# Active 3D Imaging Based on White Light Scanning: From Single Sensor to Sensor Networks

Xiang PENG*[1,2], Hailong CHEN[1,2], Dong HE[2], Xiaoli LIU[1], Bo PENG[2], Nan CHENG[2]
[1] College of Optoelectronic Engineering, Shenzhen University, Shenzhen, China;
[2] Shenzhen ESUN Co. Ltd., Shenzhen, China

## Abstract

Active 3D imaging based on white light scanning is a sort of efficient techniques which can digitize objects existed in physical world, and therefore to acquire 3D images of test targets. A light coding with a series of structured light patterns, for example, fringe patterns are usually applied for such a 3D imaging modality. A single optical 3D sensor based on fringe projection or 3D optical sensor network composed of multiple node sensors can be regarded as kinds of optical instruments that works on principle of computational imaging. In order to reconstruct a 3D image it is necessary to decode the fringe pattern to get the phase map that acts as special marks encoding each point of test object or scene, resulting in accurate determination of point correspondence. Once the homologous pairs identified precisely a 3D image can be obtained through a series of computations using a stereoscopic vision theory. This paper will briefly reviewed the working principle of an optical digitizer based on the strategy of phase-aided active imaging (PAAI) and that of 3D sensor network composed of multiple PAAI sensors. In addition, many potential applications in 3D body scanning, face scanning, and heritage digitizing will also be addressed in this paper to show great potentials of 3D optical digitizer based on the PAAI strategy and PAAI-based optical measurement network. Both academic research at the Shenzhen University and commercialized products developed at the Shenzhen ESUN Co. Ltd. with number of patented technologies will be presented.

Keywords: 3d imaging, 3d modeling, single sensor, sensor network

## 1. Introduction

There are various 3D imaging modalities based on different physical principles, including passive and active ways to acquire 3D images. Techniques based on passive 3D imaging include shape from silhouette, stereoscopy, and photogrammetry while active 3D imaging include laser triangulation, structured light projection, and time-of-flight (TOF), etc. [1-3]. Here we focus our attention on active 3D imaging because of its inherent property of high accuracy in terms of shape measurement. In the category of active 3D imaging the structured light projection, including fringe or speckle pattern projection is a main-stream technology for digitizing objects existed in physical world, and therefore to acquire 3D images of test targets. A light coding with a series of structured light patterns, for example, fringe patterns are usually applied for such a 3D imaging modality. A single optical 3D sensor based on fringe projection or 3D optical sensor network composed of multiple node sensors [4] can be regarded as kinds of optical instruments that works on the principle of computational imaging. In order to reconstruct a 3D image it is necessary to decode the fringe pattern to get the phase map that acts as special marks encoding each point of test object or scene, resulting in accurate determination of point correspondence. Once the homologous pairs identified precisely a 3D image can be obtained through a series of computations using a stereoscopic vision theory. It has been proved in recent years that 3D imaging and modeling (3DIM) has a wide spectrum of practical applications ranging from industrial inspection, reverse engineering, art and cultural heritage, medicine/forensic imaging, 3D biometrics and security, medical imaging, entertainment, to virtual reality (VR) and/or augmented reality (AR) and many more. It can be foreseen and no doubt that 3DIM will generate much more profound impact on human daily life.

The rest of paper is organized as follows: Section 2 gives a brief review to working principle of an optical digitizer based on a strategy of phase-aided active imaging (PAAI) and that of 3D sensor network composed of multiple PAAI sensors. In order to obtain accurate shape measurement results we would especially emphasize on the importance of 3D sensor and sensor network calibrations, which will be described in Subsection 2.2 and 2.3. To cover the entire object surface or scene it is usually necessary to take range-images (2.5D) acquisitions from different perspectives.   In this case

---

* xpeng@szu.edu.cn; +86- 755- 26538548; www.opto.szu.edu.cn

those partial range images in local coordinate systems must be registered or aligned in a common coordinate system (world coordinate system) and this procedure is referred to as range image registration, which will be described in Subsection 2.3. Once registration is accomplished some redundant data exist in overlapped areas of multiple range images, which should be removed by a following procedure referred to as range-images merging or integration in order to obtain a non-redundant point cloud that represents a true 3D image of object surface. The procedures for range-image registration and integration are explained in Subsection 2.4. Section 3 introduces some techniques for 3D modeling based on a complete point cloud, which include geometric representations, some manipulations of 3D model and the creation of photorealistic 3D image and/or model. A series of commercialized products developed at the Shenzhen ESUN Co. Ltd. with patented technologies are presented in Section 4. The last Section summarizes and concludes the paper.

## 2. From 2D image to 2.5D range image

### 2.1. Camera model and calibration

It is a common sense that a two-dimensional (2D) image is a projection of three-dimensional (3D) object onto a detection plane, for example, a retina of human eye or a photo-sensor of CCD camera. It is obvious that two coordinate systems involve in here: the world coordinate system (WCS) and the camera coordinate system (CCS) as illustrated in Fig. 1.
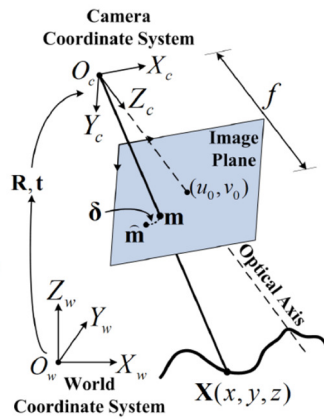


Fig. 1. 3D to 2D projection: image formation.

