

Segmentation and 3D reconstruction approaches for the design of laparoscopic augmented reality environments

P. Sánchez-González , F. Gayá , AM. Cano , and EJ. Gómez

Grupo de Bioingeniería y Telemedicina (GBT), ETSIT, Universidad Politécnica de Madrid
C/ Ciudad Universitaria s/n, 28040 Madrid, Spain

Networking Research Center on Bioengineering, Biomaterials and Nanomedicine
(CIBER-BBN), Madrid, Spain

Abstract. A trend in abdominal surgery is the transition from minimally invasive surgery to surgeries where augmented reality is used. Endoscopic video images are proposed to be employed for extracting useful information to help surgeons performing the operating techniques. This work introduces an illumination model into the design of automatic segmentation algorithms and 3D reconstruction methods. Results obtained from the implementation of our methods to real images are supposed to be an initial step useful for designing new methodologies that will help surgeons operating MIS techniques.

Keywords: Laparoscopic surgery, video analysis, illumination model, segmentation, 3D reconstruction, augmented reality.

1 Introduction

Minimally Invasive Surgery (MIS) has attained a position as the best alternative of some techniques . Opposed to open surgery laparoscopic techniques presents several advantages: smaller and less incisions, quick recovery, fewer post-operative complications and less pain. However, these new techniques introduce some limitations. Thus, surgeons lack depth information due to indirect vision through two-dimensional flat video displays . A possible way to overcome this drawback is to develop new augmented reality (AR) environments to help surgeons performing surgical techniques in the operation theatre. Additional and enhanced information will be provided and fused with the real world in this environment.

The first augmented reality system applied to medicine was intended for neurosurgery . MIS AR environments do not appear until the end of the 1990s . Last proposals combine preoperative images (CT and MR are the best alternatives) and intraoperative ones (CT, MR and US). A new alternative is to use endoscopic images as intraoperative information source. This is the approach chosen in this research: the analysis of endoscopic video images for extracting useful information that can be focused on surgical scene (organs, veins...) or on surgical tools.

MIS AR environments require the detachment, identification and posterior labeling of anatomical structures in the scene. Thus, it is needed an initial segmentation process of the surgical image. Endoscopic images features complicate the automation of this step. Among other features that difficult this task there are the lack of sharp borders, the transparencies and the color variability of each structure due to textures (then, a superposition of the color range exists between anatomical structures).

In addition, one of the major challenges for laparoscopic visualization is acquisition of the depth map associated with the image from the endoscopic camera. Determination of 3D scene structure from a sequence of 2D images is one of the classic problems in computer vision. Thus, lots of researchers had worked on this problem, developing numerous techniques for computing 3D structure. These techniques usually use cues from motion, stereo, shading, focus, defocus, contours, etc. Endoscopic images features such as the absence of straight lines or planes complicate considerably reconstruction tasks. So, although there is a multitude of reconstruction approaches, it is difficult to find one that can cope with all situations. Some of their limitations are: (1) Most of existing techniques assume some physical constraints (e.g. shape from shading are used to assume the illumination and reflectance remain constant through a scene); (2) Problems in the accurate recovery of surface structure due to occlusions or camouflages (absent zones or no detected ones); (3) Computational complexity because most of the methods are iterative ones;(4) Limited application to defined textures; (5) Limited application to continuous surfaces (e.g. have to be planes).

In this research we address both challenges (3D reconstruction and automatic segmentation) based on the study of reflected lights in the scene (analyzed in the proposed illumination model) (Fig.1).

