



**HAL**  
open science

## **From 2D towards 3D cartography of hollow organs**

Christian Daul, Walter Blondel, Achraf Ben Hamadou, Rosebet Miranda, Charles Soussen, Didier Wolf, François Guillemin

### ► **To cite this version:**

Christian Daul, Walter Blondel, Achraf Ben Hamadou, Rosebet Miranda, Charles Soussen, et al.. From 2D towards 3D cartography of hollow organs. 7th IEEE International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE 2010, Sep 2010, Tuxtla Gutiérrez, Mexico. pp.285-293. <hal-00517538>

**HAL Id: hal-00517538**

**<https://hal.science/hal-00517538v1>**

Submitted on 14 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## From 2D towards 3D cartography of hollow organs

C. Daul, W.P.C.M. Blondel, A. Ben-Hamadou, R. Miranda-Luna, C. Soussen, D. Wolf, F. Guillemin  
Centre de Recherche en Automatique de Nancy (CRAN, UMR 7039, CNRS Nancy - University),  
2, avenue de la Forêt de Haye, F-54516 Vandœuvre-Les-Nancy, France.  
E-mail: christian.daul@ensem.inpl-nancy.fr

**Abstract**—Endoscopy is a standard imaging modality commonly used in different medical fields like lesion diagnosis in hollow organs or mini-invasive surgery. Meanwhile, endoscopic data suffer from the fact that each image of a video-sequence only corresponds to a small 2D field of view. This paper presents a mosaicing algorithm leading to visually coherent large field of view maps. The ability of the algorithm to build 2D bladder maps is assessed with both phantom and patient data. This contribution describes also a 3D cartography method and gives preliminary surface reconstruction results on phantoms with bladder textures.

**Index Terms**—2D image mosaicing, 3D surface construction, 3D cartography, bladder, endoscopy.

### I. INTRODUCTION

In hollow organs of the human body (*e.g.*, bladder, oesophagus, colon), pathological lesions located on their inner wall can be detected in video-sequences acquired with cystoscopes (rigid endoscopes) or fiberscopes (fibered flexible cystoscopes). Meanwhile, each image of the sequence visualizes only a small surface of the organ (with corresponding areas smaller than  $1 \text{ cm}^2$ ), while lesions are typically spread over wide areas *i.e.* several images. The lesion diagnosis and tissular status follow up of lesions is difficult since clinicians have to mentally reconstruct the scene using the image sequence (*i.e.* they cannot observe at one go the whole region of interest). In endoscopic video-sequences, consecutive images (acquired from different viewpoints) are always partially overlapped. Image mosaicing techniques (registration of consecutive image pairs) can be used to build maps (large field of views including the whole region of interest) facilitating the lesion diagnosis.

Although image mosaicing was used in numerous fields (astrophysics, consumer photography, remote sensing, *etc.*), only few studies were devoted to the cartography of medical data. Mammography [1], ophthalmology [2], microscopy [3] and X-ray angiography [4] are among the few examples in which image mosaicing has been used. These approaches are either not fully automated or based on the use of *a priori* knowledge concerning the geometrical transformation between the images. Such knowledge is not available for endoscopic sequences since clinicians freely move the endoscope in the hollow organs. Moreover, these algorithms were not conceived to register a large number of images (typically, thousands of images have to be registered for bladders). Apart from the present contribution, the only other publication dealing with image mosaicing of hollow organs describes a cartography method for fluorescence bladder endoscope images [5],

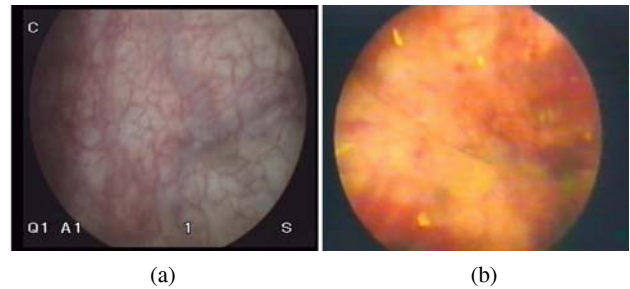


Fig. 1. Bladder images extracted from cystoscopic acquisitions of two different patients. (a) The blood vessel contours can be segmented. (b) No significant image primitive (contour, corner, *etc.*) can be extracted from the image.

[6]. This work exploits the fact that, in fluorescence endoscopy, image primitives can be segmented and used as homologous structures during the data registration. For the more widespread (or standard) white light modality, image primitives cannot usually be extracted in a robust way for hollow organs. As sketched in Fig. 1, the intra- and inter-patient bladder texture variability is very strong and images do not necessarily include significant primitives (contours, edges, corners). To be as robust and general as possible, a mosaicing algorithm of hollow organ images shall not rely on the segmentation of image primitives.

The next section describes a bidimensional (2D) cartography method leading in a robust way to 2D wide field of views. In Section III, a tridimensional (3D) reconstruction algorithm is sketched showing that it is possible to build large 3D field of views of hollow organs with data of modified endoscopes. The results presented in Section IV first show the potential of 2D bladder cartography based on both phantom and patient data, and illustrate the feasibility of 3D bladder cartography.

### II. 2D CARTOGRAPHY ALGORITHM

The registration process used to superimpose the overlapping part (*ov*) of the source image  $I_{k+1}^{ov}$  on the target image  $I_k$  is given in (1), these images being grey level data obtained from consecutive ( $k, k+1$ ) color images of the video-sequence.

$$\tilde{T}_{2D}^{k,k+1} = \arg \max_{T_{2D}^{k,k+1}} MI \left( I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}) \right) \quad (1)$$

$\tilde{T}_{2D}^{k,k+1}$  is the geometrical transformation superimposing  $I_{k+1}^{ov}$  on the target image  $I_k^{ov}$ . The registration process optimizes the parameters of  $T_{2D}^{k,k+1}$  such as the  $MI$  (mutual information, see [7]) measure reaches a maximum. The  $I_k^{ov}$  images do not

include barrel distortion and light inhomogeneities since these radial distortions and shading effects affecting endoscopic data are corrected with the algorithms described in [8], [9]. Moreover, as detailed in [10], the bladder being filled up with physiological serum during an examination, and due to the high acquisition rate (25 images/s), the bladder can be considered as being rigid between two acquisitions  $k$  and  $k + 1$ . Consequently,  $T_{2D}^{k,k+1}$  is a  $3 \times 3$  homogeneous perspective matrix corresponding to a viewpoint difference and superimposing the  $i$ -th pixel  $P_{2D}^{i,k+1}$  with coordinates  $(x_{2D}^{i,k+1}, y_{2D}^{i,k+1})^T$  in  $I_{k+1}$  on the  $i$ -th (homologous) point  $P_{2D}^{i,k}$  with coordinates  $(x_{2D}^{i,k}, y_{2D}^{i,k})^T$  in  $I_k$ . In (2), the six parameters  $a_{rs}$  (with  $r \in [1, 2]$  and  $s \in [1, 3]$ ) depend on the  $K_x$  and  $K_y$  scale factors, the  $S_x$  and  $S_y$  shearing parameters, and the 2D in plane rotation  $\varphi$  angle.  $a_{31}$  and  $a_{32}$  correspond to the perspective change due to out of plane rotations and  $a_{33} = 1$ .