Mathematically, an image formation without consideration of lens distortions can be expressed by following matrix equation that is also referred to as a camera model or a pinhole model

$$s\mathbf{m} = \mathbf{K}[\mathbf{R} \mid \mathbf{t}]\mathbf{X} \quad \mathbf{X} \to \mathbf{m} : 3D \to 2D \tag{1}$$

where **K** is a matrix of internal parameters of the camera with an imaging lens and a photo-sensor inside. **R, t** are rotation matrix and translation vector representing external parameters of the camera, respectively. s is a scale factor denoting zooming effect according to the distance from the 3D object to imaging plane. **X** stands for a 3D point on the object surface while **m** is an image point of 3D point **X** in space. If we don't consider the lens distortion, Eq. (1) is linear equation that reflects the projection or mapping from 3D point **X** to 2D image point **m**. It is worthy pointing out that **R, t, K,** and s can be determined from a specifically designed procedure which is called camera calibration [5]. Once the internal and external parameters of the camera as well as scale factor are determined, the image is uniquely formed through the linear projection model expressed by Eq. (1). It is noted that the mapping from right side to left side is unique but vice versa does not hold because 2D image point **m** is degenerated (lower dimension) from 3D point **X** (higher dimension) to 2D image point. However, if we utilize two cameras: left camera and right camera (an analogy to human eyes) there are two set of equations (the same as Eq. (1)) available for uniquely reconstruct 3D point **X** under the framework of epipolar geometry**.** This is really a principle of stereoscopy! Nevertheless, the necessity to reconstruct 3D point **X** is that the homologous point pair $\mathbf{m}_r$ and $\mathbf{m}_l$ must be accurately determined, where $\mathbf{m}_r$ is an image point of **X** on the photo-sensor plane of right camera whereas $\mathbf{m}_l$ is that on the left camera sensor. Otherwise, it is still impossible to successfully reconstruct 3D point **X**. In the community of computer vision searching for homologous point pair is an open problem and by no means a trivial issue. It needs to be carefully considered.

On the other hand, due to various aberrations of imaging lens the image point there is a deviation between the position of ideal image point $\mathbf{m}$ and that of practically detected image point $\hat{\mathbf{m}}$ as shown in Fig.1 where $\boldsymbol{\delta}$ indicates non-linear aberrations caused by lens distortion. The shift from $\mathbf{m}$ to $\hat{\mathbf{m}}$ can be mathematically expressed as

$$\hat{\mathbf{m}} = \mathbf{m} + \delta(\mathbf{m}; \theta)$$

(2)

As mentioned above the camera calibration aims to estimate the internal and external parameters of underlying camera with a given calibration target as shown in Fig. 2 where the centroids of mark coordinates on the target surface have already known in advance, for example, they can be measured by Coordinate Measurement Machine (CMM), and the corresponding image points are also known via image detection of the target.
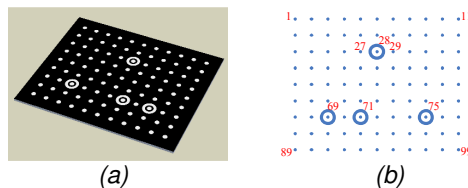


*Fig. 2. Calibration target (a) and reference point identification (b).*

With initial estimate $\mathbf{X}_W$ and $\mathbf{m}_c^d$ we are able to solve $\mathbf{R, t}$ and $\mathbf{K}$ with help of decomposition of eipipolar equation, see Ref. [5] for details. This process can be expressed as

$$\mathbf{m}_c^d, \mathbf{X}_w \implies \mathbf{K}, \mathbf{R}_c, \mathbf{t}_c$$

(3)

With $\mathbf{K}, \mathbf{R}_c, \mathbf{t}_c$ as initial estimates we can make further refinement by defining the residual error as the deviation between the estimated image point $\hat{\mathbf{m}}$ and practically detected image point $\mathbf{m}_c^d$. What we want is to make residual error minimized, which is a non-linear minimum problem and it can be solved by Levenberg-Marquet technique [5]

$$\min \sum_{ij} \left\| \mathbf{m}_{c,ij}^d - \hat{\mathbf{m}}(\mathbf{K}, \mathbf{R}_{c,i}, \mathbf{t}_{c,i}, \delta) \right\|_2$$

(4)

Where subscript c denotes the CCS involved, i indicates the ith pose of the camera whereas j the jth image point on the ith view. Superscript d implies the practically detected image point on the camera sensor plane. The minimization of Eq. (4) gives optimal estimates of parameter matrices and vectors: $\mathbf{K}, \mathbf{R}_{c,i}, \mathbf{t}_{c,i}, \delta$.