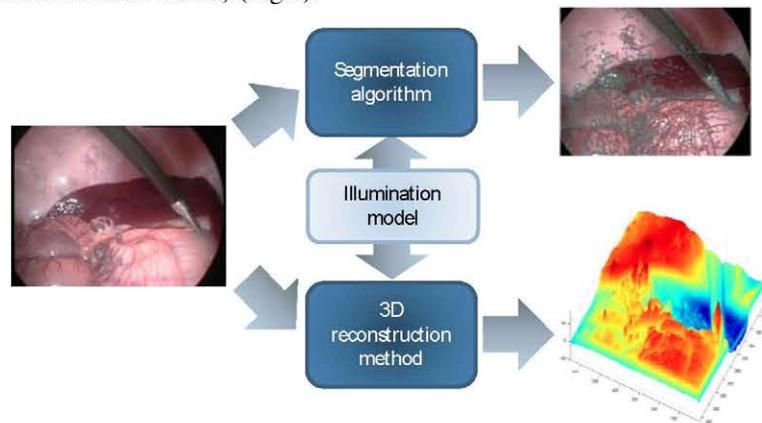


Fig. 1. Applications of the illumination model

2 Methods

This section starts with the study of the surgical scene illumination equation (2.1), and describes the two image processing algorithms designed and developed in this research: segmentation (2.2) and 3D reconstruction (2.3).

2.1 Illumination model

In this section a study of the light reflected is made. Moreover, variables that have influence on the amount of light reflected off a 3D surface to a point of observation are analyzed.

The illumination model initial hypothesis are: (1) Laparoscopic images are characterized by the fact that light source and caption system are located at the same spatial point, the endoscope tip (see fig. 2); (2) Organs surfaces are considered as lambertian ones in surgical scenes (light is scattered equally in all directions); and (3) Indirect light intensity is considered negligible.

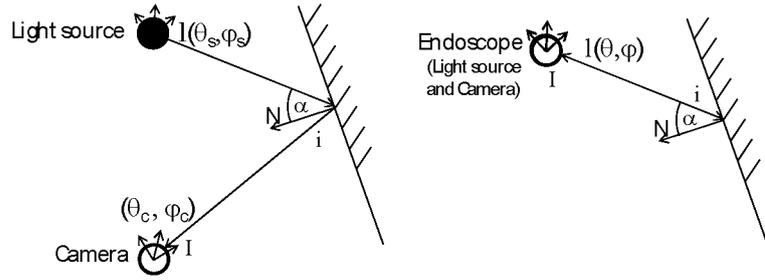


Fig. 2. General (left) and endoscope (right) scenery

Considering these hypothesis, the luminous intensity of each image pixel is computed. Let's $I(\theta_s, \varphi_s)$ be the emitted light spatial distribution in a spherical coordinates system origin located in the light source (see Fig.2). In this situation, the light emitted power L per solid angle ($[\theta_{s1}, \theta_{s2}]$, $[\varphi_{s1}, \varphi_{s2}]$) is calculated like:

$$L = \int_{\theta_{s1}}^{\theta_{s2}} \int_{\varphi_{s1}}^{\varphi_{s2}} I(\theta_s, \varphi_s) \cdot d\varphi_s \cdot d\theta_s \quad (1)$$

As mentioned in hypothesis (2), an organ is treated as a lambertian surface, which intercepts the light and reflects it. The reflected light power is:

$$i = k_d \cdot l \cdot N = k_d \cdot |l| \cdot |N| \cdot \cos \alpha \quad (2)$$

where k_d is a surface intrinsic constant and α is the angle between the light source vector (l) and the surface normal one (N) (see Fig.2).

Camera captures this reflected power. Thus, supposing the coordinates origin located in the camera focal point, power measured, per pixel in (θ_c, φ_c) direction, is:

$$I = \int_{\theta_{c1}}^{\theta_{c2}} \int_{\varphi_{c1}}^{\varphi_{c2}} i \cdot d\varphi \cdot d\theta \quad (3)$$

Based on hypothesis (1), the two mentioned coordinates systems are fitted in the same spatial point. Therefore, equation (2) can be substituted into equation (3), which results in the simplified illumination equation of the surgical scene:

$$I = k_d \cdot L \cdot \cos \alpha \quad (4)$$

where k_d is a surface intrinsic constant, L is the emitted power per unit solid angle $([\theta_{s1}, \theta_{s2}], [\varphi_{s1}, \varphi_{s2}])$, and α is the angle between the light source vector and the surface normal one. This equation asserts that light intensity received in one pixel is independent of the surface-endoscope distance. This fact implies that it is not possible to determine absolute objects distance based on solely illumination information. It causes that two homothetic surfaces with similar reflectivity are undistinguished.

