doi:10.3788/gzxb20134212.1421

用于可移动文物真实感成像的光学三维数字化仪

李阿蒙,彭翔,殷永凯,刘晓利

(深圳大学 光电工程学院 光电子器件与系统教育部/广东省重点实验室,广东 深圳 518060)

摘 要:计算机辅助光学三维数字化技术作为一种有力的工具,可应用于文物三维原真信息的获取、存储和修复.本文提出一种获取可移动文物高准确度彩色三维数字成像法,包括基于光束平差的标定技术和相位辅助的主动三维成像技术.报道了可移动文物三维成像与建模的方法构建和实现过程,并利用该方法对典型薄壁结构的青铜器进行实验研究,验证了所提出真实感三维成像方法的有效性,建立了一种用于可移动文物三维成像的仪器原型.

关键词:条纹投影轮廓术;真实感三维成像;计算机辅助光学测量;可移动文物;纹理融合

中图分类号:TP391 **文献标识码**:A **文章编号**:1004-4213(2013)12-1421-9

Optical 3D Digitizer for Photorealistic Imaging of Movable Cultural Heritage

LI A-meng, PENG Xiang, YIN Yong-kai, LIU Xiao-li

(*Key Laboratory of Optoelectronic Devices and Systems, Education Ministry of China and Guangdong Province, College Of Optoelectronic Engineering, Shenzhen University, Shenzhen, Guangdong* 518060, *China*)

Abstract: Computer-aided optical digitization technique appears to be a powerful tool for the generation, storage, and retrieval of three-dimensional digital information of cultural heritage, and therefore leading to a new generation of technology for digital preservation. In this paper, an efficient approach was presented for generating photorealistic three-dimensional image and model of movable cultural heritage. A dedicated color optical digitizer was also presented based on a strategy of modified fringe projection profilometry. The approach provided an efficient way to digitize and reconstruct photorealistic three-dimensional images and digital models of movable cultural heritage. Experiment results using bronze wares with thin-wall shape structures and complicated topologies as well as color attributes were given to demonstrate the feasibility of proposed method for digital preservation.

Key words: Fringe projection profilometry; Photorealistic three-dimensional imaging; Computer-aided optical metrology; Movable cultural heritage; Texture blending

0 Introduction

Cultural heritage is the legacy of physical artifacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present and bestowed for benefit of future generations. The movable cultural heritage is usually referred to as a class of artifacts such as bronze ware, pottery, stoneware, gold and silver ware, painting and other historical artworks as well. The deliberate act of keeping cultural heritage from the present for the future is known as Preservation or Conservation. Cultural heritage is unique and irreplaceable, which places the

Foundation item: The National Natural Science Foundation of China (No. 61201355) and Joint Research Project supported by the Sino-German Center for Research Promotion (No. GZ760)

First author: LI A-meng(1982—), male, lecturer, M. S. degree, mainly focuses on three-dimensional imaging, optical metrology, and computer vision. Email:catameng@163.com

Supervisor(Corresponding author): Peng Xiang(1955—), male, professor, Ph. D. degree, mainly focuses on optical imaging and metrology. Email:xpeng@szu.edu.cn

Received: Apr. 23, 2013; **Accepted**: Jun. 28, 2013

responsibility of preservation on the current generation. Digital preservation is the set of processes, activities and management of digital information over time to ensure its long term accessibility. In terms of digital imaging and electronic resources, preservation is no longer just the product of a program but an ongoing process. In this regard the way digital information is stored is important in ensuring its longevity. Digital preservation is defined as: long-term, error-free storage of digital information, with means for retrieval and interpretation, for the entire time span the information is required[1]. In recent years, an emerging field referred to as optics for art, architecture, and archeology has appeared, aiming to introduce modern information technology to the field of cultural heritage[2-3]. In particular, optical Three-Dimensional (3D) imaging technique has shown a great potential for generating high accurate and high resolution digital model of artifacts[4-6]. A number of efforts have been made to virtual heritage where 3D optical imaging has served as a powerful tool for digitizing the artworks and virtual reality has also been incorporated for generating augmented showcase[7-10]. Even though optical digitization techniques appear to be very promising to create 3D digital content of artifacts, it still remains a challenge for a large number of artifacts to be digitized. Among movable cultural heritage, bronze ware is one of typical examples hard to be digitized because of complicated geometric structures, topologies, and color distributions. Many of bronze wares have thin-wall shape structures and large sizes, resulting in difficulties for currently available optical imaging and digitization techniques.

In this paper, we report a new approach to generate photorealistic 3D models of movable cultural heritage by using a specifically designed optical digitizer. The rest of this paper is organized as follows. Section 2 briefly introduces the principle of optical color digitization of a specifically designed optical digitizer. The dedicated optical digitizer utilizes two cameras, one is Black-and-White (BW) and the other is a color digital camera. Both camera components of optical sensor work together with a Digital Micromirror Device (DMD) projector component, leading to a Modified Fringe Projection Profilometry (m-FPP) for range image and color image acquisition of artifacts. Once calibrated, the colored range images around the test object can be obtained with high geometric accuracy and acquired color images can be aligned with associated range images precisely. Furthermore, Section 3 discusses the techniques of alignment and fusion of multiple of range images as well as geometric representation of 3D point cloud of range images acquired by colored optical 3D digitizer. Section 4 suggests an efficient method for texture mapping and texture blending in order to a photorealistic 3D model of cultural heritage. The experimental verifications using two kinds of bronze wares are presented in Section 5, showing the feasibility of proposed approach. Finally, the Section 6 concludes major points in our current study.

