Eurographics Italian Chapter Conference (2011)
A. F. Abate, M. Nappi, and G. Tortora (Editors)

3D Reconstruction and Augmented Reality in Bronchoscopic Intervention

A. Torrisi[1], S. Livatino[2], G. Gallo[1]

[1]Department of Mathematics and Computer Science, University of Catania, Italy
[2]School of Engineering and Technology, University of Hertfordshire, United Kingdom

Abstract

The use of stereoscopic visualization has recently been proposed for many applications. Other than for entertainment, stereo viewing is being proposed for robotics and medical teleoperation. This paper proposes to exploit the stereo-camera setup available in stereo-viewing systems for 3D reconstruction. We focus on medical endoscopic applications. The advantages of having a reconstructed vision-based depth map during endoscopic navigation and intervention are many, including the possibility for generation of an augmented-reality visual scenario to support surgeons during interventions. We ran our experiments on a realistic graphic model of human bronchus to study feasibility of the proposed concept. We reconstructed depth maps of a bronchus environment and used them to generate augmented reality views of the observed scenes.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

1. Introduction

3D vision systems are currently used for enhancing depth perception and to provide a greater immersive experience for different research domains. The effectiveness and usefulness of stereo property has also been suggested in many medical applications. “Intuitive Surgical”, an American company, has proposed the integrated system “Da Vinci” [LBK*05] which permits the fulfillment of different surgical operations receiving tridimensional data from a stereoscopic optics positioned at the extremity of the surgical probe. A newer different technology [Vis10] uses a single sensor composed by many micro-lenses looking at different directions. The layout of the lenses is similar to the eye-structure of a bug. Left and right images of the observed scene are obtained with proprietary software, hence detailed technical information about this system is not easily obtainable.

Modern technologies provide flexible endoscopes that include plenty of accessories and utilities. At the same time the resolution of the integrated optics has been significantly improved. Currently, there are no companies that offer flexible stereo endoscopes while this solution seems to be promising and it will certainly will soon be on the market. A typical problem by using flexible endoscopes is that the operator loses track of the route covered during the navigation. In this regard, some medical tracking systems have been proposed. These systems calculates the position and the orientation of surgical instruments using optical or magnetic sensors [NDI11]. Other approaches attempt to reconstruct the path followed by the endoscope-tip position. However, this does not provide information on tip position at run time.

In this contribution we intend to exploit the stereo-camera setup available in stereo-viewing systems for 3D reconstruction. In particular, we aim at extracting depth information of the observed scene. We focus on bronchoscopic applications. Nevertheless, our concept can be applied to different endoscopic procedures. Lacking a real instrument, i.e., a flexible bronchoscope equipped with a stereo camera, we perform our experiments in a simulated virtual environment. In particular, the stereoscopic matchings have been done using a synthetic 3D model of the tracheobronchial tree. Given two calibrated views in this model, stereo data are exploited for the construction of a depth map with respect to the tip of “virtual” bronchoscope. The information provided by depth map representation can be used in different ways. It might be useful as a complement to others proposed methods aimed to track the tip of surgical instruments. It is also useful to improve the visual navigation and surgical intervention. An important advantage of this reconstruction is to enable the

DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2011/009-013

use of augmented reality to support the endoscope teleguide. In particular, depth maps extracted from our virtual model can be combined with incoming video images.

The paper is organized as follows: Section 2 reports the details of the extraction of the depth map. Section 3 discusses the obtained results. In Section 4 conclusions are drawn.

2. Experiments description

2.1. Rendering technique

The usage of virtual reality techniques in clinical applications is getting more wide-spread because of the availability of more detailed simulation models. Virtual simulators offer many advantages to the medical staff: they are especially valuable for training purposes, for pre-operative planning and evaluation of surgical skills [VDMS09], [HFSL*00]. Virtual reality can recreate the conditions of experimental research that would be difficult to propose in the real world. In particular, in this article, virtual reality is applied to recreate the typical environment of a bronchoscopy. To this aim we have preliminarly built a geometric model of a significative segment of the tracheobronchial tree. The synthetic model has been realized with the open source software Blender [Ble11] using real bronchoscopic images as a reference. Special care has been taken to replicate in the virtual model two of the main geometric parameters of the human pulmonary cavities: the branching degree (i.e. the rate of bifurcation of the air channels as one goes down the respiratory tree) and the decreasing rate of the tube sections after each bifurcation. Alteration of normal diameter in airways is an indicator for a better understanding of the pathology affecting the bronchial tree. One of the advantages of bronchoscopy is the accurate measurement of these sections for diagnostic purposes. This analysis is difficult to accomplish using X-ray or CT scans images. To enrich with a more realistic value our model, we have reproduced the typical pattern of the respiratory system through the use of some simple procedural textures provided by the standard Blender version. Lighting conditions have been simulated using directional spotlights properly controlled so that light fades down the pulmonary branches approximately as the it would in a real bronchoscopy.