$$\begin{pmatrix} \beta^{i,k} x_{2D}^{i,k} \\ \beta^{i,k} y_{2D}^{i,k} \\ \beta^{i,k} \end{pmatrix} = \begin{pmatrix} \underbrace{K_x \cos(\varphi)}_{a_{11}} \underbrace{-S_x \sin(\varphi)}_{a_{12}} t_x & a_{13} \\ \underbrace{S_y \sin(\varphi)}_{a_{21}} \underbrace{K_y \cos(\varphi)}_{a_{22}} t_y & t_y \\ a_{31} & a_{32} & 1 \end{pmatrix} \begin{pmatrix} x_{2D}^{i,k+1} \\ y_{2D}^{i,k+1} \\ 1 \end{pmatrix} \quad (2)$$

The  $MI$  mutual information is computed from the grey level entropies  $H_k(I_k^{ov})$  and  $H_{k+1}(T_{2D}^{k,k+1}(I_k^{ov}))$  of the overlapping parts of images  $I_k^{ov}$  and  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$  and the joint grey level entropy  $H_{k,k+1}(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}))$ . In (3),  $g_k$  and  $g_{k+1}$  are the grey levels of  $I_k^{ov}$  and  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$ , respectively.  $p(g_k)$  and  $p(g_{k+1})$  are the grey level probability density functions of  $I_k^{ov}$  and  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$  and  $p(g_k, g_{k+1})$  is a joint probability density function. The grey levels range in  $[g_{min}, g_{max}]$  (typically,  $g_{min} = 0$  and  $g_{max} = 255$ ).

$$MI(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov})) = H_k(I_k^{ov}) + H_{k+1}(T_{2D}^{k,k+1}(I_{k+1}^{ov})) - H_{k,k+1}(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov})) \quad (3)$$

with

$$\begin{cases} H_k(I_k^{ov}) = - \sum_{g_k=g_{min}}^{g_{max}} p(g_k) \ln(p(g_k)) \\ H_{k+1}(T_{2D}^{k,k+1}(I_{k+1}^{ov})) = - \sum_{g_{k+1}=g_{min}}^{g_{max}} p(g_{k+1}) \ln(p(g_{k+1})) \\ H_{k,k+1}(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov})) = \\ - \sum_{g_k=g_{min}}^{g_{max}} \sum_{g_{k+1}=g_{min}}^{g_{max}} p(g_k, g_{k+1}) \ln(p(g_k, g_{k+1})) \end{cases}$$

Two images are registered when the  $MI$ -value is maximal, *i.e.*, when the grey levels of the homologous (superimposed) pixels of  $I_k^{ov}$  and  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$  are statistically as similar as possible. The  $a_{rs}$  parameters of  $T_{2D}^{k,k+1}$  (see (2)) are determined with a stochastic steepest gradient method. In (4), the  $a_{rs}$  parameters and the  $\lambda_{rs}^j$  convergence step are updated for each  $j$  iteration of the optimization process.

$$a_{rs}^{j+1} = a_{rs}^j + \lambda_{rs}^j \frac{\partial MI(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}))}{\partial a_{rs}^j} \quad (4)$$

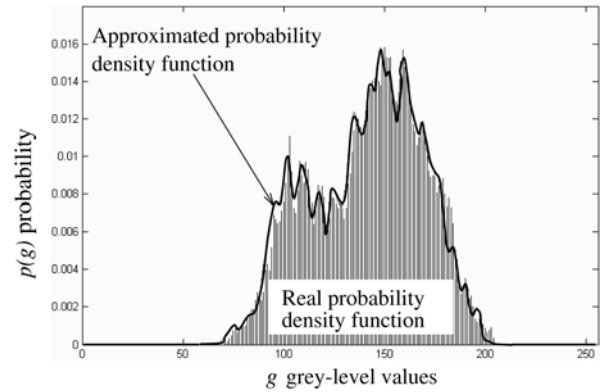


Fig. 2. Parzen's window for a bladder image coded on 256 grey level values (real and approximated probability density functions).

The mutual information cannot be computed exactly in an analytical way, because it depends on the discrete entropies  $H_k(I_k^{ov})$ ,  $H_{k+1}(T_{2D}^{k,k+1}(I_k^{ov}))$ , and  $H_{k,k+1}(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}))$ . As a consequence, the evaluation of the  $\partial IM / \partial a_{rs}$  partial derivatives is not an easy task. A first possible solution is to compute the derivatives numerically with a finite difference scheme. An alternative is to use an analytical approximation of the 1D and 2D entropies involved in the mutual information computation. We chose the second strategy since it directly leads to the computation of the  $\partial IM / \partial a_{rs}$  partial derivatives, and it is known to be more robust [11]. The  $p(g_k)$ ,  $p(g_{k+1})$  and  $p(g_k, g_{k+1})$  probability density functions thus need to be analytically approximated. Fig. 2 illustrates that the grey level histograms of bladder images consist of a sequence of peaks (or modi). As detailed in the following, such multimodal functions can be modeled as a sum of Gaussian functions computed using the Parzen's window method.

In the proposed implementation of Parzen's window, all Gaussian curve approximating a probability density function have the same standard deviation.  $\sigma_k$ ,  $\sigma_{k+1}$  and  $\sigma_{k,k+1}$  are the standard deviations to be determined during the optimisation for  $p(g_k)$ ,  $p(g_{k+1})$  and  $p(g_k, g_{k+1})$ , respectively. In practice, the  $p(g)$  probability density functions (computed with all grey levels) are replaced by the  $p^*(g)$  empirical probability density functions determined with two samples randomly chosen in the overlapping parts of  $I_k^{ov}$  (sample  $A$ ) and  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$  (sample  $B$ ).  $A$  and  $B$  include  $N_A$  and  $N_B$  pixels.  $g_A^a$  and  $g_B^b$  are the grey levels of the  $a$ -th and  $b$ -th pixel of  $A$  and  $B$ , respectively. In the following, the  $im$  subscript symbol refers either to the  $k$  target image ( $I_k^{ov}$ ) or to the  $k + 1$  transformed source image ( $T_{2D}^{k,k+1}(I_{k+1}^{ov})$ ). As mathematically formulated in (5), a  $p^*(g_{B,im}^b)$  empirical probability corresponds to the sum of values computed for grey level  $g_{B,im}^b$  with the Gaussian functions centered on the  $g_{A,im}^a$  acting as mean values.

$$p^*(g_{B,im}^b) = \frac{1}{N_A^{im}} \sum_{a=1}^{N_A^{im}} \left( \frac{1}{\sqrt{2\pi}\sigma_{im}} e^{-\frac{(g_{B,im}^b - g_{A,im}^a)^2}{2\sigma_{im}^2}} \right) \quad (5)$$

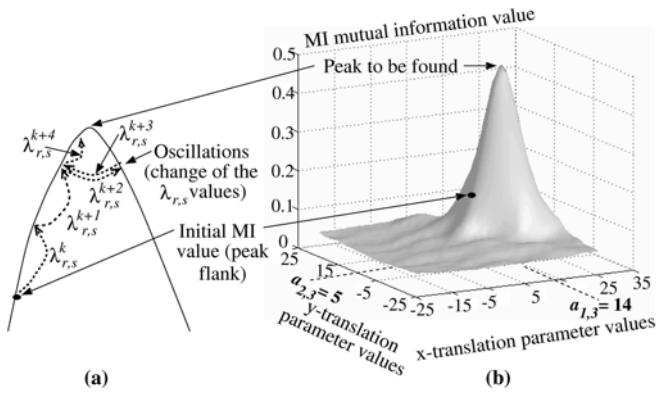


Fig. 3. Convergence step evolution. (a)  $\lambda_{r,s}^k = \lambda_{r,s}^{k+1} = \lambda_{r,s}^{k+2} = \lambda_{r,s}^{k+3}$ . After the oscillation detection, the  $\lambda_{r,s}^{k+4}$  parameters equal 85% of the  $\lambda_{r,s}^{k+3}$  parameters. (b) Mutual information computed for two bladder images. In this parameter space, the two translation parameters of the perspective registration matrix are represented. The top of the peak is localized in  $a_{1,3} = 14$  pixels and  $a_{2,3} = 5$  pixels.