## 2.2. Formation of range image (2.5D) based on phase-aided active imaging (PSSI)

This subsection briefly reviews the working principle of an optical digitizer based on the strategy of phase-aided active imaging (PAAI) and that of 3D sensor network composed of multiple PAAI sensors. Schematic illustration of the PAAI is shown in Fig. 3 where three coordinate systems are currently involved in: the world coordinate system (WCS), the left camera coordinate system (l-CCS), and right camera coordinate system (r-CCS). $\mathbf{m}_r$ and $\mathbf{m}_l$ are homologous points originated from the same 3D point $\mathbf{X}_W$ in space. If one was able to figure out the homologous points through the image detection, he would be able to reconstruct 3D point $\mathbf{X}_W$. However, it is by no means an easy task in practical to do that since the object surfaces are featureless in many cases, leading to a common problem in the technique of stereovision. In order to overcome this difficulty, we add an additional optics unit for active projection that can generate and project structured light patterns, for example, a series of fringe patterns or speckle patterns, which act as spatial carriers. When the fringe patterns are projected onto the object surface, they are phase-modulated by the height variation of object surface, resulting in deformed fringe pattern(s), in which the information of surface height variation has been encoded. If we can decode the deformed fringe pattern(s), it is possible to reconstruct the height variation of object surface, which is a range image of object surface. On the other hand, if we utilize spatial phase values to encode each point on the underlying object surface, we are able to efficiently figure out homologous points in an accurate way and therefore to precisely reconstruct the range image of object surface. Mathematically we can formulate the range image formation with following matrices equations:
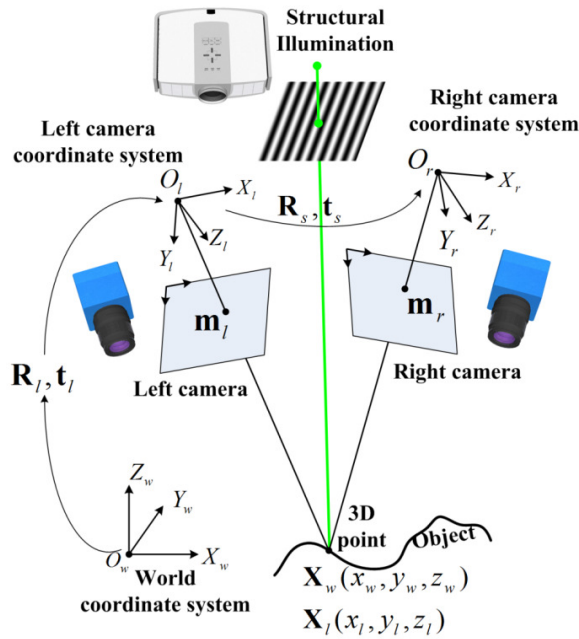
Fig. 3. Schematic illustration of PAAI layout.

$$\mathbf{m}_l \xleftrightarrow{\quad \varphi \quad} \mathbf{m}_r$$

$$\begin{cases} \mathbf{m}_l = \widehat{\mathbf{m}}_l - \delta\left(\mathbf{m}_l; \theta_l\right) \\ \mathbf{m}_r = \widehat{\mathbf{m}}_r - \delta\left(\mathbf{m}_r; \theta_r\right) \\ s_l \widetilde{\mathbf{m}}_l = \mathbf{K}_l[\mathbf{I} \mid \mathbf{0}]\widetilde{\mathbf{X}}_l \\ s_r \widetilde{\mathbf{m}}_r = \mathbf{K}_r[\mathbf{R}_s \mid \mathbf{t}_s]\widetilde{\mathbf{X}}_l \end{cases}$$

$$\begin{cases} \mathbf{R}_s = \mathbf{R}_r \mathbf{R}_l^{-1} \\ \mathbf{t}_s = \mathbf{t}_r - \mathbf{R}_r \mathbf{R}_l^{-1} \mathbf{t}_l \end{cases} \tag{5}$$

Where we have set the WCS onto the left camera. Subscript s indicates the structural relation between left and right cameras. As explained above $\mathbf{m}_r$ and $\mathbf{m}_l$ can be correlated with the aid of phase values of modulated spatial carrier(s). Mathematically original spatial carrier (fringe pattern) can be expressed as

$$I_n(\mathbf{m}) = a_0 + b_0 \cdot \cos\left[2\pi f_0 \mathbf{m} + 2n\pi/N\right], n = 0,1,\cdots,N-1 \tag{6}$$

Where vector $\mathbf{m}$ ($\mathbf{m} = (u,v)$) is an image point detected by camera(s), $f_0$ is spatial frequency of original carrier, $a_0$ is the background of the fringe pattern and $b_0$ is the contrast of fringe pattern. The term $2n\pi/N$ implies that N step phase shifts are introduced to demodulate deformed fringe pattern in order to extract phase information. The deformed fringe pattern (original carrier is modulated by surface height variation) can be expressed as

$$I_n(\mathbf{m}) = a(\mathbf{m}) + b(\mathbf{m}) \cdot \cos\left[2\pi f_0 \mathbf{m} + \varphi(\mathbf{m}) + 2n\pi/N\right], n = 0,1,\cdots,N-1 \tag{7}$$

In terms of optical metrology, the phase values $\varphi(\mathbf{m})$ can be extracted by using a strategy named phase-shifting interferometry (PSI) as described as following equation

$$\tan(\varphi'(\mathbf{m})) = \tan(\varphi(\mathbf{m})) = \frac{\sum_{n=0}^{N-1} I_n(\mathbf{m}) \sin(2n\pi / N)}{\sum_{n=0}^{N-1} I_n(\mathbf{m}) \cos(2n\pi / N)}$$

(8)

