda_{2i}\right) R_i(k) \Big/ \\ & \Bigg( \frac{1}{MT} \sum_{i=1}^{T} \left( \Phi_b^0(k) \otimes \left\{ |W(k)|^2 + |W_{ri}(k)|^2 - 2\frac{|W(k)|^2}{d} \mathrm{Re}\left[W_{ri}(k)\right] \right\} \right) + \frac{1}{M} \Phi_a^0(k) \otimes |W(k)|^2 \\ & + \frac{1}{T} \sum_{i=1}^{T} \left( m_b^2 \left\{ |W(k)|^2 + |W_{ri}(k)|^2 - 2\frac{|W(k)|^2}{d} \mathrm{Re}\left[W_{ri}(k)\right] \right\} + 2m_a m_b |W(k)|^2 \mathrm{Re}\left[1 - \frac{W_{ri}(k)}{d}\right] \right) \\ & + m_a^2 |W(k)|^2 + |S(k)|^2 \Bigg) \end{aligned} \tag{4}$$

where $\Phi_b^0(k)$ is the power spectrum of the zero-mean stationary random process $n_b^0(t)$, and $\Phi_a^0(k)$ is the power spectrum of the zero-mean stationary random process $n_a^0(t)$. $W(k)$ and $W_{ri}(k)$ are the discrete Fourier transforms of $w(t)$ and $w_{ri}(t)$, respectively. $\otimes$ denotes a convolution operator. $\lambda_{1i}$ and $\lambda_{2i}$ are the Lagrange multipliers.

If we have an input model without background noise, the optimum nonlinear distortion tolerant filter $H(k)$ becomes:

$$H(k) = \frac{\sum_{i=1}^{T} \left(\lambda_{1i} - j\lambda_{2i}\right) R_i(k)}{\left( \frac{1}{M} \Phi_a^0(k) \otimes |W(k)|^2 + m_a^2 |W(k)|^2 \right) + |S(k)|^2} \tag{5}$$

This filter can be created by using true class training targets. Once synthesized, this filter can be used to distinguish between true class non-training objects and false class objects.

## 5 PHOTON COUNTING THREE-DIMENSIONAL PASSIVE SENSING FOR ATR

In this section, we present 3D passive sensing ATR using photon counting integral imaging [16-20]. Photon counting is a passive sensing technique which has been applied in many fields such as night vision, laser radar imaging, radiological imaging, and stellar imaging. The photon counting can be binary at low light level [10, 43]. The advantage of the photon-counting detector may be enhanced by the processing of binary photon numbers which can be simpler and faster. Photon counting techniques have been applied to infrared imaging and thermal imaging [43]. Photon counting detectors have been considered for 3D active sensing by LADAR as well.

Photon counting 3D passive sensing shows significant benefits for automatic target recognition (ATR) [16-18]. The discrimination capability of the proposed system is quantified in terms of Fisher ratio and receiver operating characteristic (ROC) curves using nonlinear matched filtering. Each lenslet in the lenslet array generates a photon-limited elemental image on a photon counting detector array. Multiple perspectives of photon-limited scenes are recorded according to the corresponding lenslet.

The aim of automatic target recognition (ATR) is to identify unknown objects in a scene and categorize them into distinct classes. In this section, we review the passive sensing and recognition of 3D objects by means of photon counting integral imaging [16-18]. Nonlinear matched filtering is developed for the recognition of 3D objects. The nonlinear matched filtering is performed between a reference (irradiance image) and unknown inputs (photon-limited images). We present analysis of the statistical properties of the nonlinear correlation normalized according to the power law of the sum of photon numbers assuming the low level of photons. We define our matched filtering as the nonlinear correlation normalized with the power *v* of the photon-limited image as shown below:

$$C_{rs}(x_j;v) = \frac{\sum_{i=1}^{N_T} R(x_i + x_j)\hat{S}(x_i)}{\left(\sum_{i=1}^{N_T} R^2(x_i)\right)^{\frac{1}{2}}\left(\sum_{i=1}^{N_T} \hat{S}(x_i)\right)^{v}} = \frac{\sum_{k=1}^{N} R(x_k + x_j)}{A\left(\sum_{i=1}^{N_T} \hat{S}(x_i)\right)^{v}}, \tag{6}$$

where $A = \left(\sum_{i=1}^{N_T} R^2(x_i)\right)^{\frac{1}{2}}$, *R* is the irradiance of the reference image which is denoted by *r*, *s* represents an unknown input object from which the photon-limited image $\hat{S}$ is generated, $N_T$ is the total number of pixels, and *N* is the total number of photon detection events occurred in the scene. It is noted that $C_{rs}(x_j;v)$ has the maximum value at $x_j = 0$ in our experiments. One advantage of photon counting detection is that the computational time of the matched filtering at low light level is much faster than conventional image correlation. As shown in the second term in Eq. (6) the correlation becomes merely the sum of the reference radiance at particular pixels (photon arrivals) [18].

