

# OPTICAL THREE-DIMENSIONAL SURFACE RECONSTRUCTION AND EVALUATION USING A MULTIMODALITY IMAGING INSTRUMENTATION

Xiaoming Jiang\*, Liji Cao, Jörg Peter  
Division Medical Physics in Radiology  
German Cancer Research Center  
Im Neuenheimer Feld 280, Heidelberg, Germany  
Email: xiaoming.jiang@dkfz.de

## ABSTRACT

This paper proposes an optical surface reconstruction method from multiview projectional data acquired by a multimodality imaging instrumentation. The technique is adapted for *in vivo* small animal imaging, specifically imaging of nude mouse in the case where the influence of CT radiation doses should be eliminated. Any potential point within the field-of-view (FOV) of a nude mouse is evaluated by a proposed photo-consistency measure utilizing sensor image information. As the superposition of adjacent projections yields depth information for any point within the FOV, the three dimensional (3D) surface of the imaged object is estimated by a graph-cuts based method through global energy minimization. The reconstructed surface is evaluated by comparing the reconstructed surface and the CT volume of the nude mouse. The proposed surface reconstruction method demonstrates the feasibility of surface reconstruction from multiview projection, and there are still great possibilities to improve this method.

## KEY WORDS

Medical imaging and image processing, Optical imaging, Multi-modality imaging

## 1 Introduction

*In vivo* small animal non-contact optical imaging (OI) holds great potentials in a broad spectrum of biomedical research [1]. However, due to the characteristics of low energy photons (1.5 eV- 2.1 eV) resulting in severe scattering and attenuation thereof in tissues, techniques in *in vivo* small animal OI, whether bioluminescence imaging (BLI), fluorescence imaging (FLI) or fluorescence mediated tomography (FMT) etc., share the common characteristic of limited tissue penetration depth [2]. Moreover, the detected optical signals in OI are surface weighted, i.e. the detection is greatly influenced by the viewing angle of the data recording system [3]. Thus, classic two-dimensional planar images without the depth information of the imaged object are insufficient to reveal the real probe distribution *in vivo*. To improve the accuracies of reporter localizations and solve these two aforementioned problems, many efforts have been applied with respect to multimodality imaging

strategies and optical detections from multiview projection [2].

The utilization of a secondary imaging modality to enhance the performance of OI has been commonly accepted due to the natural complementariness between different imaging techniques. Among various combinations proposed and implemented, computed tomography (CT), an imaging modality capable of revealing the attenuation properties of the imaged object with respect to the X-ray spectrum applied, is the foremost imaging approach to be integrated with OI. This is due to that CT has a high spatial resolution, and is routinely used in biomedical research while CT/OI could be designed without consideration of the compatibility problems compared with magnetic resonance imaging (MRI) etc.. Therefore, a number of systems combining CT and optical systems are constructed, as listed in Ref. [2]. However, this combination has also some inherent shortcomings, especially in the aspect of radiation doses introduced by CT. In some specific biomedical applications the influence of radiation dosage can not be omitted.

Alternatively, the surface of the imaged object can be reconstructed optically using the optical camera system (originally for the detection of emission signal) instead of CT. Some imaging systems combining the information from optical surface have used this strategy. Most systems employ simple surface reconstruction algorithms, e.g., the threshold based back-projection method [4]. When using this technique, mice are generally fixed vertically to a rotational stage at an unnatural state so that the main part of mouse body has a convex shape because the back-projection method cannot resolve concave surface areas. In fact, surface reconstruction from multiview projection is a topic originally belonging to the field of computer vision, which has been extensively studied [5].

In this paper, we present a graph-cuts based volumetric surface reconstruction method for nude mice using a CT/OI multimodality imaging instrumentation [6]. The accuracy of the reconstructed surface is evaluated by comparing the reconstructed surface with the obtained CT volume of the imaged mouse. A comparison is also performed between the proposed method and the back-projection method.

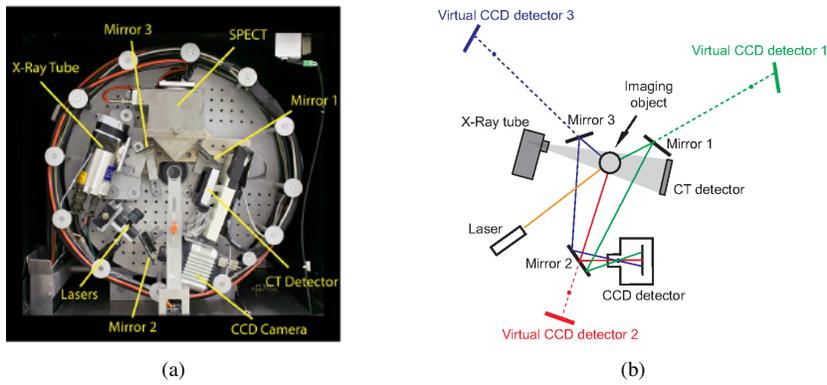


Figure 1: CT/OI bimodal imaging system. (a) photograph of the system. (b) schematic top view of the system configuration.