1 Principle of colored optical 3D digitizer

The principle of colored optical digitizer specifically designed her is based on m-FPP[11-13]. The coordinate systems of dedicated optical digitizer is shown in Fig. 1 where the DMD projector providing structural illuminations works with a B/W camera as Raw Camera (RC), to form a FPP scheme in order to acquire the range images. Another digital color camera as Texture Camera (TC), are attached to this m-FPP sensor head and used to capture associated color images of the art artifacts. For optical sensor, the DMD projector and BW camera together with texture camera TC result in m-FPP for creating partial range images as well as associated color images of artifacts to be digitized. The optical sensor can be formulated as follows. When the DMD is working, rays departing from a point $\boldsymbol{m}_p$ on the DMD chip

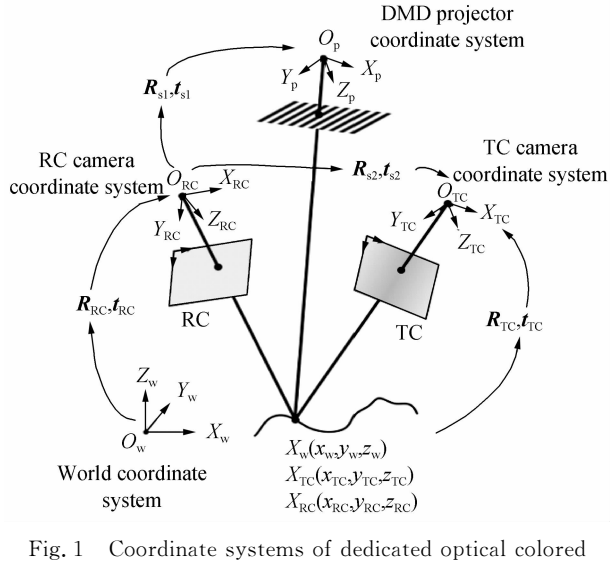

Fig. 1 Coordinate systems of dedicated optical colored 3D digitizer

plane pass through the projector lens and then irradiate the point $\boldsymbol{X}$, just like an imaging inverse process[14]. Furthermore, the exact coordinates of $\boldsymbol{m}_{\mathrm{p}}$ can be determined with the help of the camera. First, the projector sequentially projects two orthogonal sets of frequency-varying and phase-shifting sinusoidal fringe patterns onto the object and then the camera acquires those fringe patterns modulated by the object surface.

Using the phase-shifting and temporal phase unwrapping technique[15], one is able to accurately calculate the dual absolute phase maps in two orthogonal directions. For B/W camera, the absolute phase maps are denoted by $\boldsymbol{\varphi}_{\mathrm{c}}^{u}$ and $\boldsymbol{\varphi}_{\mathrm{c}}^{v}$, respectively and for projector by $\boldsymbol{\varphi}_{\mathrm{p}}^{u}$ and $\boldsymbol{\varphi}_{\mathrm{p}}^{v}$ accordingly. If $\boldsymbol{m}_{\mathrm{c}}$ is an image point of $\boldsymbol{X}$, the absolute phase values of $\boldsymbol{X}$ must be $(\boldsymbol{\varphi}_{\mathrm{c}}^{u}(\boldsymbol{m}_{\mathrm{c}}),\boldsymbol{\varphi}_{\mathrm{c}}^{v}(\boldsymbol{m}_{\mathrm{c}}))$. By searching the phase maps for the projector retina plane, if there exists a point $\boldsymbol{m}_{\mathrm{p}}$ satisfying the following conditions

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{p}}^{u}(\boldsymbol{m}_{\mathrm{p}})&=\boldsymbol{\varphi}_{\mathrm{c}}^{u}(\boldsymbol{m}_{\mathrm{c}})\\ \boldsymbol{\varphi}_{\mathrm{p}}^{v}(\boldsymbol{m}_{\mathrm{p}})&=\boldsymbol{\varphi}_{\mathrm{c}}^{v}(\boldsymbol{m}_{\mathrm{c}})\end{aligned}\tag{1}$$

The point $\boldsymbol{m}_{\mathrm{p}}$ must be the homologous point of $\boldsymbol{m}_{\mathrm{c}}$. Points pairs $\boldsymbol{m}_{\mathrm{c}}\leftrightarrow\boldsymbol{m}_{\mathrm{p}}$ correspond to the same object point $\boldsymbol{X}$, thus the projector can be modeled as an 'inverse camera', just by changing the subscript c into p. For each node sensor, the B/W camera and the projector are fixed with each other so that the structure parameters $\boldsymbol{R}_{\mathrm{s1}}$, $\boldsymbol{t}_{\mathrm{s1}}$ can be introduced to represent the rigid transformation between the B/W camera and the projector, which are defined by Eq. (2)

$$\begin{cases}\boldsymbol{R}_{\mathrm{s1}}=\boldsymbol{R}_{\mathrm{p}}\boldsymbol{R}_{\mathrm{RC}}^{-1}\\ \boldsymbol{t}_{\mathrm{s1}}=\boldsymbol{t}_{\mathrm{p}}-\boldsymbol{R}_{\mathrm{p}}\boldsymbol{R}_{\mathrm{RC}}^{-1}\boldsymbol{t}_{\mathrm{RC}}\end{cases}\tag{2}$$