Equation (5) is used to determine the  $p^*(g_{B,k}^b)$  probabilities with the  $g_{A,k}^a$  and  $g_{B,k}^b$  grey levels of the  $A_k$  and  $B_k$  samples consisting of  $N_A^k$  and  $N_B^k$  pixels randomly chosen in the  $I_k$  target image. Equation (5) is also employed to compute the  $p^*(g_{B,k+1}^b)$  probabilities of the transformed source image. However, the  $g_{A,k+1}^a$  and  $g_{B,k+1}^b$  grey levels are not randomly chosen in  $T_{2D}^{k,k+1}(I_{k+1}^{ov})$ , but are the homologous points of the pixels selected in  $I_k^{ov}$ . Consequently,  $N_A^{k+1} = N_A^k$  and  $N_B^{k+1} = N_B^k$ . The  $A_k$ ,  $B_k$ ,  $A_{k+1}$  and  $B_{k+1}$  samples are also used to compute the joint empirical probability function.  $p^*(g_{B,k}^b, g_{B,k+1}^b)$  is obtained with (5) in which the sum and the 1D Gaussian function have simply to be replaced by a double sum (one sum over the  $g_{A,k}^a$  and one sum over the  $g_{A,k+1}^a$ ) and a 2D Gaussian function, respectively.

In practice, the image entropies and the joint entropy can be approximated by the  $H_k^*(I_k^{ov})$  and  $H_{k+1}^*(T_{2D}^{k,k+1}(I_{k+1}^{ov}))$  empirical image entropies and the  $H_{k,k+1}^*(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}))$  joint empirical entropy. Equation (6) allows for the computation of the empirical entropies of the target and transformed source images ( $im$  stands again for  $k$  or  $k+1$ ). The joint empirical entropy can also be determined with (6) by replacing the sum and the 1D Gaussian function by a double sum and a 2D Gaussian function.

$$H_{im}^*(I_{im}^{ov}) = -\frac{1}{N_B^{im}} \sum_{b=1}^{N_B^{im}} \ln \left( \frac{1}{N_A^{im}} \sum_{a=1}^{N_A^{im}} e^{-\frac{(g_{B,im}^b - g_{A,im}^a)^2}{2\sigma_{im}^2}}}{\sqrt{2\pi}\sigma_{im}} \right) \quad (6)$$

The maximization of the approximated probability density functions leads to the minimization of the entropies [7]. The registration algorithm performs, at each  $j$  iteration, two sequential steps:

- 1) minimization of the empirical entropy values, and
- 2) maximization of the mutual information value.

The empirical entropies are minimized by adjusting the standard deviations of the Gaussian functions using the derivatives

of the analytical functions of  $H_k(I_k^{ov})$ ,  $H_{k+1}(T_{2D}^{k,k+1}(I_{k+1}^{ov}))$  and  $H_{k,k+1}(I_k^{ov}, T_{2D}^{k,k+1}(I_{k+1}^{ov}))$  with respect to  $\sigma_k$ ,  $\sigma_{k+1}$  and  $\sigma_{k,k+1}$  (see (7)).

$$\begin{cases} \sigma_k^{j+1} = \sigma_k^j - \lambda_k^j \frac{\partial H_k^*(I_k^{ov})}{\partial \sigma_k^j} \\ \sigma_{k+1}^{j+1} = \sigma_{k+1}^j - \lambda_{k+1}^j \frac{\partial H_{k+1}^*(T_{2D}^{k,k+1,j}(I_{k+1}^{ov}))}{\partial \sigma_{k+1}^j} \\ \sigma_{k,k+1}^{j+1} = \sigma_{k,k+1}^j - \lambda_{k,k+1}^j \frac{\partial H_{k,k+1}^*(I_k^{ov}, T_{2D}^{k,k+1,j}(I_{k+1}^{ov}))}{\partial \sigma_{k,k+1}^j} \end{cases} \quad (7)$$

$$\text{with } \frac{\partial H_{im}^*(I_{im}^{ov})}{\partial \sigma_{im}} = -\frac{1}{N_B^{im}} \sum_{b=1}^{N_B^{im}} e^{-\frac{(g_{B,im}^b - g_{A,im}^a)^2}{2\sigma_{im}^2}} \frac{(g_{B,im}^b - g_{A,im}^a)^2}{\sigma_{im}^2} - 1 \Bigg) \frac{1}{N_A^{im} \sqrt{2\pi} \sigma_{im}^2 P_{im}^*(g_{B,im}^b)}$$

The mutual information is maximized with the derivatives of the empirical entropies with respect to the  $a_{r,s}$  parameters given in (2). In (8),  $\partial g / \partial a_{r,s}$  corresponds to the grey level variations with respect to the  $a_{r,s}$  value modifications.  $G_x$  and  $G_y$  are the components along  $\vec{x}$  and  $\vec{y}$  of the grey level gradient of the point with coordinates  $(x, y)$  in  $T_{2D}^{k,k+1,j}(I_{k+1}^{ov})$ .

$$a_{r,s}^{j+1} = a_{r,s}^j + \lambda_{r,s}^j \frac{\partial H_{k+1}^*(T_{2D}^{k,k+1,j}(I_{k+1}^{ov}))}{\partial a_{r,s}^j} - \lambda_{r,s}^j \frac{\partial H_{k,k+1}^*(I_k^{ov}, T_{2D}^{k,k+1,j}(I_{k+1}^{ov}))}{\partial a_{r,s}^j} \quad (8)$$

with

$$\begin{cases} \frac{\partial H_{k+1}^*(T_{2D}^{k,k+1,j}(I_{k+1}^{ov}))}{\partial a_{r,s}^j} = -\frac{1}{N_B^{k+1}} \sum_{b=1}^{N_B^{k+1}} \left[ \frac{1}{N_A^{k+1} \sqrt{2\pi} \sigma_{k+1}^2 P^*(g_{B,k+1}^b)} \sum_{a=1}^{N_A^{k+1}} \left( - (g_{B,k+1}^b - g_{A,k+1}^a) e^{-\frac{(g_{B,k+1}^b - g_{A,k+1}^a)^2}{2\sigma_{k+1}^2}} \left( \frac{\partial g_{B,k+1}^b}{\partial a_{r,s}^j} - \frac{\partial g_{A,k+1}^a}{\partial a_{r,s}^j} \right) \right) \right] \\ \frac{\partial g}{\partial a_{r,s}^j} = \begin{bmatrix} \frac{G_x x}{a_{31}x + a_{32}y + 1} & \frac{G_x y}{a_{31}x + a_{32}y + 1} & \frac{G_x}{a_{31}x + a_{32}y + 1} \\ \frac{G_y x}{a_{31}x + a_{32}y + 1} & \frac{G_y y}{a_{31}x + a_{32}y + 1} & \frac{G_y}{a_{31}x + a_{32}y + 1} \\ \frac{\beta x}{(a_{31}x + a_{32}y + 1)^2} & \frac{\beta y}{(a_{31}x + a_{32}y + 1)^2} & \frac{\beta}{(a_{31}x + a_{32}y + 1)^2} \end{bmatrix} \\ \beta = G_x(a_{11}x + a_{12}y + a_{13}) + G_y(a_{21}x + a_{22}y + a_{23}) \end{cases}$$