The following statistical properties of nonlinear correlation peak $C_{rs}(0;1)$ have been proven in [16]:

$$\langle C_{rs}(0;1)\rangle \approx \frac{1}{A}\sum_{i=1}^{N_T} R(x_i)S(x_i), \tag{7}$$

$$\mathrm{var}\left(C_{rs}(0;1)\right) \approx \frac{1}{A^2}\left(\frac{1}{N_p}\sum_{i=1}^{N_T} R^2(x_i)S(x_i) - \sum_{i=1}^{N_T} R^2(x_i)S^2(x_i)\right), \tag{8}$$

where $\langle\cdot\rangle$ denotes the expectation operator, "var" denotes the variance operator, *S* is the irradiance of the input object *s*, and $N_P$ is a predetermined mean number of photo-counts in the entire scene. The nonlinear matched filtering shows different behaviors according to *v*. For the linear matched filtering (*v*=0), both the mean and variance of the correlation peak $C_{rs}(0;0)$ are approximately proportional to $N_P$ [16]. However, the mean of $C_{rs}(0;1)$ does not depend on the number of photons, i.e., the same correlation value can be theoretically achieved with any small number of photons. Although the variance of $C_{rs}(0;1)$ increases when using lower number of photons, this property of photon-limited images might be beneficial for pattern recognition applications.

## 6 EXPERIMENTAL RESULTS

### 6.1 Occlusion removal

Figure 4 shows two toy cars and foreground vegetation illuminated by incoherent light used in the experiments [15].

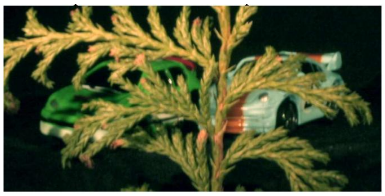


Fig. 4. 3D object used in the experiments. The blue car is a true class target, green car is a false object. Vegetation is positioned in front of the cars to partially occlude the background objects.

The pickup microlens array is placed in front of the object to form the elemental image array. The distance between the microlens array and the closest part of the occluding vegetation is around 30 *mm*, the distance between the microlens array and the front part of the green car is 42 *mm*, and the distance between the microlens array and the front part of the blue car is 52 *mm*. The minimum distance between the occluding object and a pixel on the closest background object should be equal to or greater than 9.6 *mm*, where the rhombus index number in our experiments is 7 for the green car.

This satisfies the constraint of the experimental setup to reconstruct the background objects. The background objects are partially occluded by foreground vegetation, thus, it is difficult to recognize the occluded objects from the 2D scene in Fig. 4. The elemental images of the object are captured with the digital camera and the pickup microlens array. The microlens array used in the experiments has 53 × 53 square refractive lenses. The size of each lenslet is 1.09 *mm* x 1.09 *mm*. The focal length of each microlens is 3.3 *mm*. The size of each captured elemental image is 73 pixels × 73 pixels.

With volumetric computational II reconstruction, Eq. (1), it is possible to separate the foreground occluding object and background occluded objects with the reduced interference of the foreground objects. Figs. 5 show the reconstruction of the objects at different depths, respectively.

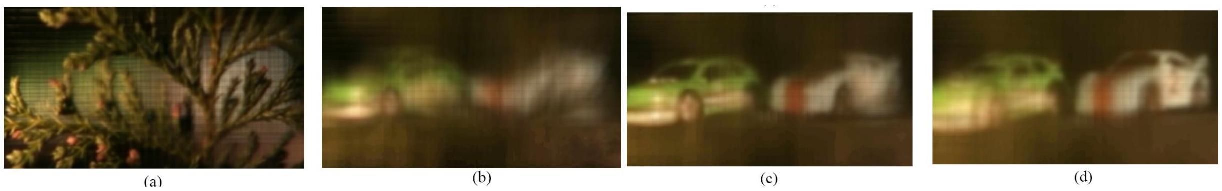


Fig. 5. Reconstructed images from the elemental images sets rotated at an angle of 32.5° at (a) $z$ = 29 mm, (b) $z$ = 45 mm, (d) $z$ = 52 mm, and (d) $z$ = 69 mm.

### 6.2 Distortion Tolerant Automatic Target Recognition

In this section we present experimental results which show recognition of 3D rotated and occluded targets in a reconstructed scene [15]. We also show the ability of the proposed technique to recognize distorted and occluded 3D non-training targets. In our experiments, we have used a blue car as a true class target, and a green car as a false object. We have obtained 7 different elemental image sets by rotating the reference target from 30° to 60° in 5° increments. The reconstructed images from the elemental image sets are shown in Fig. 6 at various angles.

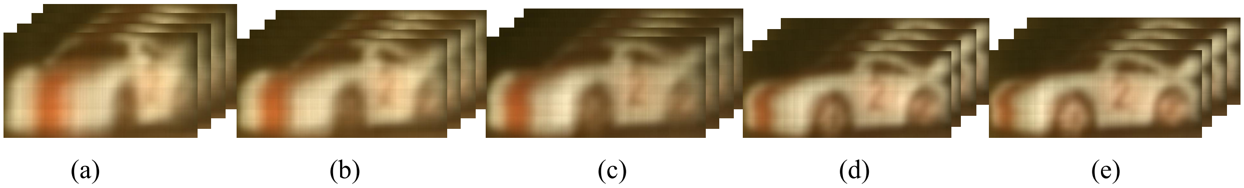
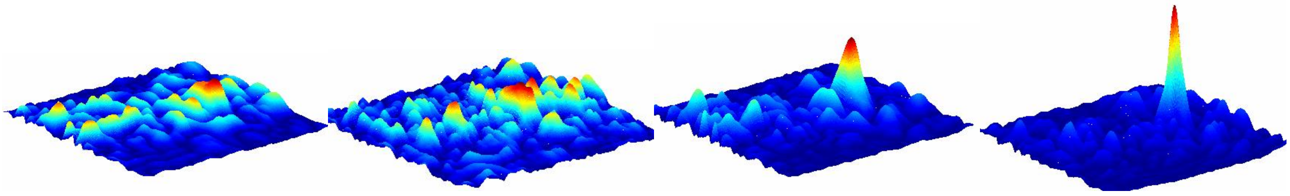