Up to now we can model optical sensor with a principle of phase-aided active stereo as following equations:

$$\begin{cases}\boldsymbol{X}_{\mathrm{RC}}=\boldsymbol{R}_{\mathrm{RC}}\boldsymbol{X}_{\mathrm{w}}+\boldsymbol{t}_{\mathrm{RC}}\\ s_{\mathrm{RC}}\tilde{\boldsymbol{m}}'_{\mathrm{RC}}=\boldsymbol{K}_{\mathrm{RC}}[\boldsymbol{I}\,|\,\boldsymbol{0}]\tilde{\boldsymbol{X}}_{\mathrm{RC}}\\ \boldsymbol{m}_{\mathrm{RC}}=\boldsymbol{m}'_{\mathrm{RC}}+\delta(\boldsymbol{k}_{\mathrm{RC}};\boldsymbol{m}'_{\mathrm{RC}})\\ s_{\mathrm{p}}\,\tilde{\boldsymbol{m}}'_{\mathrm{p}}=\boldsymbol{K}_{\mathrm{p}}[\boldsymbol{R}_{\mathrm{s1}}\,|\,\boldsymbol{t}_{\mathrm{s1}}]\tilde{\boldsymbol{X}}_{\mathrm{RC}}\\ \boldsymbol{m}_{\mathrm{p}}=\boldsymbol{m}'_{\mathrm{p}}+\delta(\boldsymbol{k}_{\mathrm{p}};\boldsymbol{m}'_{\mathrm{p}})\end{cases}\tag{3}$$

where $\tilde{\cdot}$ denotes homogeneous coordinate, $\boldsymbol{m}'_{\mathrm{RC}}$ the ideal non-distortion image point on the B/W camera, $\boldsymbol{K}_{\mathrm{RC}}$ the intrinsic parameters matrix of B/W camera, $\boldsymbol{R}_{\mathrm{RC}}$ the rotation matrix, $\boldsymbol{t}_{\mathrm{RC}}$ the translation vector, s_{RC} a scale factor and $\delta(\boldsymbol{k}_{\mathrm{RC}};\cdot)$ the lens distortion parameterized by the distortion coefficients $\boldsymbol{k}_{\mathrm{RC}}$. Once the intrinsic and extrinsic parameters $\boldsymbol{K}_{\mathrm{RC}}$, $\boldsymbol{R}_{\mathrm{RC}}$, $\boldsymbol{t}_{\mathrm{RC}}$, s_{RC}, $\boldsymbol{k}_{\mathrm{RC}}$ are determined from a procedure of optical sensor calibration, the range image $\boldsymbol{X}_{\mathrm{w}}$ can be accurately reconstructed from the combination of Eq. (2) and Eq. (3). The readers who are interested in the calibration procedure of 3D optical digitizer can find details in another recent publication[16].

The natural appearance (color attribute) of real object can be captured by TC component of dedicated optical digitizer with controlled illumination condition. Since the photos taken follow the law of the perspective, if we know the camera parameters through the system calibration, we can determine if (and where) a point on the object surface is mapped within the overlapped regions or inside the image boundaries. An efficient way to compute those pixel-to-surface correspondences is by rendering the 3D model (range images) with the same projection parameters of the given photo. In this way, the range image (depth map) can be used to discriminate whether or not a point on the surface that projects inside the color image is visible or whether it is occluded by other geometry. In this way, it is possible to safely assign a pixel color, taken form that photo, to a point onto the surface. Digital color camera and B/W camera are also fixed with each other so that the rigid transformation between digital color camera component and B/W camera component of optical sensor head can also be expressed in following equations

$$\begin{cases}\boldsymbol{X}_{\mathrm{RC}}=\boldsymbol{R}_{\mathrm{RC}}\boldsymbol{X}_{\mathrm{w}}+\boldsymbol{t}_{\mathrm{RC}}\\ s_{\mathrm{TC}}\cdot\boldsymbol{m}'_{\mathrm{TC}}=\boldsymbol{K}_{\mathrm{TC}}[\boldsymbol{R}_{s2}\,|\,\boldsymbol{t}_{s2}]\boldsymbol{X}_{\mathrm{RC}}\\ \boldsymbol{m}_{\mathrm{TC}}=\boldsymbol{m}'_{\mathrm{TC}}+\delta(\boldsymbol{k}_{\mathrm{TC}};\boldsymbol{m}'_{\mathrm{TC}})\end{cases}\tag{4}$$

where

$$\begin{cases}\boldsymbol{R}_{s2}=\boldsymbol{R}_{\mathrm{TC}}\boldsymbol{R}_{\mathrm{RC}}^{-1}\\ \boldsymbol{t}_{s2}=\boldsymbol{t}_{\mathrm{TC}}-\boldsymbol{R}_{\mathrm{TC}}\boldsymbol{R}_{\mathrm{RC}}^{-1}\boldsymbol{t}_{\mathrm{RC}}\end{cases}\tag{5}$$

2 Alignment and fusion of multiple colored-range-images

Once the configuration shown in Fig. 1 is calibrated, it can be used to acquire the range images and associated color images of 3D artifacts in a real scene. Furthermore it is able to safely assign a pixel color, taken form that photo, to a point onto the surface so that the texture mapping can be achieved. The acquisition of multiple range images and associated color images is required in order to cover the entire object surface. For the k^{th} viewpoint ($k=1,2,\cdots,K$) the color image can be captured directly with the TC. Suppose that K numbers of range images and associated color images have been acquired from K viewpoints, we

therefore obtain a set of range images $\{\boldsymbol{X}_{\mathrm{RC}k}(x,y,z),k=1,2,\cdots,K\}$ and associated color images $\{\boldsymbol{I}_k(R,G,B),k=1,2,\cdots,K\}$. As the range images and their associated color images have been aligned using the method described in Section 2, the correspondence between $\{\boldsymbol{X}_{\mathrm{RC},k}(x,y,z),k=1,2,\cdots,K\}$ and $\{\boldsymbol{I}_k(R,G,B),k=1,2,\cdots,K\}$ is now established accurately. Based on these data sets we are able to start building a colored geometric model of 3D objects in a real scene.