To perform a stereo reconstruction step in this model, we have placed a couple of aligned cameras with parallel optical axes. The cameras lie in the same vertical and depth axes. They share the same focal length and differ only in the horizontal baseline between them. Taking availability of the parameters of the cameras, it is not required to conduct a preliminary calibration step because the images are already rectified. Figure 1 (a) shows some screenshots of the model. Figure 1 (b) shows two corresponding stereo images.

2.2. Depth clues extraction

The next step in our analysis involves a stereo recording of a route inside the virtual model. Each rendering is carried out from stereo cameras looking at the scene from two different points of view. The goal is to construct a depth map of the scene from a standard stereo pair acquired by two cameras.

In order to reconstruct the scene we look at the disparity values, i.e., the differences of coordinates of homologous points lying in each image captured by two cameras. Depth is inversely proportional to the disparity, if we represent it such as a gray tone image, brightest pixels correspond to the high values of disparity and consequently to the regions near the cameras. Similarly, the darker pixels represent the deepest ones. By setting focal length and baseline, the depth of a point depends only on the disparity. For the calculation of disparity map we have used the tool in [Lan11]. There are two main motivations for this choice: this method is inspired by an algorithms according to the Middlebury stereo evaluation dataset [Hir07]. It also provides the result to support our experimental concepts.

The main problem is to establish which point in right image is the exact projection of the same point in the left image, otherwise known as the correspondences problem. The matching function used in our experiments is inspired by

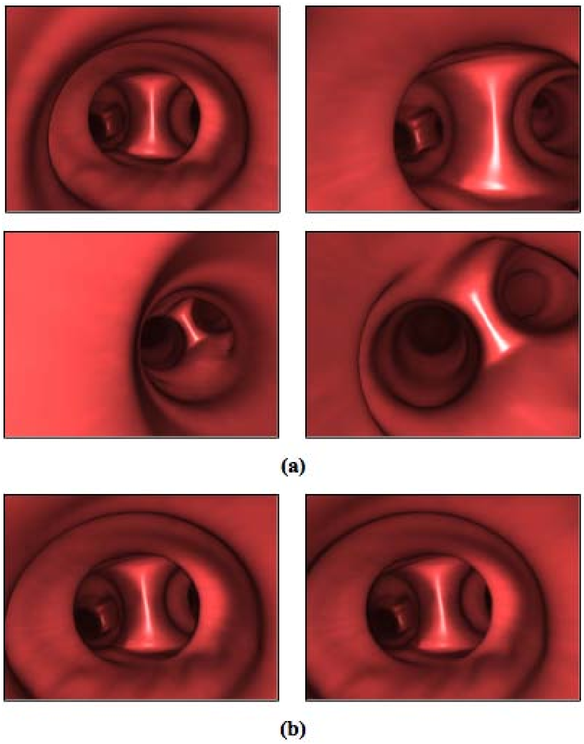

Figure 1: *(a) Examples of images extracted from the proposed virtual model. (b) Couple of stereo images.*

Klaus et al. [KSK06] and consists of a weighted combination of two outcomes: the sum of absolute intensity differences (SAD) and a measure based on gradient determining the disparity by formulating a differential equation which correlates disparity with brightness variations. This match function is defined as:

$$C(x,y,d) = (1-\omega) * C_{SAD}(x,y,d) + \omega * C_{GRAD}(x,y,d) \quad (1)$$

where

$$C_{SAD}(x,y,d) = \sum_{(i,j)\in N(x,y)} |I_1(i,j) - I_2(i+d,j)| \quad (2)$$

and

$$C_{GRAD}(x,y,d) = \sum_{(i,j)\in N_x(x,y)} |\nabla_x I_1(i,j) - \nabla_x I_2(i+d,j)| + \sum_{(i,j)\in N_y(x,y)} |\nabla_y I_1(i,j) - \nabla_y I_2(i+d,j)| \quad (3)$$

$N(x,y)$ is the correlation window at point (x,y), ∇_x and ∇_y are the gradients along the horizontal and vertical directions. $N_x(x,y)$ is a correlation window without the rightmost column, $N_y(x,y)$ a correlation window without the lowest row. The optimal value of disparity d is one that minimizes the match function C. The probability of a wrong match decreases in proportion with the size of the correlation window. A correlation window of size 3×3 pixels is optimal for the reliability of our results. Further parameters needed to estimate disparity assume the default values, as in [Lan11].