Fig. 6. Five of the seven sets of the reconstructed images from the elemental image sets ranging from z = 60 mm to z = 72 mm with 1mm increment. These 7 reconstructed image sets are spaced at 5° apart from 30°-60°. Sets shown at (a) 30°, (b) 40°, (c) 50°, (d) 55°, and (e) 60°

From each elemental image set with rotated targets, we have reconstructed the images from $z$ = 60 *mm* to $z$ = 72 *mm* in 1*mm* increments. Therefore, for each rotated angle (from 30° to 60° in 5° increments) 13 reconstructed images are used as a 3D training reference target. The input elemental images have a true class training target, or a true class non-training target and a false object (green car). A true class training and non-training target are located on the right side of the input scene and the false object is located at the left side of the scene.

The true class non-training target used in the test is distorted in terms of out-of-plane rotation, which is challenging to detect. With volumetric computational II reconstruction, it is possible to separate the foreground occluding object and background occluded objects with the reduced interference of the foreground objects. According to Eq. (4) The distortion tolerant optimum nonlinear filter has been constructed in a 4D structure, that is, $x$, $y$, $z$ coordinates and 3 color components. We set all of the desired correlation values of the training targets, $C_i$, to 1 in Eq. (3).

Figures 7(a)-7(d) are the normalized outputs of the 3D optimum nonlinear distortion tolerant filter at the depth level of the occluding foreground vegetation, the true class non-training target, and the false object, respectively. Figure 7(d) shows a dominant peak at the location of the true class non-training target. The peak value of the true class training target is higher than that of the true class non-training target. The ratio of the non-training target peak value to the training target peak value is 0.9175. The ratio of the peak value to the maximum side-lobe is 2.8886 at the 3D coordinate of the false object. It is possible to easily distinguish the true class targets and false object or occluding foreground objects. Therefore, we can easily threshold the output level to detect the 3D location of the training and distorted true class non-training targets.

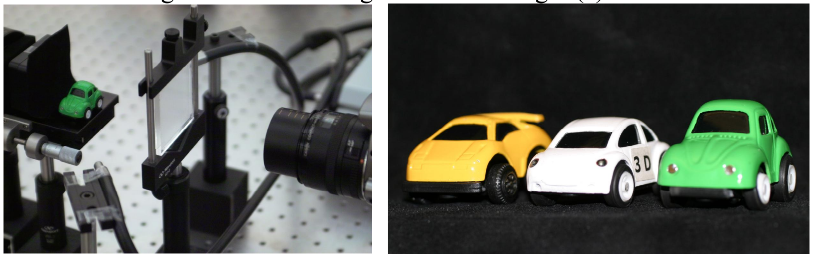


(a) (b) (c) (d)

Fig. 7. Normalized optimum nonlinear filter output for the reconstructed input scene in Figure 7 with 32.5° rotated true class non-training target and a false object at (a) $z$ = 29 mm, (b) $z$ = 45 mm, (c) $z$ = 52 mm, and (d) $z$ = 69 mm,

### 6.3 Photon Counting Three Dimensional Passive Sensing for ATR

In this section, we present the results of passive photon counting ATR [16-18]. For experiments, we use a lenslet array and a pick-up camera for the recording of elemental images as seen in Fig. 8(a).



(a) (b)

Fig. 8. (a) Experimental setup for the integral imaging, (b) three cars used in the experiments; car 1, 2 and 3 are shown from right to left.

Three sets of elemental images have been obtained for each car [see Fig. 9]. The irradiance image of the reference (r) or the unlabeled input (s) corresponds to one set of elemental images captured during the recording process. We generate photon-limited images, each with a random number of photons following Bernoulli distribution. To compute the statistical means and variances we generate 1000 images for each car. We also vary the mean value photon numbers (NP) from 10 to 1,000. The irradiance image of car 1 is used as our reference image.