In the parameter space (see Fig. 3(b)), the  $MI$ -surface consists of a peak whose top is a global maximum (small maxima are avoided by the stochastic steepest gradient). At the first iteration ( $j = 1$ ), the  $k$ -th image (target) and the  $k+1$ -th image (source) are completely superimposed ( $T_{2D}^{k,k+1}$  is an identity matrix at  $j = 1$ ) and the  $MI$ -value is systematically on peak flank due to the small endoscope displacement between two acquisitions  $k$  and  $k+1$ . The initial  $\lambda_{r,s}^1$  converge step values have to be fixed so that the optimization algorithm remains into the peak and converges towards the top. In practice, the convergence is ensured when each  $a_{r,s}^1$  taken separately leads to displacement smaller than 2% of the source image diagonal. As illustrated in Fig. 3(a), the  $\lambda_{r,s}^j$  are kept constant

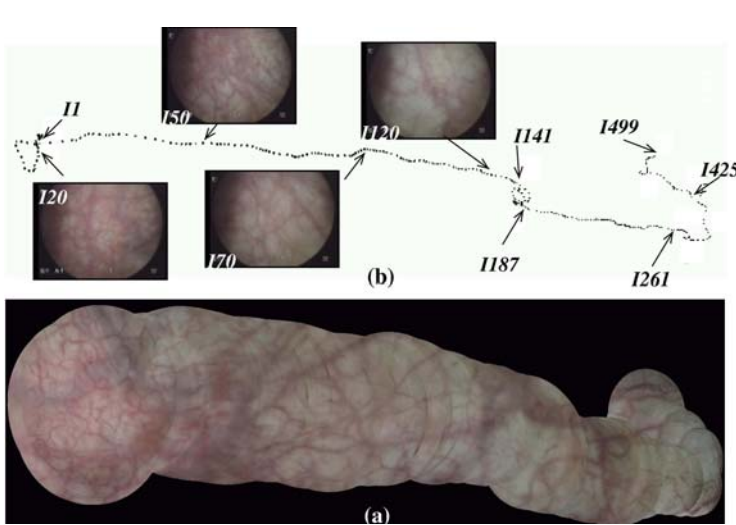


Fig. 4. 2D map example of a human bladder. (a) Mosaic built with 500 images. (b) Endoscope trajectory and positions of some images on this trajectory.

while the algorithm is progressing towards the peak top. When the algorithm oscillates around the top, the values of the  $\lambda_{rs}^j$  are updated by diminishing their current value by 15%. The peak top is supposed to be reached when the  $a_{rs}^j$  lead to displacements smaller than 1 pixel.

Finally, for a video-sequence, the 2D-map is constructed by placing all images in a common 2D coordinate system (i.e. that of the first image of the sequence) using all  $\tilde{T}_{2D}^{k,k+1}$ .  $\tilde{T}_{2D}^{1,k+1}$  is the global transformation representing the geometrical relationship between the coordinate system  $k+1$  and that of the first acquisition ( $k=1$ ).  $\tilde{T}_{2D}^{1,k+1}$  is calculated with (9).

$$\tilde{T}_{2D}^{1,k+1} = \tilde{T}_{2D}^{k,k+1} \times \tilde{T}_{2D}^{1,k} = \prod_{j=0}^{k-1} \tilde{T}_{2D}^{k-j,k-j+1} \quad (9)$$

Fig. 4 shows a result obtained for a human bladder sequence. The  $\tilde{T}_{2D}^{k,k+1}$  transformations were computed for 500 images and used in (9) to generate the 2D map of Fig. 4(a). This figure shows that it is possible to reconstruct visually coherent maps (without texture discontinuities) facilitating the diagnosis. Results quantifying the 2D-map construction accuracy and robustness are discussed in Section IV.

### III. TOWARDS 3D CARTOGRAPHY

In 2D large fields of view as represented in Fig. 4(a), whole regions of interest can be seen at one go. This type of visualization is clearly advantageous from the diagnosis point of view. However, this 2D representation suffers from limitations. Indeed, even if an organ has been *completely* scanned with the endoscope, it may happen that a region of interest is divided into two and located on two different map border parts. In brief, a single tissue of interest can be divided into several parts, located in different disconnected regions in the 2D map. To alleviate this connectivity problem,

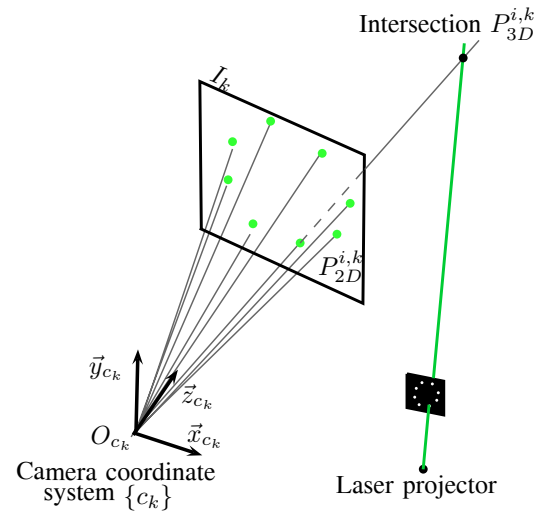


Fig. 5. 3D reconstruction of the  $P_{3D}^{i,k}$  points on the internal organ surface. For the  $k$ -th viewpoint and the  $i$ -th point, two 3D lines containing  $P_{3D}^{i,k}$  are being estimated. They issue from the camera optical centre and the diffractive lens, respectively.

and because clinicians mentally reconstruct organs when they observe 2D images, building 3D maps (3D large field of view surfaces with texture and color information) is of great interest. It can be expected that the 3D maps will not suffer from the above-mentioned connectivity issue when the whole organ is scanned.

In the following, we describe a modified endoscope relying on a combination of two light channels. For each viewpoint, the instrument allow for the reconstruction of a set of 3D points located on the internal organ surface (subsection III-A). The knowledge of a large number of 3D points (measured for all viewpoints) will allow us to sketch a 3D cartography algorithm (subsections III-B, III-C and III-D).