2.2 Segmentation algorithm

The selected image segmentation algorithm has to produce a segmentation map. This segmentation has to be totally automatic, so methods based on manual initialization (semiautomatic methods) are not suitable for our purpose. There are many features that can be used as criteria for image segmentation. For instance, color, intensity, texture and motion are the most used. Among laparoscopic image features, intensity is considered to be the most suitable criterion in our segmentation algorithm. Due to the existence of both small and large anatomical structures, split and merge methods are well suited. However, pure merging methods are computationally expensive because they start from too small initial regions (individual pixels). So, our approach should start from a good splitting, in order to obtain the minimum number of regions as possible. Taking in mind these considerations, an automatic split and merge method is designed based on the illumination model previously explained.

A common technique for segmentation is to use gradient watershed regions, which will be used as part of the splitting stage. First a gradient magnitude image has to be built, and then watershed regions are found in this image. In this step of the process, the illumination model is used. According to the illumination equation, the image logarithm is:

$$\text{Ln}I = \text{Ln}k_d + \text{Ln}L + \text{Ln} \cos \alpha \quad (5)$$

$\text{Ln}L$ term is invariant. As far as the term $\text{Ln} \cos \alpha$ is concerned, it varies in a slowly way through neighbour pixels in a same region. Similarly, k_d changes abruptly when a region ends. It makes the gradient of the logarithm be highest in region changes, which is suitable for a watershed segmentation method.

The major problem with this segmentation is that, due to the fact that k_d varies inside the organs, a certain over-segmentation is obtained. Then, we need a merging step in the segmentation process. In our merging development, Hotelling's T-square (or Mahalanobis distance) are used as the neighbour regions (or border pixels) merging metrics.

2.4 3D reconstruction method

The reconstruction method has to obtain a relief of the surgical scene from a monoscopic image (it is known by equation (4) that it is not possible to calculate absolute distances). Our approach is based on solely the illumination model. Then, it is questioned if the shape of an object can be reconstructed through measuring the light reflected by points form organs' surface.

The problem approach is shown in fig. 3. Considering α as the angle between the light source vector (L) per solid angle (θ, φ) and the surface normal vector (N), depth variation through pixel p (Δz) is proportional to the tangent of this angle. In this way, knowing this tangent, depth variation between two neighbour pixels can be obtained.

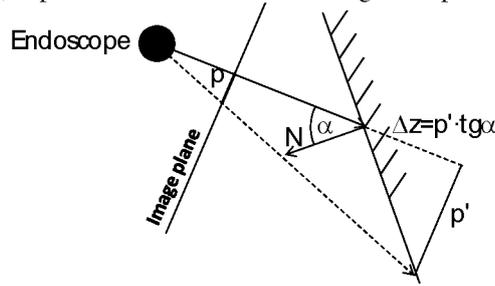


Fig. 3. Variation of the distance to the endoscope according to the tangent of the angle (α) between the visual (θ, φ) and the surface normal (N)

Thus, knowing intensity slopes (tangents of α), it is possible to calculate the surfaces' shapes adding all the slopes in the gradient direction. These intensity slopes can be calculated in a two steps process:

1. Intensity gradient calculation of each pixel in a given direction of the image.
2. Division of the intensity gradient by the intensity in the direction given.

The result of this process is equal to the gradient of the image logarithm, which is the same input to watershed algorithm explained before (2.2).

Angle α is, by definition, positive. The gradient of cosine α logarithm is:

$$d \ln \cos \alpha = -\frac{\sin \alpha}{\cos \alpha} \cdot d\alpha = -\operatorname{tg} \alpha \cdot d\alpha \quad (6)$$

The resulting sign of the equation does not originates in the $\sin \alpha$ term (which is positive), but from $d\alpha$. Therefore, it indicates the fact that α increases its value (negative sign) or decreases it (positive sign). Supposing α varies monotonically, depth variation can be accumulated, although solely depth variation absolute value is obtained.

3 Preliminary results and discussion

A new illumination model has been proposed for the laparoscopic images analysis. This illumination model has been used to design and develop both a segmentation algorithm and a reconstruction method.

The algorithms proposed are implemented in Matlab and computed in a Pentium Intel Core 2 Duo T7100 (2GHz RAM). Laparoscopic images were supplied by the Minimally Invasive Surgery Centre (Cáceres, Spain). Images were captured by an endoscope consisting of an Endocam 5506 camera (Richard Wolf) and an optical system Circon-ACMI 0°. Typical size of images processed is 576x768 pixels.