Using virtual reality we can easily obtain the field of depth on the rendered images. Then, a qualitative assessment was carried out by comparing our maps with the ground truths depicting real disparities relative to the reference image (left image). Figure 2 shows some examples of depth maps calculated with the proposed method. Ground truths images are also reported.

As shown in Figure 1, depth maps provide an adequate description of the depth of the scene. However, some problems may arise when images have satured and/or textureless regions. In these circumstances, the amount of correct matches decreases and the resulting disparity map contains inconsistent values.

2.3. Augmented reality

There are different ways to exploit depth information. One of this is augmented reality. The final step in our experiments provides for the integration of depth information in the original representation of the scene. All must be optimally developed, in a way that user has the perception of a single scene. To emphasize the effect we take into account depth information in order to meaningfully overlay colors within the image, as proposed in [LMDTM10], [Wil94]. In detail, red color is associated with the pixels with the maximum value of disparity, corresponding to areas of the scene near the cameras; likewise, blue color is associated with the deeper areas. Intermediate disparity values take gradation colors between red and blue. The colors in red-blue range have a strong impact in the operator than other colors because they are conventionally associated to the situations of danger, warning and safety respectively. In the context of the bronchoscopic images, this representation allows many advantages, including to figure out where objects are lying, so they can be easily avoided. Figure 3 shows depth maps integrated in the original representations.

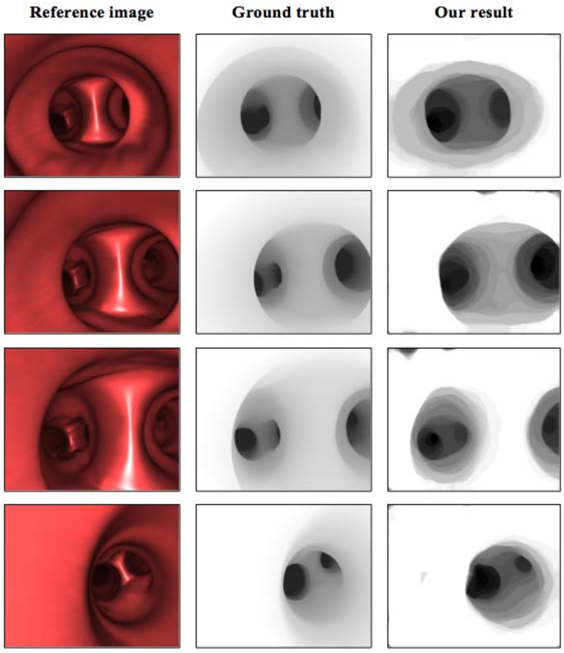

Figure 2: *Some depth maps estimated with the proposed method. The second column shows the ground truths data for the reference images.*

We have conducted some tests to verify the percentual of color information to be overlayed in the images. Final results contain the colors that best support our visual investigation and give a greater sense of depth.

3. Discussion

Our experiments show that the additional information provided by depth maps leads to a better perception of the distances in the scene. This should in turn likely provide a greater precision in the movements of the bronchoscope, minimizing the number of accidental collisions with the bronchial wall during probe navigation. This last feature provides two main benefits: the patient undergoes a less discomfort during the examination. Furthermore, the final video contains only meaningful frames to make a good diagnosis. Although at the present time precise data about the effective-

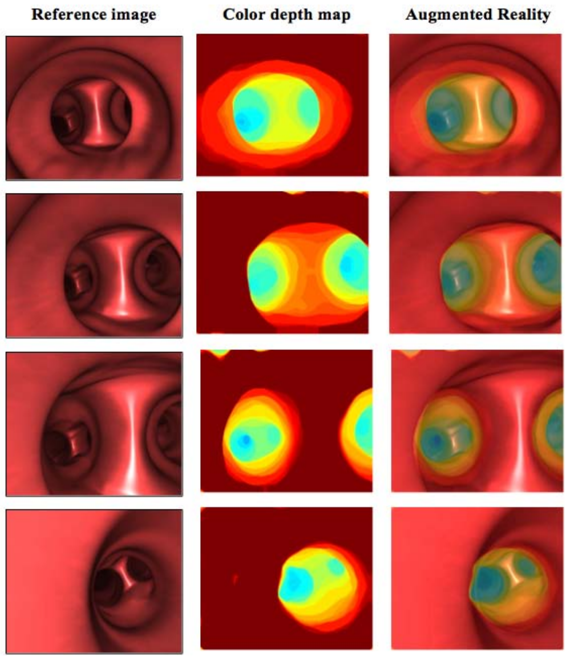

Figure 3: *Color depth maps integrated in the reference images.*

ness of the proposed set-up in reducing unwanted collision are unavailable, we believe that the present study supports the application of stereoscopic vision in bronchoscopic applications.