## 2 Methods

### 2.1 Imaging system

The instrumentation herein used for optical surface reconstruction and evaluation is the trimodal imaging system, cf. Fig. 1a and Fig. 1b, including an OI subsystem, a CT subsystem, and single emission computed tomography (SPECT) subsystem [6] (SPECT is not used in the context of optical surface reconstruction, and, therefore, it is omitted here). In the OI subsystem a light camera (ORCA-AG, Hamamatsu Photonics, Shizuoka, Japan) is configured with three mirrors distributing around the FOV. Thus, three virtual views provided by three mirrors can provide a full coverage of the FOV for the light camera, as seen in Fig. 1b. The x-ray CT (Apogee Series 500, Oxford Instruments, Oxfordshire, UK) shares the same FOV of OI so that the reconstructed optical surface could be compared with the results retrieved from segmented CT data.

### 2.2 Optical surface reconstruction

As the visibility information of the imaged object from multiview projections is not available, methods regardless of self-occlusion are preferred, e.g., Refs. [7, 8]. In addition, the light camera used herein was calibrated with a model considering the in-plane rotational angle [6], which is different from the classic calibration model used in computer vision [9]. In the Refs. [7, 8], it is assumed that there is a virtual ray going through each voxel in object space and the camera's optic center. The photo-consistency measure is calculated by finding a point along the optic ray and evaluating the similarities between the formed pixel windows by the point. However, the projection matrix of the light camera used here is not available so that the position of camera's optic center could not be calculated. Therefore, the existing algorithms in computer vision cannot be directly applied here for the optical surface reconstruction.

A different strategy is used in the step of photo-consistency measure computation concerning the specificity of the system used. In first step the photo-consistency measure of a voxel could be evaluated by calculating the

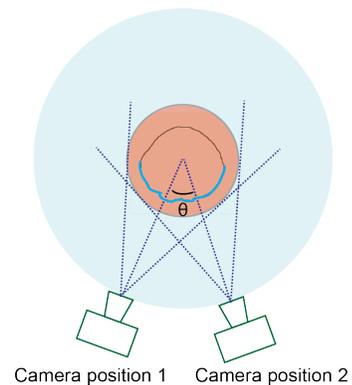


Figure 2: Illustration of the light camera configuration with one of three virtual views being used. The NCC value of voxels in object space per two projection images could be computed using two images herein taken from camera position 1 and position 2 with a tilt angle of  $\theta$  (set to be  $30^\circ$  in this paper). Due to the self-occlusion problem, voxels that are on the surface of the imaged object, and in the common FOV of two cameras, as marked in blue, process bigger NCC values than other voxels.

normalized cross correlation (NCC) using two projection images. The NCC between the formed pixel window ( $\mathbf{W}_p$ ) in one projection image and the pixel window ( $\mathbf{W}_r$ ) in the other neighboring projection image by the same voxel, as shown in Fig. 2, according to

$$\text{NCC}(\mathbf{W}_p, \mathbf{W}_r) = \frac{\sum_{i,j} (\mathbf{W}_p^{i,j} - \overline{\mathbf{W}}_p)(\mathbf{W}_r^{i,j} - \overline{\mathbf{W}}_r)}{\sqrt{\sum_{i,j} (\mathbf{W}_p^{i,j} - \overline{\mathbf{W}}_p)^2} \sqrt{\sum_{i,j} (\mathbf{W}_r^{i,j} - \overline{\mathbf{W}}_r)^2}} \quad (1)$$

where  $\overline{\mathbf{W}}_p$  and  $\overline{\mathbf{W}}_r$  represent the average value of two corresponding pixel windows, respectively. In this paper  $7 \times 7$  pixel window is used. Due to the self-occlusion problem, only voxels that are on the surface, and in the common FOV of two cameras, have bigger NCC value than the other voxels using two projection image, as illustrated in Fig. 2. In the second the calculation is extended from using two projection images to multiview projection images. As in-