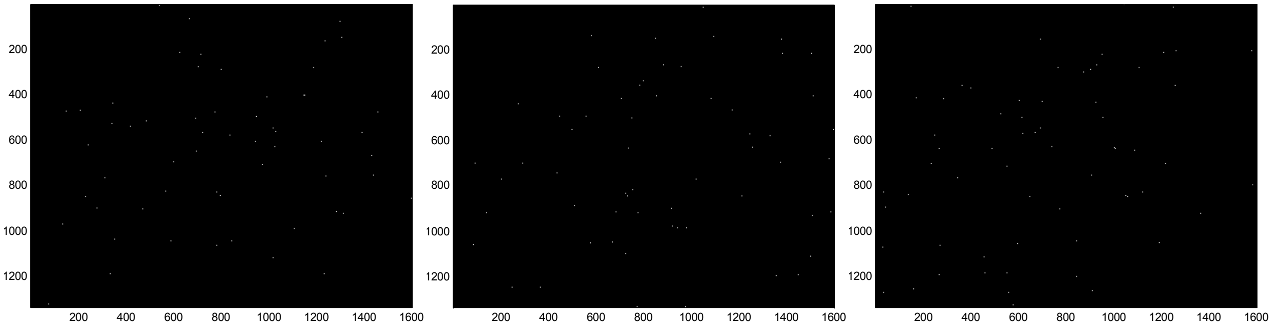
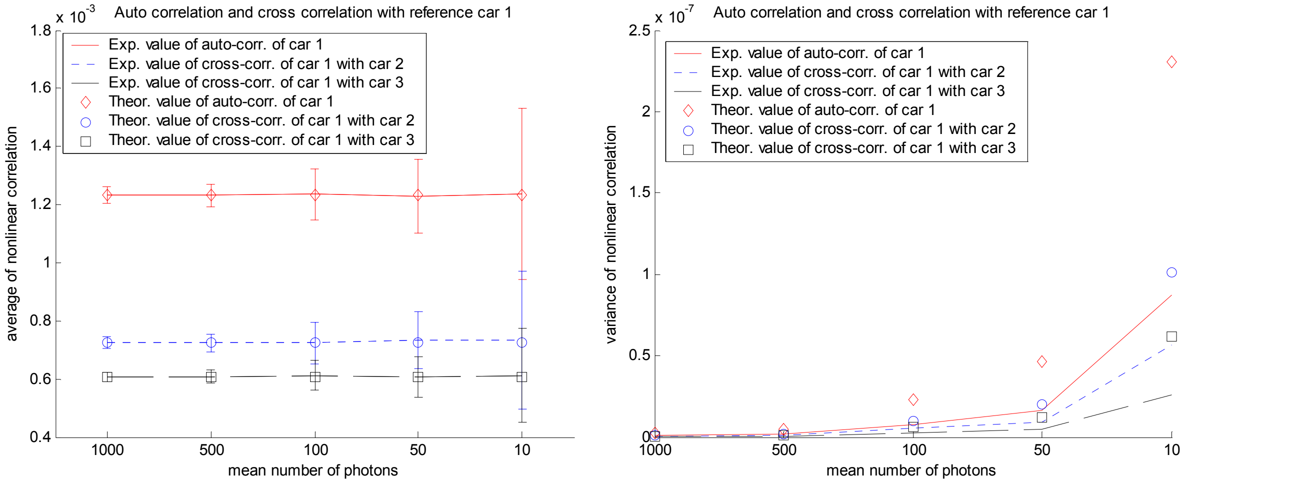


Fig. 9. Photon-limited image when $N_P$=1,000 of three cars used in the experiment, (a) car 1 (green), (b) car 2 (white), (c) car 3 (yellow).

Figure 10(a) shows the experimental results (sample mean) of correlation coefficients and their fluctuations (sample standard deviation) when v=1 with theoretical prediction in Eq. (7). The red solid line graph represents the sample mean of autocorrelation between the irradiance image and photon-limited images of car 1, and the blue dotted line graph is the sample mean of cross-correlation between the irradiance image of car 1 and photon-limited images of car 2, and the black dashed line graph is the sample mean of cross-correlation between the irradiance image of car 1 and photon-limited images of car 3. Figure 10(b) shows the sample variance of Crs(0;1) with theoretical prediction in Eq. (8). The deviation from the theoretical prediction becomes larger as the number of photons decreases as shown in Fig. 10(b).



(a) (b)

Fig. 10. Mean and variance of $C_{rs}$ (0;1), (a) sample mean and theoretical prediction, (b) sample variance and theoretical prediction.

Figures 11(a)-(d) show ROC curves corresponding to cars ($r$=1, $s$=2) and cars ($r$=1, $s$=3) for $C_{rs}$ (0;0) and $C_{rs}$ (0;1), respectively. The number of photons varies from 100 to 10.