Sample preliminary result of laparoscopic image segmentations is shown in Fig. 4. Splitting step based on the proposed illumination model results in a start point specially adapted to laparoscopic images, distinguishing between big organs (e.g. liver) and anatomic details such as vases.

In addition, an important feature is that abrupt slope variations are not present in the surgical scene. It facilitates segmentation step due to the term $\cos \alpha$ remains approximately constant.

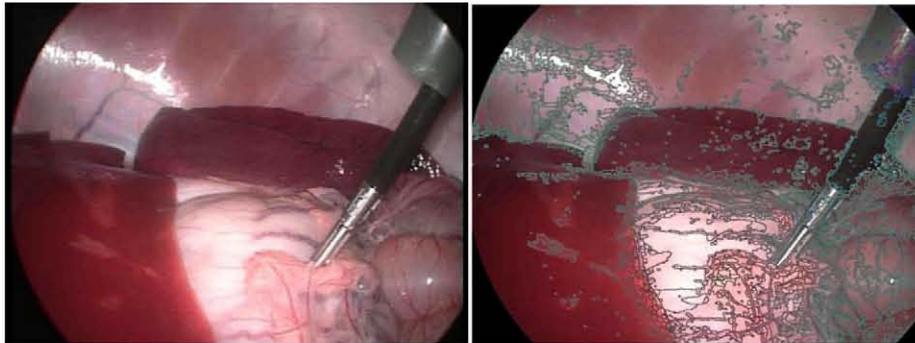


Fig. 4. Segmentation preliminary results (Initial image and split image)

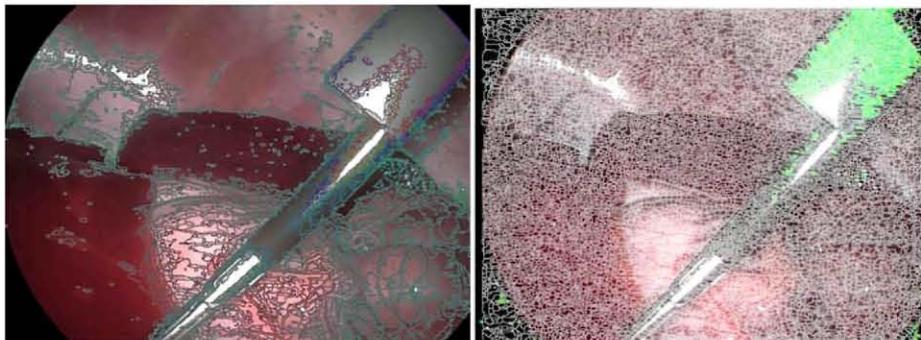


Fig. 5. (Left) Segmentation with the algorithm proposed, (Right) Watershed segmentation

Although results obtained in this study show that merging stage should be improved using other additional metrics, accuracy achieved in laparoscopic surgery images (Fig.5-Left) is higher than obtained by classic automatic segmentation methods such as typical watershed (Fig.5-Right). Moreover, with classical watershed techniques, computational cost increases considerably in merging stage (due to the great difference in number of regions after the splitting: from 3417 with our approach to 15488 with simple watershed, for image analyzed).

In addition to the previous algorithm, a 3D reconstruction method of the surgical scene has been proposed based on the illumination model. In its design, a simple light source has been considered. Nevertheless, light reflections are presented in the scene and in some situations they can be comparable to direct illumination. As red is the principal reflected color component by organs and blue and green reflected components are minimum they can be used for the developments.

Results (see Fig.6) show that although shape obtained is in a relative scale (illumination equation shows that it is not possible to extract absolute distances between organs and the endoscope through a solely use of illumination information), its main limitation is the uncertainty of the slope sign (concave/convex ambiguities).

The relative distances problem is a minor one: knowing one real depth value per continuous region and the sign of slope, it can be propagated obtaining the shape and real distance between surfaces and the endoscope.

Both problems could be addressed mixing information provided by other shape-from-X techniques such as motion (based on the matching of features from different frames) that can provide absolute distances or textures which can deal with the ambiguity of slope sign. Shape-from-shadows provides no information due to the absence of shadows, and the monoscopic nature of laparoscopic images does not allow stereo approaches.