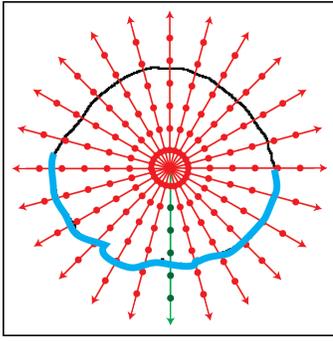


Figure 3: Illustration of a radial sampling process. Knowing that voxels on the surface and in the common FOV of two projection images have bigger NCC values, as contours marked in blue, the radial sampling is performed to increase the contrast of reconstructed image. The sampling process begins from a point of estimation inside of the imaged object (coordinates origin here), one of which is marked in green. Voxels on the surface have biggest NCC value along the radial direction are remained, which could be added up using multiview projection images, while other voxels are omitted.

indicated in Ref. [7, 8], the direct average of NCC values would decrease the resolving ability of surface. Thus, a radial sampling process is performed to increase the contrast of NCC values for the calculated volume using two projection images, as shown in Fig. 3. The sampling process begins from a point of estimation inside of the imaged object in every radial direction. Voxel on the surface processing biggest NCC value along the radial direction are remained, which could be added up using multiview projection images, while other voxels are omitted. It is notable that the radial sampling could be performed under the hypothesis that in preclinical optical imaging the volume of interest is limited [10] between thorax and abdomen, i.e., the imaged object is with a simple geometry.

The photo-consistency measure of a voxel  $V_{i,j,k}$  is calculated to be

$$\varphi(V_{i,j,k}) = \exp(-\mu\sigma) \quad (2)$$

where  $\mu$  is a rate-of-decay parameter (set to be 0.5 in this paper) and  $\sigma$  is the accumulative NCC value using multiview projection images. By calculating all the photo-consistency measures a photo-consistency volume of the imaged object is obtained. The problem of surface reconstruction is applied onto the formed volume, and could be further solved by the graph-cuts based method [11].

### 2.3 Reconstructed surface evaluation

To evaluate the performance of the proposed method, it is compared with the state-of-the-art back-projection. Meanwhile two measures are used comparing the differences between the reconstructed optical surface and the segmented CT data of the mouse. The first measure is error rate. Given

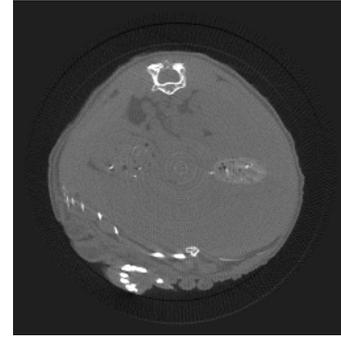


Figure 4: One reconstructed transaxial CT slice of the nude mouse.

the segmented CT binary volume  $V_{\text{phantom}}$  of the imaged object, the binary volume  $V_{\text{recon}}$  via optical surface reconstruction could be compared according to

$$\gamma = |V_{\text{phantom}} \oplus V_{\text{recon}}| / |V_{\text{phantom}}| \quad (3)$$

where  $\oplus$  represents an exclusive or operation (XOR) between two binary volumes and  $|\cdot|$  is a  $l_1$  norm operation of binary data.

The second measure is concerning the accuracies of reconstructed surface. The segmented CT binary volume could be further represented by a triangle mesh, and be used as the reference mesh for the evaluation. Provided the triangle mesh of the reconstructed surface, which is usually described by vertices and faces, the evaluation is conducted by looping each vertex of reconstructed mesh and finding the minimal distance between the fixed vertex in the reconstructed mesh and faces from the obtained reference mesh. The mean value and standard deviation of these minimal distances for vertices could be used for the measurement of dispersion between two meshes.

## 3 Results

The proposed method has been investigated and evaluated using a nude mouse. First, a CT scan was conducted to extract the surface of the nude mouse from its CT volume. One reconstructed slice of the CT data is shown in Fig. 4. The reconstructed CT data was segmented into binary data (the nude mouse and the background), which is later used as the ground truth for the evaluation of the reconstructed surface.

Afterwards, optical images were acquired from 120 projectional angles (the step angle is set to  $3^\circ$ ). An example projection image is shown in Fig. 5. Due to the specific design of the multimodality imaging system, there are three virtual views in each projection image. The region used for surface reconstruction from multiview projection is marked in red. In order to perform the back-projection method, a semi automatic segmentation method is employed to extract the imaged mouse from the background in projection image, as shown in Fig. 6.