#### A. 3D endoscope principle

Typical endoscopes consist of two channels: the input channel illuminates the tissue with white light whereas the output channel allows for the acquisition of the reflected light with a camera. In [12], an additional channel was used to obtain 3D information with endoscopes. A green collimated laser source passes through a diffractive optics (holographic binary phase lens) built to generate a given light pattern (a dot matrix in [12]). This light pattern is projected through the third channel onto the surface to be reconstructed. Classical projector [12] and camera calibration [13] methods lead to the evaluation of the instrument parameters allowing for the computation of the 3D laser dot center positions. A detailed description of the 3D dot reconstruction method is given in [12]. Before going further, we summarize the principle of this reconstruction algorithm, since it is the starting point of our 3D cartography algorithm.

Let  $n$  be the number of green dots projected for each endoscope position. Consider a given viewpoint  $k$  and the cor-

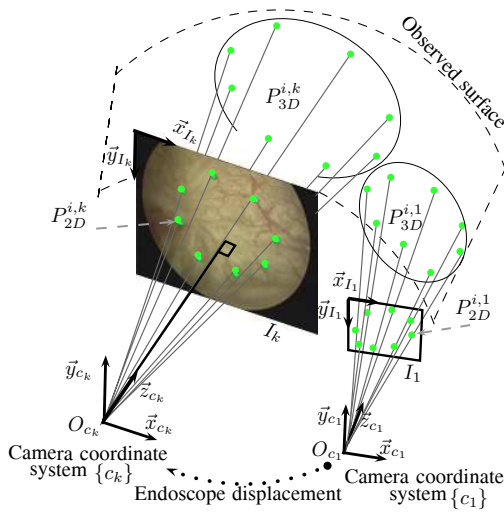


Fig. 6. Available data subset for each  $k$  position of the endoscope: one  $I_k$  distortion free grey level image, eight  $P_{2D}^{i,k}$  points, and eight  $P_{3D}^{i,k}$  reconstructed points.  $I_k$  is represented with the color of the acquired image to show that the high hue contrast between the green dots and the reddish organ tissue facilitates the segmentation of the  $P_{2D}^{i,k}$  points.

responding image  $I_k$ . In order to reconstruct the 3D position of the  $i$ -th green dot center ( $i = 1, \dots, n$ ), its 2D position must first be estimated in image  $I_k$  by an image segmentation procedure (see Fig. 5). The camera parameters are then used to determine the 3D line passing through the optical center of the camera and the  $i$ -th dot. Similarly, the 3D line issuing from the diffractive lens and passing through the dot center can be computed. The total number of 3D lines is equal to  $2n$  ( $n$  lines per channel) since there are  $n$  projected dots. The 3D position of the  $i$ -th dot (resulting from the dot projection onto the organ surface) is finally found by researching its corresponding pair of lines (association of two lines among  $2n$ ) and intersecting them.

In this contribution, an experimental set-up based on the principle of the endoscope described in [12] is used to project  $n = 8$  green dots onto the internal organ surface (see Fig. 6, and [14]). Due to the weak number of dots and their small size, each green dot projection is only spread over a few image pixels. Thus, the main texture and color information remain available in the acquired images. It is also noticeable that, due to the weak endoscope displacement in bladders (some  $mm/s$ ) in comparison with the acquisition speed (25 images per second), we can assume that the bladder (filled up with water) surface does not change between two consecutive image acquisitions.

As illustrated in Fig. 6, a data sequence consists of  $M$  images  $I_k$  (for  $k \in \{1, \dots, M\}$ ) and additional computed information (eight  $P_{2D}^{i,k}$  2D dot projection centers extracted from  $I_k$  and eight  $P_{3D}^{i,k}$  reconstructed dot centers, with  $i \in \{1, \dots, 8\}$ ). For the  $k$ -th viewpoint, the coordinates of the  $P_{3D}^{i,k}$  points are known in the local coordinate system  $(O_{c_k}, \vec{x}_{c_k}, \vec{y}_{c_k}, \vec{z}_{c_k})$ , where the  $O_{c_k}$  origin corresponds to the camera optical center.

To solve the 3D surface reconstruction problem, it is necessary to place all  $P_{3D}^{i,k}$  points in a common coordinate system,

for instance that of the first data set  $(O_{c_1}, \vec{x}_{c_1}, \vec{y}_{c_1}, \vec{z}_{c_1})^1$ . The 3D point coordinate transformation is guided by the successive registrations of consecutive  $I_k$  and  $I_{k+1}$  image pairs. Similar to the 2D mosaicing method, and prior to the registration, the images are being converted in grey levels and their distortions are corrected [14] (in the following,  $I_k$  stands for these distortion free grey level images). However, the final 3D map which will be superimposed on the reconstructed surface will include the color information.

### B. Geometrical considerations

The  $P_{2D}^{i,k}$  and  $P_{3D}^{i,k}$  points are mathematically related by the homogeneous perspective matrix  $\mathcal{K}$  fully defined by the  $f$ ,  $S_x$ ,  $S_y$ , and  $(u_0, v_0)$  intrinsic camera parameters. In (10),  $f$  (line in bold in Fig. 6) refers to the camera focal length,  $S_x$  and  $S_y$  correspond to the size of the pixels along the  $x$  and  $y$  axes, respectively, and  $(u_0, v_0)$  are the coordinates of the  $O_{c_k}$  optical centre projection onto the  $I_k$  image plane.

$$z_{3D}^{i,k} \begin{bmatrix} P_{2D}^{i,k} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{S_x} & 0 & u_0 & 0 \\ 0 & \frac{f}{S_y} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} = \mathcal{K} \begin{bmatrix} P_{3D}^{i,k} \\ 1 \end{bmatrix} \quad (10)$$

$$\text{with } P_{2D}^{i,k} = \begin{pmatrix} x_{2D}^{i,k} \\ y_{2D}^{i,k} \end{pmatrix} \text{ and } P_{3D}^{i,k} = \begin{pmatrix} x_{3D}^{i,k} \\ y_{3D}^{i,k} \\ z_{3D}^{i,k} \end{pmatrix}.$$

The  $T_{3D}^{k,k+1}$  geometrical transformation between  $(O_{c_k}, \vec{x}_{c_k}, \vec{y}_{c_k}, \vec{z}_{c_k})$  and  $(O_{c_{k+1}}, \vec{x}_{c_{k+1}}, \vec{y}_{c_{k+1}}, \vec{z}_{c_{k+1}})$ , corresponding to two consecutive endoscope viewpoints, is a rigid transformation involving three rotations ( $R_x$ ,  $R_y$ , and  $R_z$ ) and three translations ( $T_x$ ,  $T_y$ , and  $T_z$ ). As formulated in (11),  $T_{3D}^{k,k+1}$  can be used to compute the new coordinates of point  $P_{3D}^{i,k+1}$  (reconstructed by the endoscope for the  $k+1$ -th viewpoint) into the coordinate system of the  $k$ -th viewpoint. The point displaced from the  $k+1$  coordinate system into that of the  $k$ -th acquisition is denoted by  $\hat{P}_{3D}^{i,k}$ . It is worth noticing that  $\hat{P}_{3D}^{i,k}$  computed with (11) differs from the  $P_{3D}^{i,k}$  point provided by the endoscope for viewpoint  $k$  since the laser dot locations on the internal wall surfaces change between two endoscope positions  $k$  and  $k+1$  (no homologous 3D points are reconstructed with the endoscopic data).