It is worthy to point out that the phase value extracted by using the PSI are wrapped with a module of $2\pi$ and they need to be unwrapped with the phase unwrapping algorithm(s). Here we utilize a modified temporal phase unwrapping algorithm to get unwrapped phase values, which can be summarized as following equation [6] (interested readers can refer to Ref [6] for more details)

$$\varphi_{t_2} = \varphi'_{t_2} - 2\pi \cdot NINT\left[\frac{\varphi'_{t_2} - k \cdot \varphi_{t_1}}{2\pi}\right], \ k = t_2/t_1$$

(9)

To this point, we have obtained a continuously distributed and unwrapped phase map whose values can be used to encode object surface, and therefore correlate the homologous points with identical phase values in horizontal (u) and vertical (v) directions, respectively. Searching for homologous points in this way can reach an accuracy with sub-pixel level [7]. This search process can be illustrated by Fig. 4.
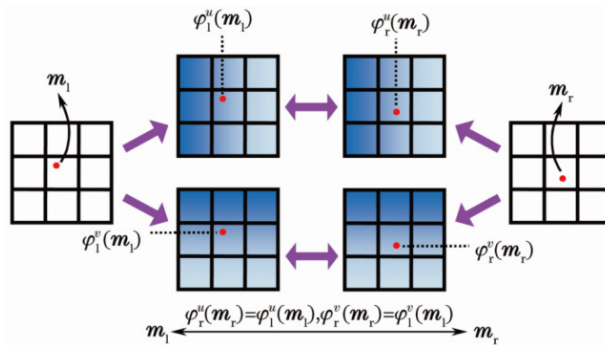


*Fig. 4. Illustration of searching for homologous points.*

As we discussed earlier one can reconstruct range image (2.5D) of object surface with homologous points and the decomposition of epipolar equation (decompositions of fundamental matrix and essential matrix lead to initial estimate of camera parameters) [5]. This process is also called Euclidian reconstruction in terms of multi-view geometry of computer vision.

In order to further improve the accuracy of initial estimate, we have recently proposed a new strategy used in the calibration procedure for PAAI, which makes use of the bundle adjustment (BA) technique [8] when optimizing the residual errors [9]. The comparison of residual errors of PAAI with and without using BA strategy in the optimization is shown in Fig. 5, which clearly shows dramatic reduction in residual errors, resulting in higher measurement accuracy. The error distribution described with blue color represents that the BA technique is applied whereas that with red color indicates the BA is not used. Dramatic error reduction can be observed and proved.
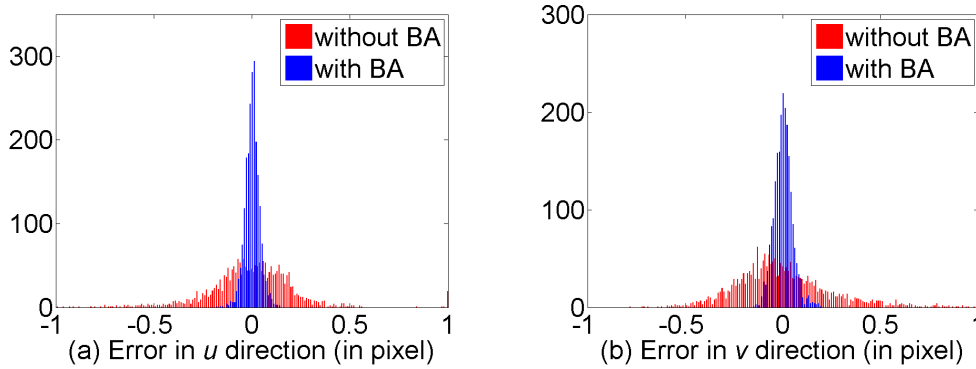
Fig. 5. Residual error distributions before and after using BA strategy.

Up to now we have already obtained a 2.5D image of underlying object surface with high accuracy from single perspective. The 3D sensor based on PAAI has been developed at the Shenzhen University and commercialized at Shenzhen ESUN Co. Ltd., whose a prototype is illustrated as Fig. 6.



Fig. 6. A prototype of 3D sensor based on PAAI.