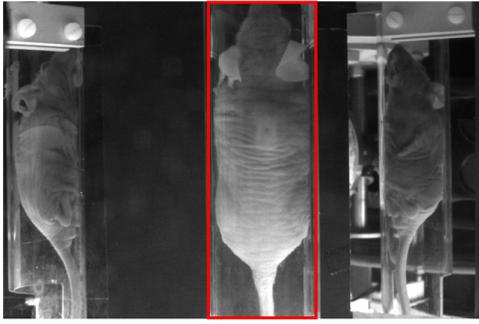


Figure 5: An example image of the nude mouse inside the FOV of the optical system. With the use of mirrors, the nude mouse placed in a glass tube is imaged simultaneously by three virtual cameras. The region marked in red is used for multiview surface reconstruction.

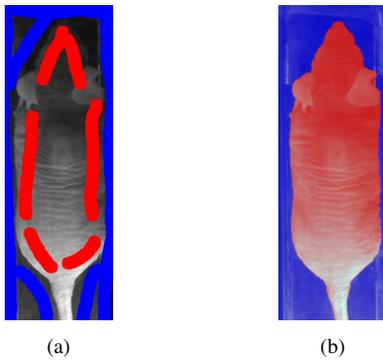


Figure 6: An example of semi automatic segmentation for the projection image marked by the red box in Fig. 5. In (a) the scribbles in red and blue provide the initial information for the nude mouse and background, respectively. The corresponding segmented result is shown in (b). Regions in red illustrate the result of segmentation for the nude mouse.

A volumetric surface reconstruction is then performed on a  $128 * 128 * 146$  volume with a voxel size of 0.26 mm. A slice of volume with the same height as the previously shown CT slice in Fig. 4 is calculated according to the proposed photon-consistency measure, as shown in Fig. 7. The photon consistency volume is segmented by the graph-cuts method. A comparison among the segmented CT slice and the extract slices at the same height using the proposed method and the back-projection method is shown in Fig. 8. The overall error rate between the reconstructed surface volume using the proposed method and the CT volume is calculated to be 12.44% while the result using the back-projection method yields an error rate of 9.70%. The measured accuracies are  $0.936 \pm 0.336$  mm and  $0.760 \pm 0.296$  mm for these two methods, respectively.

The corresponding rendered surfaces from the CT data and the reconstructed volume are shown in Fig. 9 (a), (b) and (c), respectively. Although some differences exist between the reconstructed surface and the ground truth, most concave areas of the nude mouse are resolved using

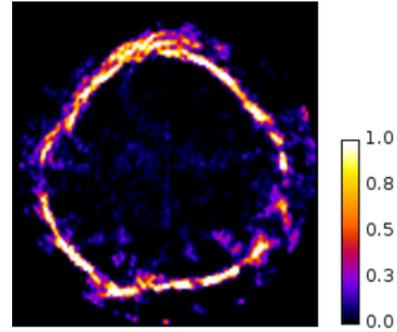


Figure 7: The reconstructed trans-axial slice of photon-consistency measure at the same height as the CT slice in Fig. 4.

the proposed method while not using the back-projection method, as seen in Fig. 9.

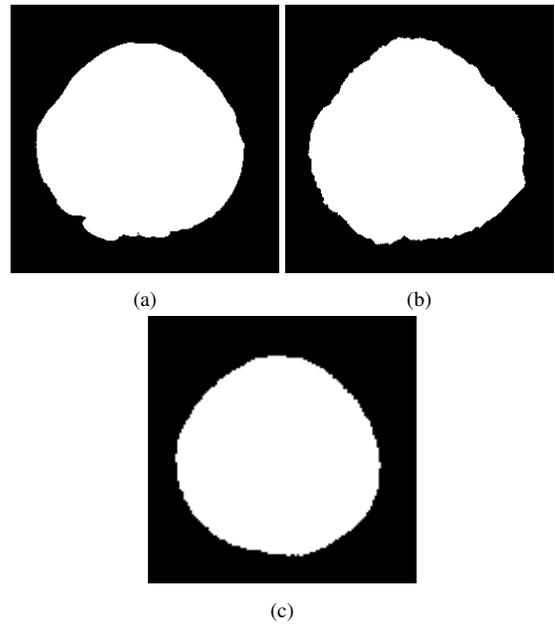


Figure 8: A comparison of the reconstructed trans-axial slices using different methods. (a) is the result from CT segmentation. (b) is the segmented volumetric slice at the same height using the proposed method. (c) is the slice derived by the back-projection method.