In order to get a complete model it needs to align partial range images obtained from different viewpoints in the world coordinate system, this registration process can be mathematically described as follows

$$\begin{cases}\boldsymbol{X}_{\mathrm{W}_1}=\boldsymbol{R}_{\mathrm{RC}_1}\boldsymbol{X}_{\mathrm{RC}_1}+\boldsymbol{t}_{\mathrm{RC}_1}\\ \vdots\\ \boldsymbol{X}_{\mathrm{W}_k}=\boldsymbol{R}_{\mathrm{RC}_k}\boldsymbol{X}_{\mathrm{RC}_k}+\boldsymbol{t}_{\mathrm{RC}_k}\quad(k=1,\cdots,K)\\ \vdots\\ \boldsymbol{X}_{\mathrm{W}_M}=\boldsymbol{R}_{\mathrm{RC}_M}\boldsymbol{X}_{\mathrm{RC}_M}+\boldsymbol{t}_{\mathrm{RC}_M}\end{cases}\tag{6}$$

where $\boldsymbol{X}$ denotes range image points in 3D space and the subscript RC indicates the coordinate system of B/W camera. $\boldsymbol{R}_{\mathrm{RC}_k}$ and $\boldsymbol{t}_{\mathrm{RC}_k}$ represent the transformations from RC coordinate system to world coordinate system for k^{th} viewpoint. In our approach, the initial transformations of local range images into global coordinate system are determined by the calibration process. Then the coarse registration with initial estimate of sensor poses is further refined by a procedure of an Iterative Closest Point algorithm (ICP)[17]. It is worth to point out that the registered 3D image contains redundant data within the overlapped region. Therefore a further integration procedure is required to remove the redundancy and to obtain a complete and non-redundant data point-cloud. There are different approaches to integrate multiple range data sets, with no clearly superior approach so far. A key issue is how to manage the fusion of data in overlapped region. In our approach the integration of multiple range images has been accomplished by a stitching-based technique[18].

3 Geometric representation of 3D colored point cloud

The conversion of integrated multiple colored range-images into a mesh model can be symbolically described as follows

$$\{\boldsymbol{X}_{\omega k}(x,y,z;R,G,B)\}\rightarrow\boldsymbol{M}\{\boldsymbol{V}_i(x,y,z;R,G,B)\}\quad(i=1,\cdots,N)\tag{7}$$

where $\boldsymbol{M}$ stands for a mesh model that is consisted of a complete and non-redundant point cloud and connectivity representing the topology of the model. The vertices $\boldsymbol{V}_i$ of point cloud comes from data set of range images, and each vertex is associated with a given color with Red-Green-Blue (RGB) format coming from associated color images. In our approach we apply a triangular mesh to build up a geometric model from the point-cloud. This model consists of a number of triangle facets, which are denoted by $\Delta_j\{\boldsymbol{V}_{1j},\boldsymbol{V}_{2j},\boldsymbol{V}_{3j}\,|\,j=1,\cdots M\}$. Each triangle facet Δ_j is composed of three adjacent vertices and its orientation defined by normal vector n. Geometric model of object can therefore be described as

$$\boldsymbol{M}\{\Delta_j\{\boldsymbol{V}_{1j},\boldsymbol{V}_{2j},\boldsymbol{V}_{3j}\}\,|\,n_j\}\quad(j=1,\cdots M)\tag{8}$$

On the other hand, the correspondence between each vertex and its color in input image has been established so that Eq. (8) can be rewritten as

$$\boldsymbol{M}\left\{\Delta_j\left\{\begin{matrix}\boldsymbol{V}_{1j}(R_1,G_1,B_1),\\ \boldsymbol{V}_{2j}(R_2,G_2,B_2),\\ \boldsymbol{V}_{3j}(R_3,G_3,B_3)\end{matrix}\right\}|\,n_j\right\}\quad(j=1,\cdots M)\tag{9}$$

It will be seen later that this geometric and color representation particularly fit into the algorithm of texture blending or stitching.

4 Texture stitching and photorealistic model generation

Because of the variation of environmental lighting and inaccuracy of geometric model reconstruction, the incoherence exists within the overlapped region and in the borders when more than one texture images taken from different viewpoints are mapped onto the same triangle facet. Those color discontinuities or will lead to visual incoherence. Therefore, the colored vertices within the overlapped regions and in the image-to-image borders should be blended or stitched in order to remove the incoherence or color discontinuities in textured 3D model. Various approaches have been proposed for selecting most correct color which has to be applied to each part of 3D geometric model. Several approaches were suggested for texture blending[19-20]. Some of them averaged the colors of texture image pixels projected from sample points of range image within the overlapped region whereas others utilized bilinear interpolation technique to smooth color variation across the edges of texture triangle patches.