### 2.3. Measurement Network composed of multiple 3D node sensors (MN-3D)

It should be noted that each 3D sensor based on PAAI can merely capture a range image of underlying object from a specific perspective (position and orientation of the camera) to form a 2.5D image. To cover entire object surface either moving the sensor around object or rotating object in front of the sensor in order to acquire a series of range images taken from different perspectives. These multi-view range images are captured in their local coordinate systems (frame coordinate systems). Our goal is to obtain a true 3D image instead of 2.5D image of underlying object, which can cover entire object surface without any shadows and occlusions. For this purpose, multiple range images taken from different perspectives must be aligned or registered in a common coordinate system (world coordinate system). This procedure is referred to as the registration of multiple range images in terms of computer graphics. Iterative Closest Point (ICP) based algorithms [10] is a sort of techniques for registering multiple range images. ICP based algorithms need an initial estimate for rough registration and then take a constraint optimization process to get fine registration. Original ICP algorithm [10] was under an assumption of rigid body transformation and initial registration was usually done by virtue of surface texture or specific features on the object surface through a human-machine interactions. In this sense the automation for multi-view range images registration has been a long-term effort to be made. Recently we proposed an efficient solution with a measurement network (MN-3D) composed of multiple 3D node sensors [11], whose topological structure was specifically designed according to the application at hand. An exemplary layout of MN-3D is illustrated in Fig. 7 where only three-node sensors are included, but it is easy to extend to include N node sensors in a straightforward way if necessary.
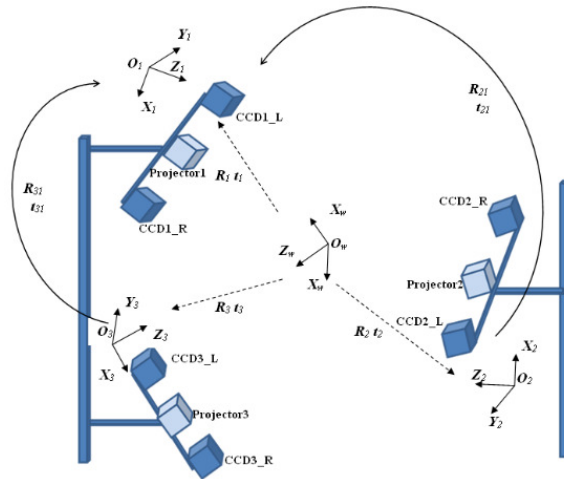
*Fig. 7. Schematic layout of a MN-3D with three node sensors.*

## 2.4. Calibration of MN-3D

Again, to enable the MN-3D work well it is necessary to be able to perform the calibration. However there are yet several difficulties exist when calibrating the MN-3D. First, if the size of underlying object is large and topology complex, the measurement volume becomes large. There is no suitable calibration target available on the market to calibrate such a large measurement volume. Second, the pose and orientation of each node sensor must be accurately determined simultaneously in order to make the MN-3D work. To address these issues we recently proposed an approach for the calibration of MN-3D [12], which first created a benchmark point set (virtual calibration reference points) with help of the technology of photogrammetry and then adopted a hierarchical strategy to calibrate the MN-3D in a successive way: from camera to 3D node sensor and then to the MN-3D. If the reader was interested in this technique, he would look into the Ref. [12] for more details. The result of MN-3D calibration is shown in Fig. 8 where red triangles represent camera poses while blue point set indicates the benchmark points to be used to calibrate the MN-3D.
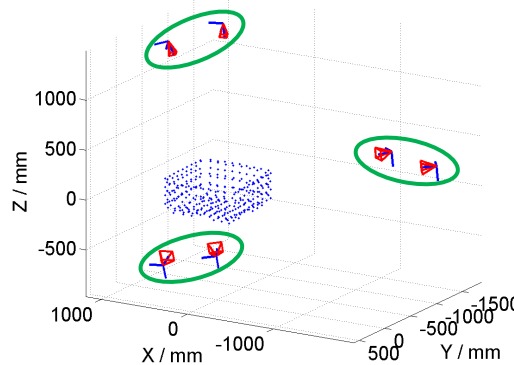


*Fig. 8. Calibration result of a MN-3D with three node sensors.*

## 2.5. From 2.5D to truly 3D

It is interesting to note that the position and orientation of each 3D node sensor is now determined after the calibration of MN-3D as described in subsection 2.4. These parameters provide fairly good initial estimate for the registration of multiple range images captured from each 3D node sensor. Then the ICP algorithm can be applied to optimize rough registration and gain a refined registration result. Fig. 9 is an example for the registration of multiple range images where a human face is used as test object.

The left and right are range images taken from two side-views of node sensors whereas the middle gives a result of registration of left and right range images. It is worth to note that there is no any marks needed to stick onto the object surface during the whole process of registration of multiple range images, which is preferable for many application scenarios such as a cultural heritage surface. It is absolutely not allowed to stick landmarks on such a surface in order to prevent from surface damage.

*Fig. 9. Registration result of a MN-3D with two node sensors.*

As mentioned in the Introduction section, once registration is accomplished some redundant data exist in overlapped areas of multiple range images, which should be removed by a following procedure referred to as range-image integration or merging in order to obtain a non-redundant point cloud that represents a true 3D image of object surface. We have developed two different approaches for merging multiple range images: an approach based on mesh stitching [14, 15] and approach based on the solution of general iso-surface equation [16], to which interested readers may refer for more technique details.