The colored depth map overlayed on the original representation is only one of the potential augmented-reality visualizations. With this analysis we have to experience with depth map and augmented reality visualization based on color in endoscopic context. However, informations that can be integrated on real bronchoscopic images are several. Hence, augmented reality in bronchoscopic environment can actually provide an useful instruments to overcome surgeon's perceptual skills.

These tasks may be proposed using only information provided by depth-map images. Additional tools can be developed combining depth informations into a tridimensional mesh, using dense surface stereo reconstruction techniques. In this context, a potential application involves the shape reconstruction, in a post-operative scenario, of route taken by the physician during the examination. In this way the expert can analyze which regions have been explored. This reconstruction can also be used for educational purposes to develop training-oriented systems for the simulation of bronchoscopic examinations.

Depth maps in Figure 2 provides a detailed description about the depth of the scene. The high number of correct matches is due to the ideal conditions provided by the virtual environment. In the real case, the situation is most likely more challenging as the bronchoscopic images may present a more articulated or smoother texture, which may make harder to solve the correspondence problem. Defocus regions may also be present due to sudden movements by the operator during the navigation. The situation is worsened by the presence of satured regions of color due to the led light of the surgical probe. In order to obtain the same results proposed in virtual environment, it is necessary to use appropriate denoising image processing algorithms.

Stereo reconstruction in this kind of images is a difficult issue. The complications are due to the nature of the images that often include radial distortion. In addition, the matching function used for extract disparity often rely on the use of epipolar rectified images. To overcome these problems, an accurate calibration step is needed in order to obtain information on the perspective view of the scene and to bring images in a standard stereo form.

The problems listed above can be addressed using appropriate image processing techniques and the proposed approach can therefore successfully be applied to support endoscopic navigation and intervention. A first step in this direction is to develop a real prototype of a flexible bronchoscope whose tip is equipped with two aligned miniature cameras. This technological step is under development at the labs of the "School of Engineering and Technology, University of Hertfordshire".

4. Conclusion

This paper proposed vision-based 3D reconstruction of bronchoscopic images. The experiments have been conducted in a virtual environment using a synthetic 3D model allowing for stereoscopic viewing. The extraction of depth maps was performed which was considered very relevant. A depth map can support endoscopic navigation and it can represent a valid base for the generation of an augmented reality visualscenario that would support surgeons during intervention. The proposed approach can be extended to most endoscopic procedures. Part of our ongoing work includes implementation of the proposed algorithms on a real bronchoscope prototype.

References

[Ble11] BLENDER: free open source 3d content creation suite, 2011. http://www.blender.org. 2

[HFSL*00] HENG P., FUNG P., SAK LEUNG K., QIU SUN H., WONG T.: Virtual bronchoscopy. *The International Journal of Virtual Reality 4*, 4 (2000). 2

[Hir07] HIRSCHMÜLLER H.: Evaluation of cost functions for stereo matching. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2007). 2

[KSK06] KLAUS A., SORMANN M., KARNER K.: Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *Pattern Recognition, 2006.*

ICPR 2006. 18th International Conference on (0-0 2006), vol. 3, pp. 15 –18. 3

[Lan11] LANKTON: Fast 3d stereo vision, 2011. http://www.shawnlankton.com/2008/04/stereo-vision-update-with-new-code. 2, 3

[LBK*05] LEVEN J., BURSCHKA D., KUMAR R., ZHANG G., BLUMENKRANZ S., DAI X., AWAD M., HAGER G. D., MAROHN M., CHOTI M., ET AL.: Davinci canvas: A telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability. In *MICCAI* (2005), Duncan J. S., Gerig G., (Eds.), vol. 3749 of *Lecture Notes in Computer Science*, Springer, pp. 811–818. 1

[LMDTM10] LIVATINO S., MUSCATO G., DE TOMMASO D., MACALUSO M.: Augmented reality stereoscopic visualization for intuitive robot teleguide. In *Industrial Electronics (ISIE), 2010 IEEE International Symposium on* (july 2010), pp. 2828 –2833. 3

[NDI11] NDI: 3d real-time measurement enabling computer-assisted surgery and therapy, 2011. http://www.ndigital.com/medical. 1

[VDMS09] VAN DER MEIJDEN O., SCHIJVEN M.: The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surgical Endoscopy 23* (2009), 1180–1190. 10.1007/s00464-008-0298-x. 2

[Vis10] VISIONSENSE: Stereo vision endoscope, 2010. http://www.visionsense.com. 1

[Wil94] WILLIAMS R.: *The non-designer's design book: design and typographic principles for the visual novice*. Peachpit Press, Berkeley, CA, USA, 1994. 3