Simplicity and low computational cost are two important characteristics of the reconstruction method proposed. Moreover, thanks to the local characteristic of this method (it uses gradient as a local operator) it is not affected by the distortions correction processing, which can be made before or after the reconstruction.

Moreover, a more complex evaluation should be done. For this evaluation a new surgical video sequence should be acquired using an optical marker, in order to get the real shape of the surface, and then compare it with our results.

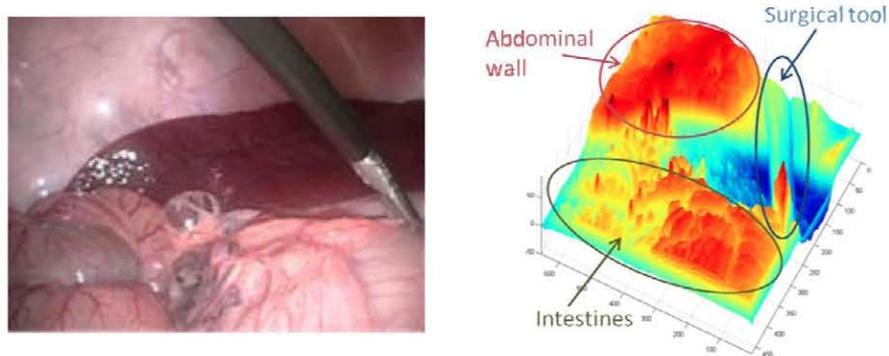


Fig.6. Original image and its depth map

4 Conclusions

Laparoscopic video analysis is a good alternative to enhance intraoperative information in order to help surgeons performing the operating techniques. In this research an illumination model has been proposed and used for designing a segmentation algorithm and a 3D reconstruction method. These are two of the main steps in the development of augmented reality environments.

In this paper it has been shown how this illumination model improves segmentation step comparing to traditional segmentation methods. Thus, accuracy and computational cost are improved.

Moreover, results obtained from application of our reconstruction method to real images show that there are some limitations that make difficult its real application. However, it can be a good starting point to design new 3D reconstruction methods that should combine illumination, texture and motion information in order to reach the final goal: to design and develop new efficient augmented reality environments, without including complex systems in the operating room.

References

- A. Cuschieri, Laparoscopic surgery: current status, issues and future developments, *Surgeon* 3, 125–30, 132–3, 135–8. (2005).
- Usón J, Pascual S, Sanchez FM. Aprendizaje y formación en cirugía laparoscópica. Bilbao E, Pascual S (eds.), *Cirugía laparoscópica del Reflujo Gastroesofágico. Técnica de Nissen*, Centro de Cirugía de Mínima Invasión, Cáceres, 27-61 (2002).
- K. H. Fuchs, *Minimally Invasive Surgery, Endoscopy*; 34: 154-159. (2002).
- Kelly MD, P. J., Kall, B., and Goerss, S. Computer-assisted stereo-taxic resection of intra-axial brain neoplasms. *Journal of Neurosurgery*, 64:427-439. (1986).
- D. Roberts, J. Strohbehn, J. Hatch, W. Murray, and H. Kettenberger, A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope., *J Neurosurg*, 65 pp. 545–9. (1986).
- W. Freysinger, A. Gunkel, and W. Thumfart, Image-guided endoscopic ent surgery, *European archives of Otorhinolaryngology*, 254 pp. 343–346. (1997).
- P. Sánchez González, F. Gayá Moreno, A.M. Cano González, E.J. Gómez Aguilera, Reconstrucción tridimensional en imágenes laparoscópicas a partir de la iluminación de la escena quirúrgica. XXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), Cartagena, pp. 463-366. (2007).
- Thomas Fromherz and Martin Bichsel, Shape from Multiple Cues: Integrating Local Brightness Information, *Fourth International Conference for Young Computer Scientist, ICYCS 95*, Beijing, P.R. China, (1995).
- Cavanagh, P., Reconstructing the Third Dimension: Interactions Between Color, Texture, Motion, Binocular Disparity, and Shape, *CVGIP(37)*, No. 2, pp. 171-195. (1987).