## 3. Three-dimensional modeling and creation of photorealistic 3D image

### 3.1. Three-dimensional (3D) modeling

3D point cloud is a set of 3D scattered data, in which there is no any connectivity among the data points. For such data points it is impossible to manipulate, for example, data compression and decompression. In order to manipulate 3D data set, it is necessary to represent 3D point cloud with topology or connectivity. In other words we need a geometric representation of those scattered 3D data points either with a mesh representation (triangle mesh or other polygon mesh) or with a parameterized representation (e.g. B-spline or isosurface or subdivision surface, etc.). Such a geometric representation is a 3D model that represents object surface with a mesh or a parametric surface, and this kind of 3D models can be manipulated such as mesh reduction (to obtain a simplified model) and refinement of a simplified model (see Refs. [14] - [16]) for more details. Fig. 10 give an example of 3D mesh model where 53,000 vertices and 106,000 facets are involved in. The data file for stereolithography (STL) is with a size of 5 Mb.
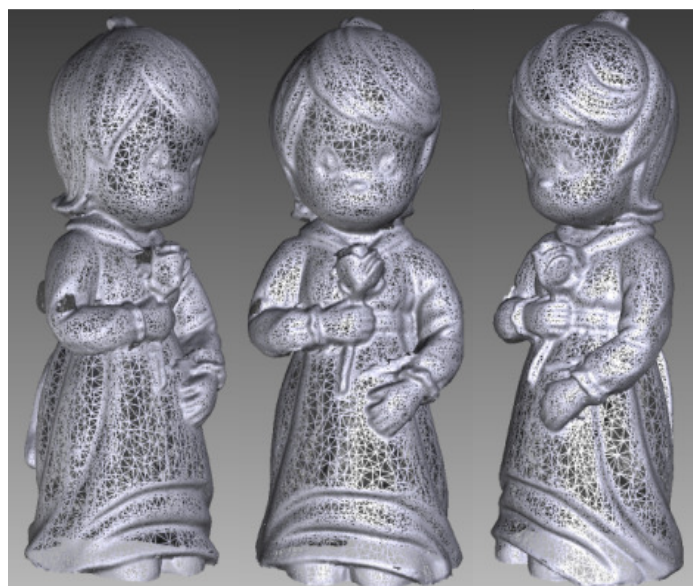


*Fig. 10. Example of 3D model with triangle mesh.*

Fig. 11 gives an example for mesh reduction and Fig. 12 shows an example of refinement from the simplified model of Fig. 11. For simplified mesh model there are 5,000 vertices and 11,000 facets are involved in comparison with original model. Now the size of STL data file is 500 Kb. It is really a compressed version in three-dimension! Another interesting thing to note is that major features of object are remained in this simplified model.
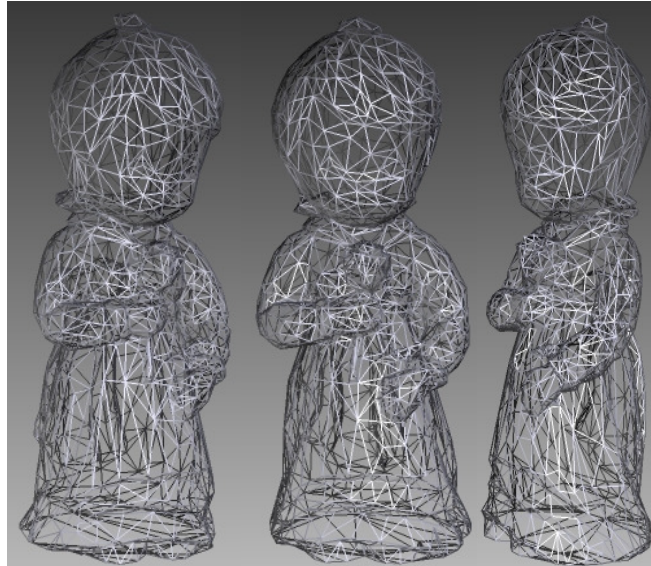


*Fig. 11. Example of simplified 3D model.*



*Fig. 12. Example of refinement of a simplified 3D model (Fig. 11).*

The simplified model (Fig. 11) can be refined to get back fine details by using, for example, subdivision technique. The refined model as shown in Fig.12 include 22,102 vertices and 44,208 facets. After refinement, the size of 3D data file becomes 2 Mb, where some of fine details of object have been come back.

The mesh reduction and simplified mode refinement will be very useful in terms of storage and transmission of 3D model. Otherwise, huge data storage is required and terrible bandwidth will be needed to transmit and handle so large 3D data files.

### 3.2. Creation of photorealistic 3D image

A photorealistic model of a real-life object or a scene has a great potential in a number of application areas such as game development, computer animation, film production, virtual reality, cultural heritage protection, digital content creation, and three-dimensional television (3DTV), to name just a few. In comparison with the generation of geometric models that have received much attentions in past decades, the issue of generating three-dimensional (3D) textured models or photorealistic 3D images becomes increasingly interesting in recent years with the advancement of various 3D optical digitizers that bridge real life to a virtual world [17, 18]. Furthermore, the dimension and shape of an object can be represented through a geometric model whereas the natural appearance of the object is represented with color texture. Therefore, a color-textured model with photorealistic quality not only needs precise shape information but also the information of the natural appearance.

The generation of a textured model requires three crucial steps: (1) calibration of 3D sensor, (2) acquisition of range images and textured images, and (3) generation of a photorealistic mode. Because the cameras are basic components of 3D optical sensor and/or measurement network with multiple 3D sensors (MN-3D), the camera calibration is deemed essential in the system calibration of 3D optical sensor and sensor network, i.e. MN-3D. Calibration is necessary not only to get the coordinates of digitized surface but also to set proper correspondences between these coordinates and the texture captured by a separate color camera. In addition to range images acquisition an extra color camera attached to 3D sensor is used to acquire texture images associated with range images at each viewpoint. The generation of a photorealistic image/model involves in several considerations, including the projective transformations from range image to texture image and environmental lighting. One approach for texture blending is to weight the colors of texture pixels according to some specific rules; others have adopted image mosaic technology and achieved texture blending by removing the uncontinuity of mosaic patches. Recently we proposed alternative approach to generate an accurate and photorealistic model [18 ]. In this approach, we acquire the range images and associated texture images from different perspectives with dedicated 3D optical sensor. For more technical details, the readers may refer to Ref [18] and references therein. Here we adopt the same strategy on MN-3D where the cameras, 3D node sensors, and whole MN-3D have been accurately calibrated in a hierarchical way as described in Subsection 2.4. In current situation, the registration can be implemented automatically and texture blending can be done in a very efficient way. Some of photorealistic models generated by our approach will be given in next Section.