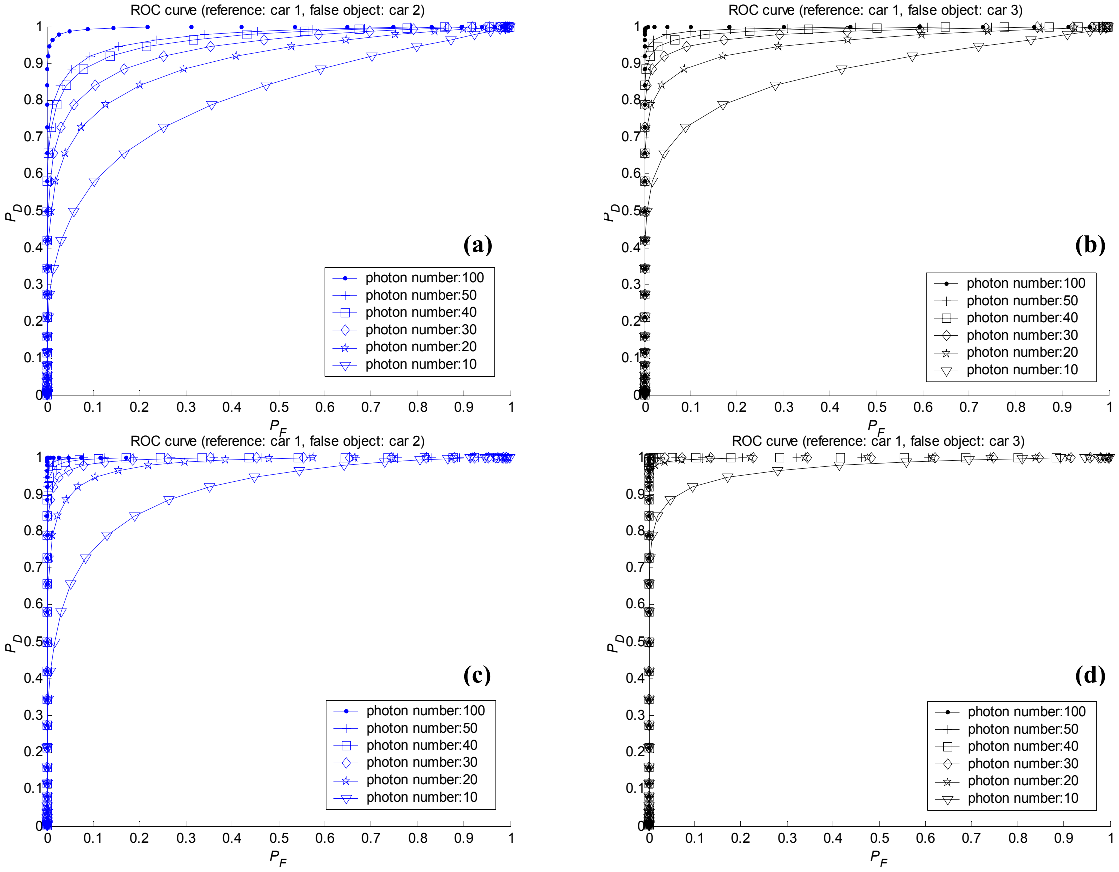


Fig. 11. ROC curves when $r$ = 1, (a) $s$ = 2 for $C_{rs}$(0;0), (b) $s$ = 3 for $C_{rs}$(0;0), (c) $s$ = 2 for $C_{rs}$ (0;1), (d) $s$ = 3 for $C_{rs}$ (0;1).

Table 1 shows the Fisher ratio defined for $C_{rs}(0;0)$ and $C_{rs}(0;1)$ [16]. Fisher ratio decreases when using a lower number of photo-counts, but for photo-counts greater than one hundred, the Fisher ratios show a good separability between the reference and false objects when $v = 1$.

**Table 1. Fisher ratios when $r = 1$**

| $N_P$ | | 1000 | 500 | 100 | 50 | 10 |
|---|---|---|---|---|---|---|
| $v = 0$ | $s = 2$ | 77.41 | 38.63 | 7.53 | 3.59 | 0.76 |
| | $s = 3$ | 131.04 | 66.05 | 12.98 | 6.46 | 1.33 |
| $v = 1$ | $s = 2$ | 204.14 | 105.54 | 19.90 | 9.5 | 1.77 |
| | $s = 3$ | 377.83 | 191.69 | 37.72 | 18.16 | 3.5 |

### 6.4 Underwater 3D Integral Imaging

We have conducted introductory modifications to the integral imaging reconstruction algorithms to conduct ranging experiments with an SAII system viewing objects submerged in water [28]. To the best of our knowledge, we are the first to report on the implementation of integral imaging systems for underwater 3D imaging applications.

When imaging underwater, one must consider refraction caused by water's ~1.33 index of refraction differing from the ~1.00 index of refraction of air. An object located in water at a physical distance $z_{water}$ from the air-water interface will appear to have a distance $z_{water}' = z_{water}/n_{water}$ from the interface to an observer looking perpendicular to the interface. The overall pathlength from observer to object is equal to the distance between the observer and the air-water interface added to the apparent water distance. To accurately reconstruct to an underwater plane, magnification must be modified as $M_r=[z_{air} + z_{water} / n_{water}] / f_r$, with $z_{air}$ the distance between the lens and air-water interface, $z_{water}$ the physical distance between the air-water interface and underwater object, and reconstruction focal length $f_r$ again equal to the focal length of the acquisition lens, $f_l$. Additionally, a camera looking perpendicular to the air-water interface will experience a change in its angle of view such that $\theta'_{HAOV} = \arcsin(\sin(\theta_{HAOV})/1.33)$.

A scene was set up in fish tank consisting of a sign placed 230mm away from the viewing wall, a treasure chest at 340mm, a toy fish at 380mm, and a Lego deep sea creature from 435mm to 495mm. The tank was filled with water and illuminated from above by diffuse incoherent light to simulate sun light. Data was collected by translating a single camera along a transverse x-y 5mm x 5mm grid with the lens flush with the fish tank, and $z_{air}$ was considered to be 0mm. Overall, 9 horizontal nodes and 7 vertical nodes were used for 63 elemental images in a 40mm x 30mm plane. A full frame 35.8mm x 29.9mm CMOS imaging sensor with 4,368 x 2,912 pixels and 8.2µm pixel pitch was used with a 50mm lens. The lens was stepped down to its smallest aperture to achieve the maximum depth of field. A central 2D image of the unoccluded scene is shown in Fig. 12(a), and Fig. 12(b) shows the same scene after foreground occlusion has been added. Three-dimensional computational reconstructions of the occluded scene are shown in Fig. 12(c)-(d). Each object is clearly in focus at its corresponding reconstruction plane and appears blurred at other reconstruction planes. The objects are successfully seen through the occlusion that troubled the 2D image shown in Fig. 12(b).