In our recent approach, we suggested a

composite-weight blending function to compute texture color[21]. The basic idea behind this approach was to create a composite-weighted blending function that operates in texture image space, capable of mixing data from various texture images by weighting them with respect to the quality of each contribution. With this composite-weight function we were able to determine the colors of geometric triangle facets, resulting in a blended texture triangle patch.

Unlike those methods reported in previous literatures, in this paper, we are focusing on a more general framework for texture stitching and address two issues with respect to color mappings from photographic data sets. By exploiting the benefit of per-vertex color encoding we present alternative texture stitching method by introducing a two-step approach: 1) creating texture mosaic onto the range images using a strategy of Markov Random Field (MRF) optimization, and 2) stitching textured fragments based on a simple and straightforward method to remove the residue incoherence or color discontinuities in final color textured 3D model.

4.1 Creation of texture mosaic on 3D model

The core problems are to detect those pixels in each texture photo which sample the object surface, and identify which the coordinates of the sampled surface points (range image points) those pixels belong to. As mentioned earlier, it is possible to assign a pixel color, taken form that photo, to a point onto the surface. However, the real problem arises when the same surface point (range image point) can take the color from many different pixels since there is more than one candidate available in this case. Obviously it is not suitable to simply choose the color from a single texture image, ignoring the other values. That is because not all sampled colors yield the same degree of the quality. Since each pixel in the source texture images has a specific quality, it is therefore necessary to take into account this quality when stitching the texture patches.

We use the range image as a starting point because some discontinuities on the photo are more easily detectable when looking at the 3D geometric model (or range images). We need to consider how much each pixel in the source texture images will contribute the final color of 3D point it maps on. Various metrics can be applied to evaluate the quality of source texture images. Since we have not only the source color images but also a faithful geometric representation of the object, we can effectively use this information to perform a higher quality evaluation. Each metric can take into account a different characteristic of the source texture image. Using all the redundancy contained in the data set of source images, we are able to obtain a high quality color mapping. Since the color blending function relies on redundancy to determine the most correct color to apply, a good level of overlapping is essential to overcome the color discrepancies due to the illumination changes or color biasing between different texture images. A better error evaluation can be done regarding the color coherence. Every time we evaluate a blending function for a 3D point we sample the contributing input images to retrieve the source colors and then we compute the correct color as weighted mean. The difference between the sampled color and the final color is a measure of how much the source image was coherent with the other neighbors. We can determine, for each texture image, how severe this incoherence is by keeping track of all those mapping errors. If we detect that error value is above a threshold, it is possible to mark that texture image as discordant. In this case, it is also possible to provide a correction for problematic texture image. Recently Lempitsky and Ivanov proposed alternative approach for texture stitching[22]. The attractive thing of this approach is that it starts by back projecting original views onto the obtained object surface. A texture is formed by mosaicking from these back-projections, whereas the quality of mosaic is maximized within the process of Markov Random Field energy optimization. Finally the residual seams between mosaic components are removed via seam leveling procedure. However, the geometric shape of the object were obtained by using a strategy so called shape-from-silhouette cues instead of using 3D digitizing. In our approach here, we move forward along this direction, but build high quality 3D geometric models with our dedicated optical digitizer. In addition, we make use of a simpler scheme to correct color discontinuities in order to reduce the algorithm complexity. We formulate the issue of texture mosaic following Lempitsky and Ivanov's framework and notation system.

Consider a 3D model mesh with facets $F_1, F_2, \cdots, F_K$, and set of texture fragments $V^1, V^2, \cdots, V^N$, each corresponding to one of the initial views. Then the texture mosaic is defined by a labeling

vector $\boldsymbol{M}=\{m_1, m_2, \cdots, m_N\}\in\{0\cdots N\}^K$, prescribing to texture the facet F_i from the fragment V^{m_i}. Obviously, not all fragments containing a facet are equally suitable for texturing it. It is assume that the quality of the facet F_i in the fragment V^j is defined by the cost value w_i^j, which is smaller with better quality. If part of the facet does not fit into the fragment then the w_i^j is set to ∞. For other fragments, the cost values can be computed with different strategies. Thus w_i^j may be computed based on the angle between the local viewing direction and the face normal, giving preference to the less oblique back-projections. A number of approaches chose the fragment for each facet based on such quality measure, which corresponds to the labeling vector $\boldsymbol{M}=\{m_i \mid m_i=\mathrm{argmin}_j w_i^j\}$. However, it should be noted that such "best fragments" approach to mosaic does not take into account the visibility of the seams when two adjacent faces are textured from different fragments. Therefore Lempitsky and Ivanov proposed to minimize the visibilities of such seams simultaneously with the maximization of fragment quality, thus measuring the fragment quality of each possible mosaic $\boldsymbol{M}$ with a two-termed energy:

$$E(\boldsymbol{M})=E_Q(\boldsymbol{M})+\lambda E_S(\boldsymbol{M}) \tag{10}$$

Here the first term is the energy term corresponding to the quality of the fragments, while the second term corresponds to the distinguishable capability of the seams. Assume that two adjacent facets F_i and F_j share an edge E_{ij}, the visibility of a seam between F_i and F_j is measured by the integral difference of the colors of this edge in corresponding fragments:

$$w_{i,j}^{m_i,m_j}=\int_{E_{ij}} d(\mathrm{Pr}_{m_i}(X),\mathrm{Pr}_{m_j}(X))\mathrm{d}X \tag{11}$$