$$\begin{bmatrix} \hat{P}_{3D}^{i,k} \\ 1 \end{bmatrix} = T_{3D}^{k,k+1} \begin{bmatrix} P_{3D}^{i,k+1} \\ 1 \end{bmatrix} \quad (11)$$

Since in (11),  $\hat{P}_{3D}^{i,k}$  and  $P_{3D}^{i,k+1}$  refer to the same physical point located on the surface (they are only expressed in two different coordinate systems) their 2D projections ( $\hat{P}_{2D}^{i,k}$  and  $P_{2D}^{i,k+1}$ , respectively) refer to homologous points. The geometrical transformation aligning  $P_{2D}^{i,k+1}$  on  $\hat{P}_{2D}^{i,k}$  is a 2D perspective transformation  $T_{2D}^{k,k+1}$ :

$$\beta^{i,k} \begin{bmatrix} \hat{P}_{2D}^{i,k} \\ 1 \end{bmatrix} = T_{2D}^{k,k+1} \begin{bmatrix} P_{2D}^{i,k+1} \\ 1 \end{bmatrix} \quad (12)$$

<sup>1</sup>This placement in a common coordinate system is referred by discrete surface construction.

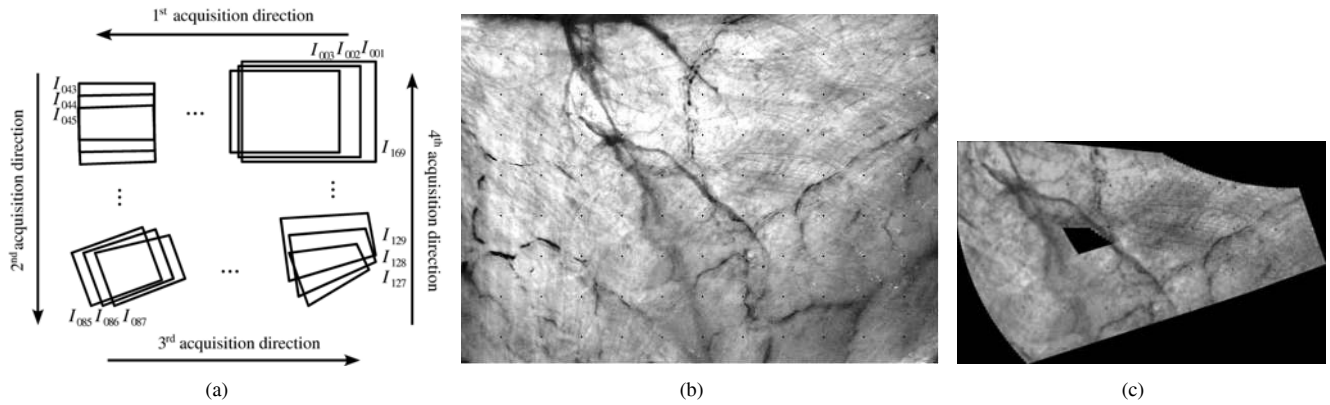


Fig. 7. 2D cartography test protocol. (a) Image acquisition path applied to the bladder photograph phantom for quantifying mosaicing precision (169 images :  $I_{001}, I_{002} \dots I_{169}$ ). (b) Regular grid of black dots overprinted on the pig bladder photograph phantom. (c) Panoramic image ( $1947 \times 1187$  pixels) constructed from the 169 acquired images.

$T_{2D}^{k,k+1}$  is mathematically defined by the  $3 \times 3$  homogeneous matrix of (2),  $\beta^{i,k}$  being a value due to the perspective change.

### C. Relationship between the $T_{2D}^{k,k+1}$ and $T_{3D}^{k,k+1}$ matrices

Applying (10) to the point  $\hat{P}_{3D}^{i,k}$  obtained with (11), we have:

$$\hat{z}_{3D}^{i,k} \begin{bmatrix} \hat{P}_{2D}^{i,k} \\ 1 \end{bmatrix} = \mathcal{K} T_{3D}^{k,k+1} \begin{bmatrix} \hat{P}_{3D}^{i,k} \\ 1 \end{bmatrix}. \quad (13)$$

From (12) and (13) we finally deduce (14). For at least four  $P_{3D}^{i,k+1}$  points (in practice  $i \in (i, \dots, N)$  and  $N = 8$ ), the unique solution  $T_{2D}^{k,k+1}$  of (14) is an over-determined system of  $2N$  equations which corresponds to a given  $T_{3D}^{k,k+1}$ .

$$\frac{1}{\hat{z}_{3D}^{i,k}} \mathcal{K} T_{3D}^{k,k+1} \begin{bmatrix} P_{3D}^{i,k+1} \\ 1 \end{bmatrix} = \frac{1}{\beta^{i,k}} T_{2D}^{k,k+1} \begin{bmatrix} P_{2D}^{i,k+1} \\ 1 \end{bmatrix}. \quad (14)$$

It is worth noticing that  $T_{2D}^{k,k+1}$  aligns all homologous pixels of  $I_{k+1}$  on  $I_k$  (not only the laser dot projections); two consecutive images  $I_{k+1}$  and  $I_k$  share large overlapping regions corresponding to homologous 2D data. This remark is the starting point of our 3D large field of view map construction algorithm (determination of the  $T_{3D}^{k,k+1}$  transformations), which enables all  $P_{3D}^{i,k}$  points to be placed in a common coordinate system. We now describe this positioning algorithm.

### D. Positioning of the 3D points in a common coordinate system (3D discrete surface construction)

The 3D rigid transformation between two cystoscope positions  $k$  and  $k+1$  is computed using an optimization method.

At the  $j$ -th iteration,  $T_{3D}^{k,k+1}$  is actualized by minimizing a squared error corresponding to the superposition of the source image  $T_{2D}^{k,k+1}(I_{k+1})$  and the  $k$ -th target image  $I_k$ :

$$\mathcal{E}(T_{3D}^{k,k+1}) = \left\| I_k^{ov} - T_{2D}^{k,k+1}(I_{k+1}^{ov}) \right\|^2. \quad (15)$$

As in Section II, the superscript *ov* refers to the overlapping parts of both images. The error  $\mathcal{E}$  is defined as a sum of quadratic grey level differences between both 2D images  $I_k$

and  $T_{2D}^{k,k+1}(I_{k+1})$ . It is a function of  $T_{3D}^{k,k+1}$ : when  $T_{3D}^{k,k+1}$  is set, the corresponding  $T_{2D}^{k,k+1}$  perspective transform can be deduced in the following way:

- 1)  $T_{3D}^{k,k+1}$  is used to position the  $P_{3D}^{i,k+1}$  points in the coordinate system of acquisition  $k$ , yielding  $\hat{P}_{3D}^{i,k}$ .
- 2) The homogeneous perspective matrix  $\mathcal{K}$  is then used to project the  $\hat{P}_{3D}^{i,k}$  points in the  $k$ -th image plane, yielding  $\hat{P}_{2D}^{i,k}$ .
- 3) Finally, the  $T_{2D}^{k,k+1}$  parameters are being estimated by solving the over-determined linear system (12).