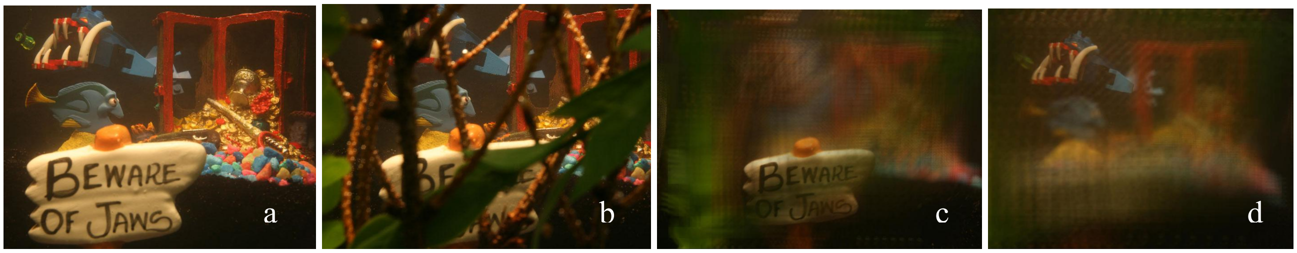


Fig. 12 Underwater Synthetic Aperture Integral Imaging example. a) A central 2D elemental image of the unoccluded underwater scene. b) Same scene as (a) with added foreground occlusion. 3D computational reconstructions of the occluded scene at c) z = 230mm, bringing the occluded foreground sign into focus, d) z = 440mm with jaws in focus

### 6.5 Sensitivity of Integral Imaging to position measurement uncertainty

We present the results an analysis on the sensitivity of 3D passive II with respect to the position measurement uncertainty during pickup [37]. SAII experiments are used to demonstrate the quantitative and qualitative degradation of computational 3D reconstructions by introducing sensor position uncertainty in the pickup process. The experimental scene is composed of two toy cars and a model helicopter located at 24cm, 32cm and 40cm from the sensor [see Fig. 13(a)]. The scene is illuminated with diffused incoherent light.

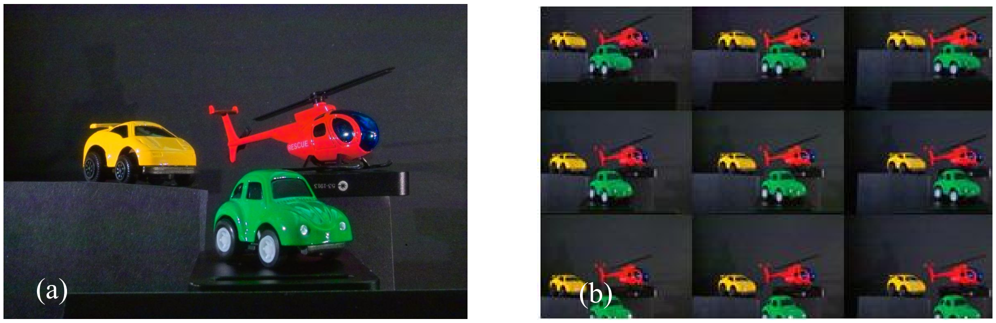


Fig. 13. (a) A 2D image of the 3D scene, (b) subset of elemental images for 3D scene in (a).

The experiment is performed by moving a digital camera transversally in an *x-y* grid with the pitch of $S_p$=5mm in both *x* and *y* directions. At each node, an elemental image is captured from the scene. The imaging sensor is 22.7×15.6mm and has a 10μm pixel pitch. Effective focal length of the camera lens is about 20mm; and elemental images are captured in a planar 16×16 grid. A subset of elemental images can be seen in Fig. 13(b) each conveying different perspective information. In Fig. 14 we show the 3D reconstruction of the scene in three different distances of the objects in Fig. 13(a) according to Eq. (1). As is clear, at each distance one of the objects is in focus while the others appear washed out.

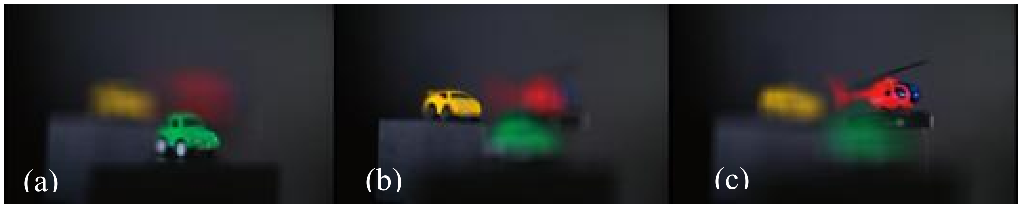


Fig. 14. 3D scene reconstruction at distances (a) z=*24 cm*, (b) *z=30 cm* and (c) *z=36 cm*.

It has been shown in [37] that Mean Square Error (MSE) can be calculated as:

$$E\left|err(x,y,z)\right|^2 = \frac{1}{R^4}E\left\{\left|\sum_{k}^{K-1}\sum_{l}^{L-1} I_{kl}(x,y) - I_{kl}(x+\frac{\Delta p_k}{M}, y+\frac{\Delta p_l}{M})\right|^2\right\}, \qquad (9)$$

where $E\{.\}$ denotes expectation operation, $I_{kl}$ is the kl-th elemental image; ($\Delta P_x$,$\Delta P_y$) are the random variables describing the sensor position error and are modeled as two independent Gaussians, $\Delta P_{x,y} \sim N(0,\sigma^2)$. We define the fraction $100\sigma^2/S_p$ to be the pitch error percentage that represents a normalized positioning error measure.