Here Pr_m is a projection operator for the view m, and $d(\cdot,\cdot)$ is some metric for colors or intensities (for example, Euclidean distance in RGB color space can be used for that purpose). Naturally, if $\mathrm{Pr}_{m_i}(X)=\mathrm{Pr}_{m_j}(X)$ then $w_{i,j}^{m_i,m_j}$ equals zero. In this case, texturing from the same fragment does not produce seams. If N denotes a set of pairs of adjacent facets, then the second term in Eq. (11) can be written as

$$E_S(\boldsymbol{M})=\sum_{\{i,j\}\in N} w_{i,j}^{m_i,m_j} \tag{12}$$

and the overall energy Eq. (10) can be written as

$$E(\boldsymbol{M})=\sum_{i=1}^{K} w_i^{m_i}+\sum_{\{i,j\}\in N} w_{i,j}^{m_i,m_j} \tag{13}$$

The functional similar to Eq. (13) is common in computer vision, as they are associated with probability distributions of pair wise MRF. In our approach, the MRF is mesh-based, its nodes correspond to mesh facets, and node interactions are defined by facets adjacency. The minimization of Eq. (13) can be achieved by using three state-of-the art techniques: max-product Belief Propagation (BP), α-expansion Graph Cut, and convergent tree reweighted message passing. The optimized mosaics give textures, which are visually superior to the "best fragment" approach. MRF mosaicking decreases the distinguishable ability of seams significantly. However, in the presence of photometric differences between original images it can make the seams completely unnoticeable. Therefore the remaining intensity discontinuities require further processing. In an approach[21], the authors proposed a so called seam leveling procedure to perform the post processing of the texture mosaicking. In our approach here, we propose alternative approach to remove the residue intensity discontinuities in order to reduce the algorithm complexity of seam leveling.

4.2 Removal of residue seams and other incoherence

Here we describe how to remove the residue incoherence or color discontinuities in final color textured 3D model of real object. A number of gradient-domain methods for seamless stitching, which avoids intensity blending, were proposed for planar images. Recently, Lempitsky and Ivanov present a new method referred to as seam leveling for an adaptation of those method to the problem of fragment combination of on arbitrary manifold, e. g. meshes. However, this technique must first formulate a piecewise continuous function f on an arbitrary smooth orientated manifold where seam is a submanifold of codimension formed by the discontinuity points, then define a piecewise continuous seam leveling function g with same discontinuity as f. This approach tried to find out g through solving an optimization problem of variation so that the sum $f+g$ is an everywhere continuous function on underlying manifold. Mathematically, this approach is a beautiful formulation to the problem of seamless stitching. However, it would be tedious and troublesome in terms of numerical implementation and practical applications. In this section, we propose a simpler and more straightforward method for removing the residue seams and other incoherence.

To achieve seamsless stitching, we take into account another procedure to remove the remaining seams and other artifacts. To do this, the boundary triangles should first be detected in the

boundaries of fragments on the mesh model. A boundary triangle is defined when three vertices of underlying triangle are textured from different views (e.g. texturing from two or three images). Without loss of generality, we assume that a boundary triangle Δ_i being composed of three vertices $\boldsymbol{V}_a$, $\boldsymbol{V}_b$, $\boldsymbol{V}_c$ are textured from three input images $\boldsymbol{I}_a$, $\boldsymbol{I}_b$, $\boldsymbol{I}_c$, then a resample step is performed to get the sample points p_i, which can be written as

$$\boldsymbol{p}_i = \alpha \boldsymbol{V}_a + \beta \boldsymbol{V}_b + \gamma \boldsymbol{V}_c \tag{14}$$

where $\alpha + \beta + \gamma = 1$, the sample points $\boldsymbol{p}_i$ is represented by three weighted vertices on the mesh model. By using the perspective projection it is easily to determine the texture coordinates of $\boldsymbol{p}_i$ in a straightforward way. Furthermore, we assume that the colors of vertices on associated input image are denoted by $\tilde{\boldsymbol{v}}_a$, $\tilde{\boldsymbol{v}}_b$, $\tilde{\boldsymbol{v}}_c$, and the color of a sample point on the texture map is denoted by $\boldsymbol{C}_{\mathrm{p}}$. Thus, $\boldsymbol{C}_{\mathrm{p}}$ can be written as

$$\boldsymbol{C}_p = \alpha \tilde{\boldsymbol{v}}_a + \beta \tilde{\boldsymbol{v}}_b + \gamma \tilde{\boldsymbol{v}}_c \tag{15}$$

This projective operation is traversed over the entire textured mesh model, then overall colored sample points on the texture map are obtained in this way and they are saved into a data bank for later processing. The textured triangle facets are composed of colored vertices taken from the data bank and they are packed into the texture images, which will be used to generate a photorealistic 3D model by back projection. The key of this seams removal method is to handle those triangles adjoining adjacent fragments on the mesh model and textured triangles within the overlapped regions on the texture image.