## 4. Commercialization at Shenzhen ESUN Co. Ltd.

In this Section, three products of optical digitizers based on patented PAAI techniques developed at Shenzhen University and technically transferred to the ESUN Co. Ltd. are introduced. The PAAI single sensor and measurement network with multiple 3D node sensors are commercialized at the Shenzhen ESUN Co. Ltd.[19]. These commercial products include the desktop 3D color scanner, the 3D face camera, and the full-body scanner.

### 4.1. Desktop 3D color scanner (ESUN 3D+™)

Figure 13 shows a commercial product of desktop 3D color scanner manufactured by ESUN.



*Fig. 13. Commercial product of desktop 3D color scanner.*

In this instrument, there are two 3D node sensors and motorized rotation stage involved in. The object to be scanned is placed on the stage and two node sensors take the range image from different perspectives while the stage is rotated under automatic control, leading to a complete 3D point cloud and 3D model. It should be noted that the registration of multiple range images is accomplished automatically. Fig. 14 shows an example of photorealistic model of "turtle immortal", which is observed from different perspectives as shown in Fig. 14 (a)-(e).
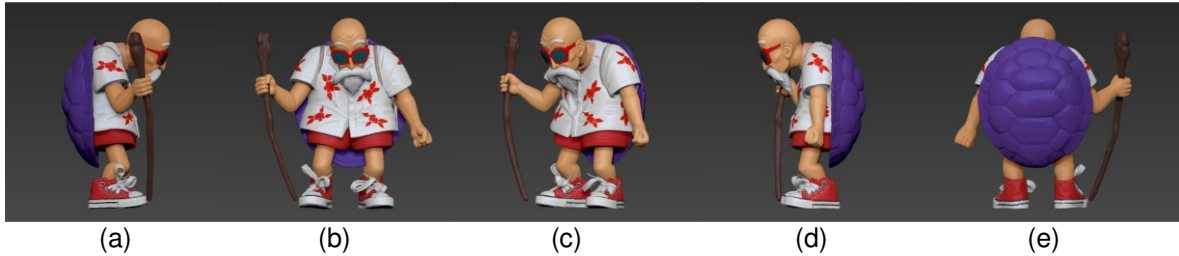


|        (a)        |        (b)        |        (c)        |        (d)        |        (e)        |

*Fig. 14. "turtle immortal" observed from different perspectives.*

### 4.2. 3D face camera (ESUN 3D+<sup>TM</sup>)

**4.2. 3D face camera (ESUN 3D+$^{TM}$)**

Figure 14 shows a commercial product of 3D face camera designed and manufactured by ESUN. Fig. 14 (a) schematically illustrates the design of a prototype of 3D face camera while (b) is the illustration of manufactured instrument.
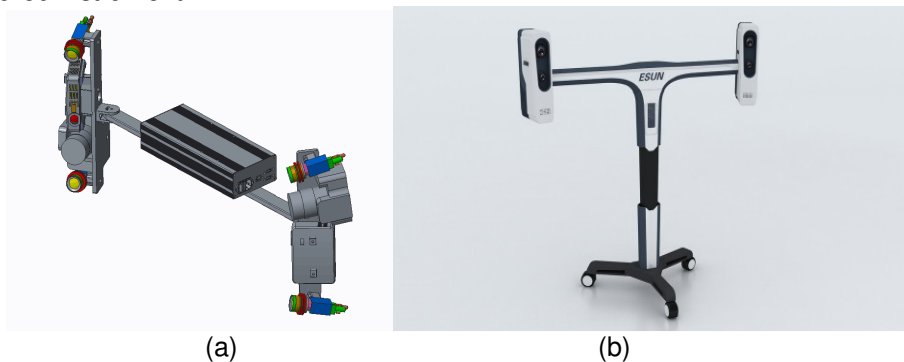


|                (a)                |                (b)                |

*Fig. 15. 3D face camera (ESUN 3D+$^{TM}$): (a) design of prototype, (b) instrument manufactured.*

Fig. 16 (a)-(c) show a 3D image of the face of a real person.



*Fig. 15. 3D face imaging of a real person.*

### 4.3. Full-body scanner (ESUN 3D+$^{TM}$)

Figure 16 shows a commercial product of full-body scanner designed and manufactured by ESUN. This product is composed of eight 3D node sensors installed on four stand rods respectively. The acquisition of multiple range images is synchronized by software and implemented in a pipeline fashion.
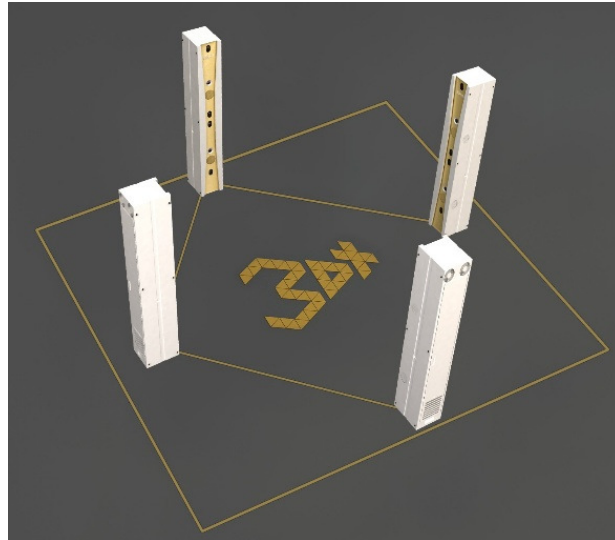
*Fig. 16. Full-body scanner designed and manufactured (ESUN 3D+$^{TM}$).*