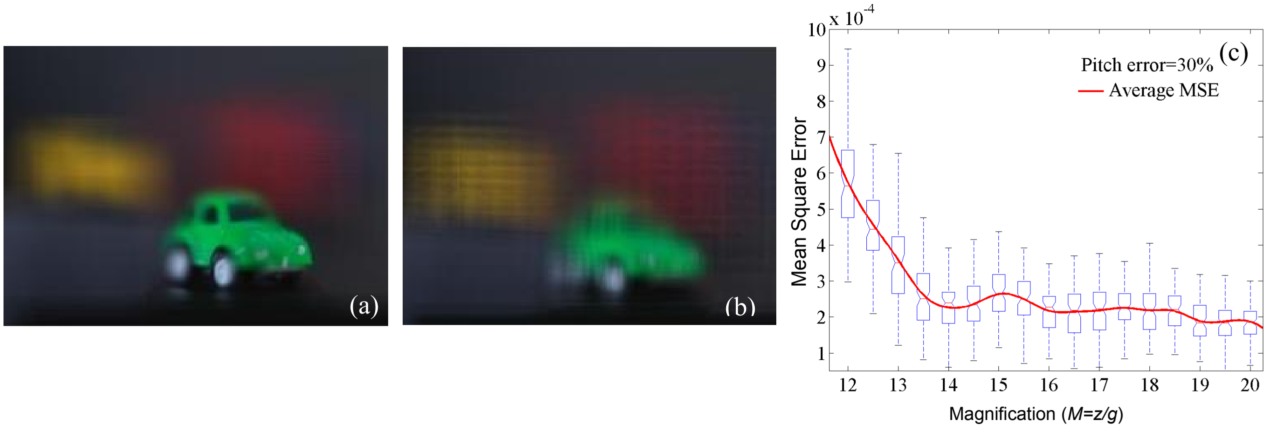


Fig. 15. Reconstruction at *z= 24 cm* using (a) original camera position, and (b) using distorted camera position with 30% pitch error. (c) Box-and-whisker diagram of the MSEs for *z*=24*cm (M*=12*)* to *z*=40*cm (M*=20*)* with pitch error of 30%

Monte Carlo simulation is used to study the degradation effect of sensor position uncertainty during the pickup process on the reconstructed images. We perform computational reconstruction with Eq. (1) to reconstruct a plane of the 3D scene at the specific distance *z* by utilizing the distorted camera positions. To measure the error, the MSE of the reconstruction results compared with the ones using the correct positions on the equally spaced grid is calculated. Figure 15(a)-(b) shows the result of reconstruction using known and random positions respectively at $z$=24cm with 30% pitch error. Dislocated position of the camera is a random vector from which we choose 500 samples and utilize them in Eq. (1) to computationally reconstruct one plane of the 3D scene at the distance *z*. As a result, 500 MSE values are obtained (via Eq. (9)) for each reconstruction distance. Figure 15(c) shows the box-and-whisker diagram of all MSEs for $z=24cm$ ($M$=12) to $z=40cm$ ($M$=20).

Figure 15(c) also shows the statistical properties of MSEs at a specific distance $z=z_0$ ($M=M_0$). The blue box shows the variance of the MSE which is limited to its upper and lower quartiles and dotted blue line is limited to the smallest and largest computed MSEs. At each plane, the average of 500 MSEs is computed and shown with the solid red line in Fig. 15(c). This average for each particular plane of the scene is a reasonable estimation of the error one can expect due to a 30% camera positioning error.

## 7 CONCLUSION

We have presented an overview of both active and passive 3D sensing techniques for ATR. We have highlighted a 3D sensing and imaging technique based on passive sensing, and its applications in ATR, under water 3D imaging, photon counting sensing ATR, and removal of partial obscurations for ATR. The benefits and weaknesses of both active and passive 3D sensing techniques have been discussed. For small, inexpensive platforms such as small UAVs, it is likely that 3D passive imaging will be preferable to active 3D imaging in terms of cost, size, and complexity.

## REFERENCES

[1] Sadjadi, F., Mahalanobis, A., “"Target-adaptive polarimetric synthetic aperture radar target discrimination using maximum average correlation height filters,” Applied Optics 45, 3063-3070 (2006)

[2] Refreigher, P., Laude, V., Javidi, B., “Nonlinear joint-transform correlation: an optimal solution for adaptive image discrimination and input noise robustness,” Opt. Lett. 19, 405-407 (1994).

[3] Mahalanobis, A, Van Nevel, A. “An Integrated approach for Automatic Target Recognition using a Network of Collaborative Sensors”, Applied Optics 45(28), 7365-7374 (2006).

[4] Javidi, B., Wang, J., “Optimum distortion-invariant filter for detecting a noisy distorted target in nonoverlapping background noise,” J. Opt. Soc. Am. A 12, 2604-2614 (1995)

[5] Sadjadi, F., “Improved target classification using optimum polarimetric SAR signatures,” IEEE Trans. on Aerosp. Electron. Syst. 38, 38-49 (2002).

[6] Mahalanobis, A., Muise, R., R., Stanfill S., R., and Nevel, A. V., “Design and application of quadratic correlation filters for target detection,” IEEE Trans. on Aerosp. Electron. Syst. 40, 837-850 (2004).

[7] Javidi, B., ed., “Image recognition and classification: algorithms, systems, and applications”, Marcel Dekker, New York, (2002).

[8] Javidi, B., ed., “Optical imaging sensors and systems for homeland security applications”, Springer, New York, (2006).

[9] Goodman, J. W., “Introduction to Fourier optics, 2$^{nd}$ edition”, McGraw-Hill, New York, (1996).

[10] Goodman, J. W., “Statistical optics”, Jonh Wiley & Sons Inc., (1985).

[11] Javidi, B., Ponce-Diaz, R., Hong, S.-H., “Three-Dimensional Recognition of Occluded Objects Using Volumetric Reconstruction,” Optics Letters, 31, 1106-1108 (2006)

[12] Hong, S., Javidi, B., “Detecting 3D Location and Shape of Noisy Distorted 3D Objects using LADAR Trained Optimum Nonlinear Filters,” Applied Optics 43(2), 324-332 (2004).