5 Experiment verifications for cultural heritage digitization

A prototype of 3D color optical digitizer developed in our laboratory used in the experiment is shown in Fig. 2. This dedicated optical digitizer

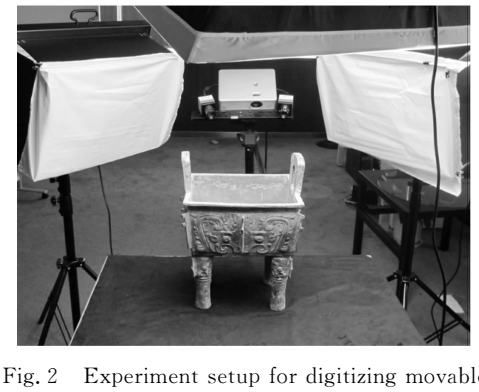

Fig. 2　Experiment setup for digitizing movable cultural heritage

is based on m-FPP, which is composed of a digital color camera (The Imaging Source DFK-41AF02 with 1394 interface and 1 280×960 resolution), a B/W camera (The Imaging Source DMK-41BF02 with 1 394 interface and 1 280×960 resolution), and a DMD projector (LG HX300G-JE with 1 024×768 resolution) as basic units.

Once optical digitizer is calibrated, the m-FPP sensor can be used to acquire the range images and associated color images, and the mapping relationship between the mesh model and associated texture images has been established in a precise way. Therefore it is possible to back project the associated texture images onto the mesh model (rendering in terms of computer graphics) accurately to produce textured fragments. Two kinds of bronze wares are taken as example artifacts to verify our approach.

The first experiment was performed by taking a bronze ware Ding as exemplary artifact of movable cultural heritage to verify our approach. The Ding with thin-wall shape structure not only has complicated geometric shape and topology, but also unique color attribute. A photorealistic 3D image or model should correctly reflect these characteristics of the artifact. The length-width-height of Ding are with 292 mm, 221 mm, and 377 mm, respectively. In the experiment the Ding was placed on an optical table as show in Fig. 2, and its range images and associated color images were acquired from 44 views by fixing optical digitizer while changing the poses of the Ding with respect to optical digitizer. Then the multi-view colored range images were registered and integrated to form a complete and colored 3D mesh model using the method discussed in Section 3. Fig. 3 (a) shows the photo of this bronze ware, (b) and (c) show partial range images with color from two different viewpoints, (d) shows registered range images of this bronze ware, (e) the reconstructed 3D model, and (f) the color textured 3D model obtained by our approach, which is a combination of the MRF mosaic and residue seams removal. From experiment results shown in Fig. 3 we can see that the proposed approach is able to generate a photorealistic model of movable cultural heritage, like Ding used in this demonstration.

The second experiment was to perform with another kind of bronze wares, *Jue*, which has totally different geometric shape, topology and color attribute. The height and maximum diameter

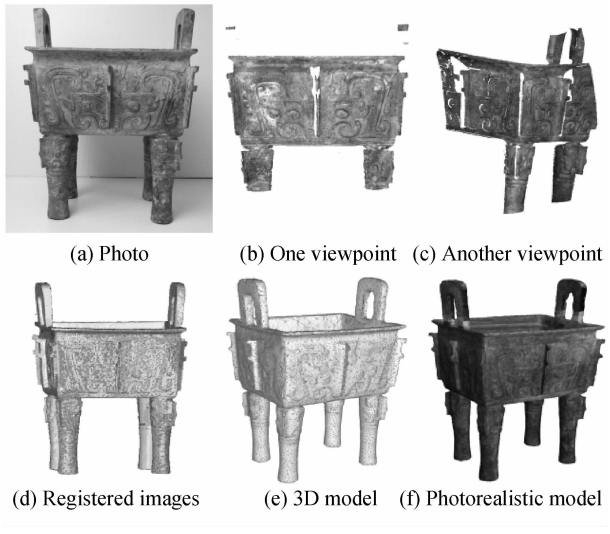

Fig. 3 Experimental verification of our approach using *Ding*

are 197 mm and 60 mm, respectively. During the acquisition, 32 views were captured by altering the poses of the object, and multi-view partial range images were then registered and integrated into a complete model. Fig. 4(a) shows the photo of this bronze ware, (b) and (c) show partial range images with color from two different viewpoints, (d) shows registered range images of this bronze ware, (e) the reconstructed 3D model, and (f) the color textured 3D model obtained by our approach. Again, this experiment verifies the feasibility of proposed approach.

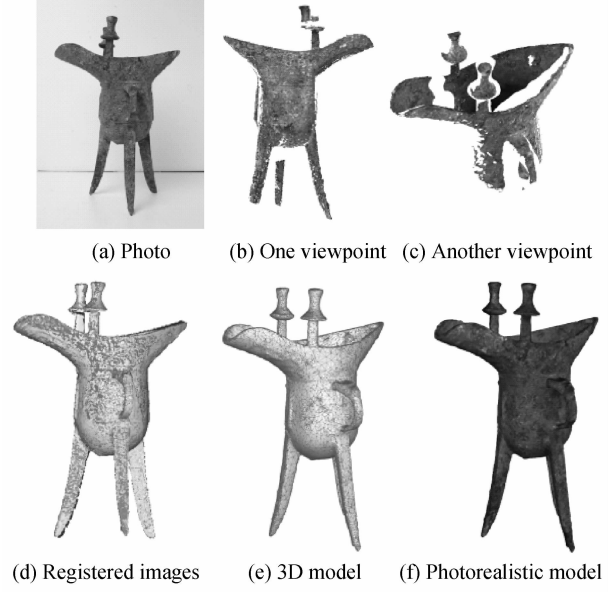

Fig. 4 Experimental verification of our approach using *Jue*

Both experiments show that the proposed approach is feasible to create photorealistic models of cultural heritage. The calibration technique described in our recent paper[16] guaranteed that the views were accurately registered with geometric model and back-projected onto surface in a precise way. Once these prerequisite conditions are satisfied, we are able to generate a high quality and high resolution photorealistic 3D model of real artifacts with our proposed approach. Fig. 5 and Fig. 6 present the observations from different view angles of reconstructed photorealistic model of *Ding* and *Jue*, respectively.