## 4 Discussions

Our proposed method yields a reasonable accuracy of about 1 mm, which is in agreement with the similar studies using the back-projection method [4, 12]. The results indicate that the overall performance of the back-projection method is slightly better than that of the proposed method. The main reason is that there are only a few small concave areas on the surface of nude mouse, and the improvement using the proposed method is not so apparent. However,

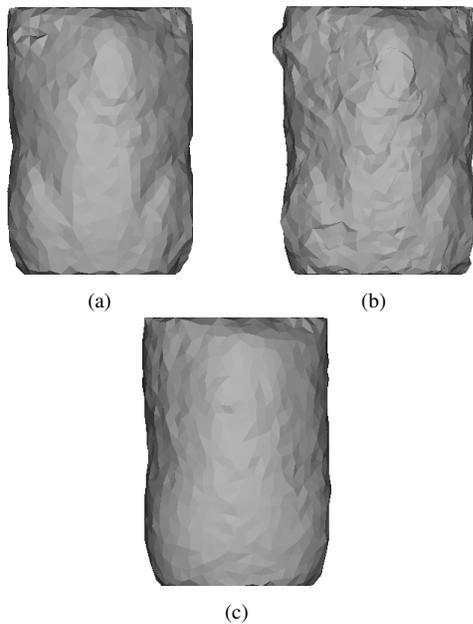


Figure 9: Rendered surfaces of the imaged nude mouse using the CT segmentation data (a), the data obtained by the proposed method in this paper (b), and the back-projection method, respectively.

compared with the classic back-projection methods, which require mice to be tied to rotational stages, the handle of animals in the experiment (placing a mouse in a glass tube) is comparatively easy. In addition, the proposed method does not require extra segmentation step for projection images. In the context of this experiment, due to the use of glass tube as an animal holder, some parts of mouse skin remain attached to the glass tube. So the direct application of back-projection is not possible because of the difficulty for the segmentation of the mouse in projection images. A semi segmentation method [13] is used to overcome the restriction. But a tremendous amount of work has to be done drawing seed scribbles for over one hundred images, an example of which could be seen in Fig. 6.

There are possibilities to further improve this proposed approach. Firstly, the geometrical calibration of the multimodality imaging system used is based on an off-line strategy [6]; the last calibration of the system was performed half a year ago whereas the proposed method requires very accurate calibration. Secondly, in this experimental study a rather faint environmental light was used to illuminate the nude mouse. The images were acquired with a long exposure time, which made the obtained image a bit blurred. This could be observed by the calculated slice of photo-consistency measure in Fig. 7. There are some gaps in the areas which are in fact on the surface of the nude mouse. Thirdly, we only used the information from one virtual camera. In the future the surface might be reconstructed using virtual view images at three different scales.

## Conclusion

In conclusion, we have proposed a surface reconstruction from multiview projection and demonstrated its feasibility using a multimodality imaging instrumentation by experimental validation.

## References

- [1] V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder, *Nature biotechnology*, 23 (3), 2005, 313–320.
- [2] C. Darne, Y. Lu, and E. M. Sevick-Muraca, *Physics in medicine and biology* 59(1), 2014, R1–R64.
- [3] T. F. Massoud and S. S. Gambhir, *Genes & development* 17(5), 2003, 545–580.
- [4] H. Meyer, A. Garofalakis, G. Zacharakis, S. Psycharakis, C. Mamalaki, D. Kioussis, E. N. Economou, V. Ntziachristos, and J. Ripoll, *Applied optics* 46(17), 2007, 3617–3627.
- [5] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York City, USA, 2006, 519–528.
- [6] L. Cao, M. Breithaupt, and J. Peter, *Physics in medicine and biology* 55(6), 2010, 1591–1606.
- [7] G. Vogiatzis, C. Hernández, P. Torr, and R. Cipolla, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 29(12), 2007, 2241–2246.
- [8] X. Jiang, L. Cao, and J. Peter, *Optical Engineering* 53(2), 2014, 023104–023104.
- [9] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library* (Sebastopol California: O’reilly, 2008).
- [10] N. Ducros, A. Bassi, G. Valentini, G. Canti, S. Arridge, and C. D’Andrea, *Journal of biomedical optics* 18(3), 2013, 020503–020503.
- [11] Y. Boykov, O. Veksler, and R. Zabih, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23(11), 2001, 1222–1239.
- [12] G. Hu, H. Li, J. Yao, and J. Bai, *2008 IEEE International Conference on BioMedical Engineering and Informatics* Hainan, China, 2008, 76–80.
- [13] M. Tang, L. Gorelick, O. Veksler, and Y. Boykov, *2013 IEEE International Conference on Computer Vision*, Sydney, Australia, 2013, 1–8.