[13] Hong S., Javidi, B., “Optimum Nonlinear Composite Filter for Distortion Tolerant Pattern Recognition,” Journal of Applied Optics 41, 2172-2178, (2002).

[14] Yeom, S., K., and Javidi, B., “Three-dimensional Distortion-tolerant Object Recognition using Integral Imaging,” Optics Express 12, 5795-5809 (2004).

[15] Hong, S. -H., Javidi, B., “Distortion-tolerant 3D recognition of occluded objects using computational integral imaging,” Opt. Express 14, 12085-12095 (2006).

[16] Yeom, S., Javidi, B, and Watson, E., "Photon counting passive 3D image sensing for automatic target recognition," Opt. Express 13, 9310-9330 (2005).

[17] Yeom, S., Javidi, B., and Watson, E., "Three-dimensional distortion-tolerant object recognition using photon-counting integral imaging," Opt. Express, 15(4) (2007).

[18] Yeom, S., Javidi, B., and Watson, E., "Photon-counting passive 3D image sensing for reconstruction and recognition of occluded objects," Optics Express 15, 16189-16195 (2007).

[19] Yeom, S., Javidi, B., Watson, E., "Photon counting 3D passive sensing and recognition using computational integral imaging," Computational Optical Sensing and Imaging 2007, Vancouver BC, Canada (2007).

[20] Yeom, S., Javidi, B., and Watson, E., "Three-dimensional distortion-tolerant object classification using photon-counting linear discriminant analysis," Proc. of SPIE 6778, Boston, Massachusetts (2007).

[21] Benton, S. A., "Selected Papers on Three-Dimensional Displays", SPIE Optical Engineering Press, Bellingham, (2001).

[22] Arimoto, H., Javidi, B., "Integral three-dimensional imaging with digital reconstruction," Opt. Lett. 26, 157-159 (2001)

[23] Jang, J. -S., Javidi, B., "Three-dimensional synthetic aperture integral imaging," Opt. Lett. 27, 1144–1146 (2002).

[24] Hong, S.-H., Javidi, B., "Three-Dimensional Visualization of Partially Occluded Objects Using Integral Imaging", IEEE/OSA Journal of Display Technology 1, 354-359 ( 2005)

[25] Hwang, Y., S., Hong, S., Javidi, B., "Free view 3D visualization of occluded objects by using computational synthetic aperture integral imaging," IEEE Journal of Display Technology 3, (2007).

[26] Hong, S., and Javidi, B., "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," Optics Express 12, 4579-4588 (2004).

[27] Hong, S., Jang, J.-S., Javidi, B., "Three-dimensional volumetric object reconstruction using computational integral imaging," Optics Express, 12, 483-491 (2004),

[28] Schulein, R., T., and Javidi, B., "Underwater multi-view three-dimensional imaging," to appear in IEEE Journal of Display Technology (2008)

[29] Wilburn, B., Joshi, N., Vaish, V., Barth, A., Adams, A., Horowitz, M., Levoy, M., "High performance imaging using large camera arrays," Proc. of the ACM 24, 765-776 (2005).

[30] Ng, R., "Fourier slice photography," in Proc. of ACM SIGGRAPH, pp. 735–744 (2005).

[31] Levoy, M., Hanrahan, P., "Light field rendering," in Proc. of ACM SIGGRAPH, pp. 31–42, New Orleans (1996).

[32] Levoy, M., "Light fields and computional imaging," IEEE Computer**,** 46–55 (2006).

[33] Javidi, B., Okano, F., eds., "Three Dimensional Television, Video, and Display Technologies", Springer, Berlin,, (2002)

[34] Javidi, B., Hong, S., H., Matoba, O., "Multidimensional optical sensor and imaging system," Applied Optics, 45, pp. 2986-2994 (2006).

[35] Martìnez-Cuenca, R., Saavedra, G., Martínez-Corral, M., and Javidi, B., "Extended Depth-of-Field 3-D Display and Visualization by Combination of Amplitude-Modulated Microlenses and Deconvolution Tools," IEEE J. Display Tech. 1, 321–327 (2005).

[36] Stern, A., Javidi, B., "Three-dimensional image sensing and reconstruction with time-division multiplexed computational integral imaging" Appl. Opt. 42, 7036-7042 (2003).

[37] Tavakoli, B., DaneshPanah, M., Javidi, B., Watson, E., "Performance of 3D integral imaging with position uncertainty," Opt. Express 15, 11889-11902 (2007)

[38] Okoshi, T., "Three-dimensional imaging techniques", Academic Press, New York, (1976).

[39] Okano, F., Hoshino, H., Arai, J., and Yuyama, I., "Real-time pickup method for a three-dimensional image based on integral photography," Applied Optics 36, 1598-1603 (1997).

[40] Jang, J.-S., Javidi, B., "Improved viewing resolution of three-dimensional integral imaging by use of nonstationary micro-optics," Opt. Lett. 27, 324-326 (2002).

[41] Lippmann, G., "La photographic intergrale," C. R. Acad. Sci. 146, 446-451 (1908).

[42] Ives, H. E., "Optical properties of a Lipmann lenticulated sheet," J. Opt. Soc. Am. 21, 171-176 (1931).

[43] Watson, E., A., Morris, G., M., "Comparison of infrared upconversion methods for photon-limited imaging," J. Appl. Phys. 67, 6075-6084 (1990).