Fig. 5 Observations of texture model of *Ding* with 4 different viewing angles

Fig. 6 Observations of texture model of *Jue* with 4 different viewing angles

6 Conclusion

In conclusion, we have verified that a computer-aided optical digitizer can be used to acquire photorealistic 3D models of movable cultural heritage. This dedicated optical digitizer is able to complete the whole process from colored range image acquisition to photorealistic 3D model creation. In particular, we have suggested a texture stitching technique that combines two steps: creation of texture mosaic on 3D model with MRF optimization and removal of residue seams and other incoherence. Two kinds of bronze wares have been taken as example artifacts to verify the feasibility our proposed approach. It should also be pointed out that the geometric models obtained by 3D optical digitizer possess much higher quality than by shape-from-silhouette. We believe that our approach should be useful for the generation, storage, and retrieval of 3D digital information of cultural heritage in the field of digital preservation. More experimental verifications using other artifacts such as pottery and porcelain are ongoing and will be reported in the near future.

Reference

[1] GLADNEY H. Pciples for digital preservation [J]. *Communications of the ACM*, 2006, **49**(2): 111-116.

[2] PETROV M, TALAPOV A, ROBERTSON A, *et al*. Optical 3D digitizers: bringing life to the virtual world[J]. *IEEE Computer Graphics and Applications*, 1998, **18**(3): 28-37.

[3] GROVES R, LI A, LIU X, *et al*. 2. 5D virtual reality visualisation of shearography strain data from a canvas painting [C]. SPIE, 2009, **7391**: 739109-1-8.

[4] CHEN F, BROWN G, SONG M. Overview of three-dimensional shape measurement using optical methods [J]. *Optical Engineering*, 2000, **39**(1): 10-22.

[5] BLAIS F. Review of 20 Years of range sensor development [J]. *Journal of Electronic Imaging*, 2004, **13**(1): 231-240.

[6] GORTHI S, RASTOGI P. Fringe projection techniques: Whither we are[J]. *Optics and Lasers in Engineering*, 2010, **48**(2): 133-140.

[7] GUIDI G, BERALDIN J, ATZENI C. High-accuracy 3D modeling of cultural heritage: the digitizing of Donatello's "Maddalena"[J]. *Image Processing*, 2004, **13**(3): 370-380.

[8] CIGNONI P, SCOPIGNO R. Sampled 3D models for CH applications: A viable and enabling new medium or just a technological exercise[J]. *Journal on Computing and Cultural Heritage*, 2008, **1**(1): 2.

[9] REMONDINO F, GIRARDI S, RIZZI A, *et al*. 3D modeling of complex and detail cultural heritage using multi-resolution data[J]. *Journal on Computing and Cultural Heritage*, 2009, **2**(1): 2.

[10] BERNARDINI F, RUSHMEIER H, MARTIN I, *et al*. Building a digital model of Michelangelo's Florentine Pieta [J]. *IEEE Computer Graphics and Applications*, 2002, **22**(1): 59-67.

[11] GENG J. Structured-light 3D surface imaging: a tutorial[J]. *Advances in Optics and Photonics*, 2011, **3**(2): 128-160.

[12] TUTSCH R, PETZ M, FISCHER M. Optical three-dimensional metrology with structured illumination [J]. *Optical Engineering*, 2011, **50**(10): 101507-101510.

[13] REICH C, RITTER R, THESING J. 3-D shape measurement of complex objects by combining photogrammetry and fringe projection [J]. *Optical Engineering*, 2000, **39**(1): 224-231.

[14] ZHANG S, HUANG P. Novel method for structured light system calibration[J]. *Optical Engineering*, 2006, **45**(8): 083601.

[15] PENG Xiang, WEI Lin-bin, QIU Wen-jie, *et al*. Phase Reconstruction Based on Generalized Fringe Pattern Sequence Encoding[J]. *Acta Optica Sinica*, 2006, **26**(8): 1156-1161.

[16] YIN Y, PENG X, LI A. Calibration of fringe projection profilometry with bundle adjustment strategy [J]. *Optics Letters*, 2012, **37**(4): 542-544.

[17] BESL P, MCKAY H. A method for registration of 3-D shapes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, **14**(2): 239-256.

[18] TURK G, LEVOY M. Zippered polygon meshes from range images[C]. SIGGRAPH 94, Orlando: ACM, 1994.

[19] BERNARDINI F, MARTIN I, RUSHMEIER H. High-quality texture reconstruction from multiple scans[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2001, **7**(4): 318-332.

[20] CALLIERI M, CIGNONI P, CORSINI M, *et al*. Masked photo blending: Mapping dense photographic data set on high-resolution sampled 3D models[J]. *Computer Graphics*, 2008, **32**(4): 464-473.

[21] LIU X, PENG X, YIN Y, *et al*. Generation of photorealistic 3D image using optical digitizer[J]. *Applied Optics*, 2012, **51**(9): 1304-1311.

[22] LEMPITSKY V, IVANOV D. Seamless mosaicing of image-based texture maps[C]. CVPR 2007, Minneapolis USA: IEEE, 2007: 1-6.