Performing steps 1) to 3) is equivalent to iteratively solve the over-determined system (14).  $\mathcal{E}$  is optimized with respect to  $T_{3D}^{k,k+1}$  using the simplex algorithm and the preceding estimated transform  $T_{3D}^{k-1,k}$  as initial solution.

Once all  $T_{3D}^{k,k+1}$  have been estimated, it is straightforward to deduce the global transform  $T_{3D}^{1,k}$  from the  $k$ -th to the first camera coordinate systems:

$$T_{3D}^{1,k} = \prod_{j=1}^{k-1} T_{3D}^{j,j+1}. \quad (16)$$

For each viewpoint  $k$  and for each 3D point  $P_{3D}^{i,k}$ ,  $i \in \{1, \dots, n\}$ ,  $T_{3D}^{1,k}$  is used to place  $P_{3D}^{i,k}$  in the first camera coordinate system (considered as the global system). In practice, because the number of viewpoints is huge and there are  $n = 8$  3D points per viewpoint, positioning all 3D points results in a large number of points in the global coordinate system. The points are finely distributed on the organ surface.

## IV. RESULTS AND DISCUSSION

### A. 2D cartography

In order to quantify the mosaicing precision of the 2D cartography, a fine control of the cystoscope movements and the precise knowledge of its position coordinates relative to the bladder wall are needed (such information is not available in clinical *in vivo* conditions). Therefore, a dedicated acquisition protocol on a phantom with precise control of the cystoscope



Fig. 8. Panoramic image from a sequence of cystoscopic images obtained in clinics.

positions was undertaken. A 3D motorized micrometric positioning system allowed for applying in-plane translations and scale factor variations along  $\vec{x}$ ,  $\vec{y}$  and  $\vec{z}$  axis ( $0.1 \mu\text{m}$  precision), as well as in- and out-of-plane rotations ( $0.01^\circ$  precision). The phantom is a printed photograph ( $10 \times 12 \text{ cm}$ ) of the inner surface of a widely opened pig bladder, after excision. The pig bladder was chosen for its similarities to the human one in terms of texture and anatomy. The image acquisition sequence was performed following the predefined path sketched in figure 7(a) and implying 2 rotations and 3 translations of the cystoscope. The 1<sup>st</sup> acquisition direction included a combined translation of 42 mm in  $\vec{x}$  and 42 mm along  $\vec{z}$  (scale-factor variation). The 2<sup>nd</sup> direction involved a movement of 42 mm translation along  $\vec{y}$  coupled to an in-plane rotation from 0 to  $20^\circ$ . The 3<sup>d</sup> direction movement is composed of a -42 mm translation along  $\vec{x}$  with an off-plane rotation from 0 to  $16^\circ$  (perspective modification). The last acquisition path involved translations along  $\vec{y}$  and  $\vec{z}$ , and in- and off-plane rotations simultaneously, returning the device to its starting position. 42 images were acquired along each direction with the first ( $I_{001}$ ) and last ( $I_{169}$ ) images taken at the same cystoscope position.

A grid of regularly spaced black dots was printed over the photograph of the bladder (see figure 7(b)) and exploited for calculating an error  $e_m$  representing a measure of both the registration accuracy and the panoramic image construction accuracy (mosaicing accuracy). The panoramic image obtained after applying our mosaicing algorithm to the acquired image sequence is shown in figure 7(c). Automatic computation of the registration/mosaicing error was carried out after rigid registration and segmentation of every visible dots (31 dots) in the panoramic image. The error measure  $e_m$  was calculated as the mean Euclidean distance between the coordinates  $(x_m^j, y_m^j)$  of  $N_p$  dot centroids segmented in the panoramic image and the coordinates  $(x_r^j, y_r^j)$  of the homologous dot

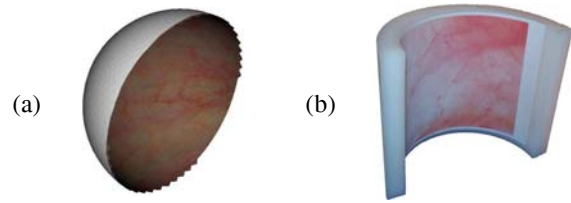


Fig. 9. Two phantoms with different realistic bladder textures. (a) Simulated sphere. (b) Real half cylinder.

centroids segmented in the photograph, so that:

$$e_m = \frac{1}{N_p} \sum_{j=1}^{N_p} \sqrt{(x_m^j - x_r^j)^2 + (y_m^j - y_r^j)^2}. \quad (17)$$

The mean  $e_m$  registration error between two consecutive images  $I_{k+1}$  and  $I_k$  amounts to 1.047 pixels. For such errors, two superimposed images present no texture discontinuities so that the registered images are always visually coherent. As illustrated by the patient data results of Figs. 4(a) and 8, the 2D maps are without texture discontinuities as long as the cystoscope trajectory do not include closed loops. For closed loops, due to registration error accumulation over the image sequence, using only the mutual information based registration algorithm leads to texture discontinuities between the first and last images of the loop. For the phantom sequence of Fig. 7, the  $e_m$  mean mosaicing error between images  $I_{001}$  and  $I_{169}$  amounts to 43.14 pixels. To diminish strongly this error, we proposed an algorithm [10] that superimposes the first and last loop images while keep the visual coherence of the images located into the loop. The mutual information based method, associated to this accumulation error correction algorithm, was used to build the map of Fig. 7(c). In this panoramic image the  $e_m$  mean value is now 3.35 pixels. As seen in Fig. 7(c), visual coherence is ensured for the loop. Considering that the size of the panoramic image is  $1947 \times 1187$  pixels, the maximum error represents less than one percent of the image height.

Our mosaicing algorithm was applied to clinical cystoscopic sequences for demonstrating its applicability to real conditions (*in vivo*) and for performing a qualitative validation on clinical data [15]. These panoramic images were analyzed by urologists who validated their visual coherence. Although the mosaicing errors were not measured here, the clinicians estimated they were small enough to allow them to not distort the quality of image interpretation and analysis.

### B. 3D surface construction

Since the  $T_{3D}^{k,k+1}$  and  $T_{2D}^{k,k+1}$  transformations are unknown for real bladder data, two phantoms were used to assess the 3D textured surface construction algorithm. The first phantom, consisting of a simulated surface superimposed by real human bladder texture, was used to obtain quite realistic data and to test the inherent accuracy and the robustness of the 3D surface reconstruction. The second phantom, consisting of a real surface with bladder texture, allows for the acquisition of data with the active vision system described in section III-A. Both



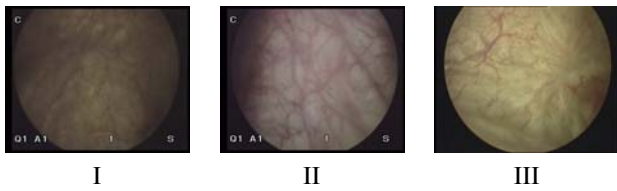


Fig. 10. I, II and III : three images extracted from real endoscopic exams during the robustness evaluation tests. The chosen images present strong differences both in texture and illumination.

phantoms exhibit known parametric 3D shapes allowing for a quantitative assessment of the surface construction algorithm.

### B.1 Algorithm robustness and inherent accuracy

The simulated surface is a 5 cm radius sphere (see Fig. 9(a)).