*Fig. 17. 3D images/models of human body (ESUN 3D+$^{TM}$).*

Figure 17 illustrates the 3D images/model of a human body observed from different perspectives.

## 5. Conclusion

3D imaging and modeling based on phase aided active stereovision has been reviewed, which includes the concepts and methodologies of the camera model and calibration, the formation of 2.5D range image, the calibration of range finder, the measurement network with multiple 3D node sensors (MN-3D), the calibration of MN-3D, the registration of multiple range images, the integration of multiple range images, the manipulations of 3D model including simplification and refinement, the generation of a photorealistic 3D image/model. Several case studies have been given to validate our approach. In addition, the commercialization of those technologies at Shenzhen ESUN Co. Ltd. are also introduced with a number of illustrations. We believe that 3D imaging and modeling techniques will continuously generate more profound impact on human life and the investigations on both academic circle and industrial sectors will keep going in succession. More exciting advances are expected, for example, 3D imaging on mobile devices is a very promising research and development direction.

(A revised version of this paper has been submitted to OSA journal for publication)

## Acknowledgments

## References

[1] F. Blais, "Review of 20 years of range sensor development." Journal of Electronic Imaging, 2004, 13(1): 231-240.

[2] S. Gorthi, P. Rastogi, "Fringe projection techniques: Whither we are?" Optics and Lasers in Engineering. 2010, 48(2):133-140.

[3] G. Sansoni, M. Trebeschi, F. Docchio, "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation." Sensors, 2009, 9(1): 568-601.

[4] X. Peng, X. Liu, Y. Yin, A. Li, "Optical measurement network for large-scale and shell-like objects," Opt. Lett. 36(2), 157-159 (2011).

[5] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. (Cambrige University, 2003).

[6] X. Peng, Z. L. Yang, H. B. Niu, "Multi-resolution reconstruction of 3-D image with modified temporal unwrapping algorithm," Optics Communications, 224, 35–44 (2003).

[7] S. Dong, X. Zhao, Y. Yin, J. Tian, X. Peng, "Sub-Pixel Matching with Consideration of Lens Distortion." Proc. SPIE 7155, 2008: 71552H.

[8] B. Triggs, P. McLauchlan, R. Hartley, W. Fitzgibbon, "Bundle Adjustment - A Modern Synthesis," in Proceedings of the International Workshop on Vision Algorithms: Theory and Practice, (ICCV, 2000), pp.1883: 298-372.

[9] Y. Yin, X. Peng, A. Li, X. Liu and Z. Gao, "Calibration of fringe projection profilometry with bundle adjustment strategy," Optics Letters, 37, 542-544 (2012).

[10] P. J. Besl, N. D. McKay, "A method for registration of 3D shapes, in Proceedings of IEEE Conference on Transactions on Pattern Analysis and Machine Intelligence," (IEEE, 1992), pp.239-256.

[11] X. Peng, X. L. Liu, Y. Yin, and A. Li, "Optical measurement network for large-scale and shell-like objects," Optics Letters, 36 (2), 157-160 (2011).

[12] Y. Yin, X. Peng, X. Liu, A. Li, X. Qu, "Calibration strategy of optical measurement network for large-scale and shell-like objects," Opt. Commun. 285(8), 2048-2056 (2012).

[13] X. Liu, X. Peng, Y. Yin, A. Li and X. Zhao, A method for global registration of range data combined with markers, ACTA OPTICA SINICA 29(4),1010-1014 (2009).

[14] Z. Zhang, X. Peng and D. Zhang: "Transformation Image into Graphics," in Book: *Integrated Image and Graphics Technologies*, Chapter 7, 111-129, edited by David Zhang, Mohamed Kamel and George Baciu, Kluwer Academic Publishers, USA. 2004.

[15] X. Peng and J. Tian, "Three-Dimensional Multi-resolution Imaging and Modeling", in Book: *New Development in Laser and Electro-Optics Research*, Chapter 4, Edited by William T. Arkin, Nova Science Publisher, USA, 2006.

[16] Z. Liu, Y. Ding, X. Peng. "Generalized isosurface equation and its application to multi-view range images integration." Optik, 2004, 115(2): 71-76.

[17] I. Stamos, L. Liu, C. Chen, G. Wolberg, G. Yu, and S. Zokai, "Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes," Int. J. Comput. Vis.78, 237–260 (2008).

[18] X. M. Liu, X. Peng, Y. K. Yin, X. L. Liu, and W. Wu, "Generation of photorealistic 3D image using optical digitizer," Appl. Opt., 51, 1304-1311(2012).

[19] www.es-display.com

[20] www.bellus3D.com