For the robustness assessment, three images (see Fig. 10) exhibiting various textures and illumination conditions were extracted from human endoscopic sequences and mapped onto the inner surface of the simulated half sphere. Image acquisitions were simulated using realistic camera parameters ( $\mathcal{K}$  matrix and distortion coefficients values) and by applying known 3D displacements to the virtual 3D endoscope. The simulated images and the  $P_{3D}^{i,k}$  laser points were used by the surface construction algorithm to find the 3D transformations. The transformations calculated with our surface construction algorithm are ideally equal to the known transformations used to simulate the data (2D image sequence and 3D laser dots). These experiments allow for the assessment of the largest endoscope viewpoint change leading to successful 3D point displacements from the  $k$ -th + 1 into the  $k$ -th coordinate system. Table I provides the parameter value intervals. The algorithm robustness is acceptable since the interval limits of table I correspond to transformations being by far larger than those of real endoscope displacements. Indeed, the 25 images per second acquisition rate and the small endoscope displacement speed (few  $mm/s$ ) lead to translations and rotations typically smaller than 0.3 mm and  $1^\circ$ , respectively.

The inherent reconstruction accuracy was tested by assessing the ability of the algorithm to precisely reconstruct surface shapes. A large field of view pig bladder photography was projected onto the inner surface of the half sphere (for urologists, pig bladder textures are visually very close to those of human bladders). As for the previous robustness evaluation experiment, image acquisitions were performed using a virtual endoscope. 65 instrument displacements were

3D transformation parameters	Transformation value intervals	
	2D	3D
Translation ( $T_x$ and $T_y$ )	$\pm 76$ pixels	$\pm 5$ mm
Scale factor ( $T_z$ )	$\pm 30\%$	$\pm 12.5$ mm
In plane rotation ( $R_z$ )	$\pm 20^\circ$	$\pm 20^\circ$
Out of plane rotations ( $R_x$ and $R_y$ )	$\pm 7^\circ$	$\pm 7^\circ$

TABLE I

PARAMETER VALUE INTERVALS OBTAINED FOR THE INNER HALF SPHERE SURFACE MAPPED WITH THE IMAGES OF FIGURE 10

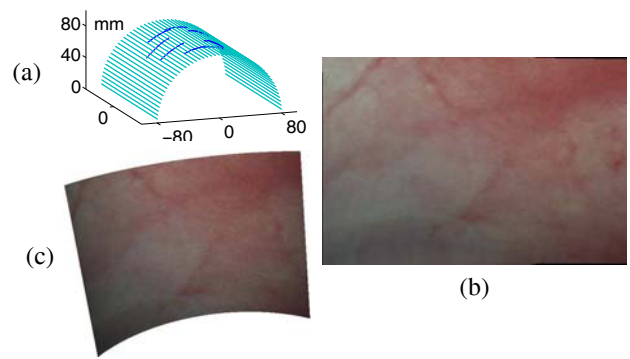


Fig. 11. Results obtained with real (acquired) data. (a) Reconstructed 3D points and interpolated cylinder. (b) 2D panoramic image to be stuck in the cylinder. (c) 3D map constructed for the sequence of 40 acquisitions: the textured image (b) is stuck on the cylinder using the known  $P_{2D}^{i,k}/P_{3D}^{i,k}$  point correspondence.

simulated by setting  $\tilde{T}_{3D}^{k,k+1}$  and  $\tilde{T}_{2D}^{k,k+1}$  transformations, the  $\sim$  symbol denoting the reference (simulated) transformations. Each  $\tilde{T}_{3D}^{k,k+1}$  transformation is a combination of the following elementary translations and rotations: 0.66 mm for  $T_x$ ,  $T_y$ , and  $T_z$  translations and  $2^\circ$  for  $R_x$ ,  $R_y$ , and  $R_z$  rotations. These parameters actually lead to displacements that are larger than those typically observed with endoscopes (simulation of extreme acquisition conditions).

After applying the 3D surface construction method to the simulated sequence, we computed the Euclidean distances between all 3D points of the discrete surface and the sphere centre. The computed mean radius and standard deviation amount to  $5.02 \text{ cm} \pm 0.26 \text{ cm}$  (ideally  $5 \text{ cm} \pm 0 \text{ cm}$ ). Such errors are by far sufficient to apply 3D cartography to hollow organ lesion diagnosis (it makes medically no sense to reconstruct bladder surfaces with a high precision).

### B.2 Accuracy assessment using real (acquired) data

Another pig bladder photography was used to map the inner surface of the half cylinder (see Fig. 9(b)). Forty acquisitions (each acquisition yields an image and eight  $P_{3D}^{i,k}$  points) were performed using an experimental set-up similar to that of [12] (a white light source endoscope, a CCD camera, and a miniaturized diffractive optics generating eight laser beams). This phantom is quite realistic (real bladder textures) and the illumination conditions are very close to those of real bladder scenes viewed by 3D endoscopes.

As sketched in Fig. 11(a), we computed the cylinder passing at best through the constructed point cloud, all  $P_{3D}^{i,k}$  points being placed in the coordinate system of the first acquisition ( $k = 1$ ). The discrete surface construction error is defined as the difference between the exact cylinder radius (8 cm) and the average distance between the main cylinder axis and the 3D points all placed in the first coordinate system.

This error amounts to 0.8 mm. Such a 3D representation of hollow organs with sub-millimetre accuracy is by far sufficient for clinicians. The panoramic image obtained as output is shown in Fig. 11(b). It does not include texture discontinuities

and the visual coherence is quite good. Also, the laser spots which are included in the data images have been removed from the panoramic image since the transformed images are overlapped and pixels containing laser points can be replaced by homologous (grey level) pixels of other images. The 3D map is finally generated by sticking the panoramic image on the reconstructed 3D surface (see Fig. 11(c)).

## V. CONCLUSION AND FURTHER WORK

Even if scanners provide accurate 3D anatomical and/or functional data, they cannot be used for lesions diagnosis in hollow organs or during mini-invasive surgery. For such tasks, endoscopy is the appropriate image modality. The drawback of cystoscopes or laparoscopes lies in the small field of view of the acquired 2D images. This paper is among the few contributions of the literature dealing with wide field of view construction of human organs using endoscopic data.

The proposed 2D bladder cartography algorithm is robust and accurate, and leads to visually coherent maps facilitating lesion diagnosis in hollow organs (as confirmed by the urologists the authors are working with). The aim of further work is to diminish the computation time of the proposed algorithm (one minute is required to register two images). One solution to reach this goal is to predict the endoscope trajectory in order to select the images to be registered. Indeed, currently all images of the sequence are registered even if two consecutive images have large overlapping parts (typically 90% of the image surface). The endoscope displacement prediction can be used to select only the images presenting a minimal image overlapping that ensures an accurate registration.

Preliminary results were also given for a 3D cartography algorithm of inner bladder walls. The cartography algorithm consists of a first step placing in a common coordinate system all 3D points acquired from different viewpoints by a 3D active vision based instrument. Tests performed with phantoms exhibiting simple and analytically known shapes demonstrate the potential of the surface construction algorithm. The aim of a further work is compute, using the discrete surfaces, continuous bladder surfaces with textures. Bi-cubic splines can for instance be used in a piecewise fitting scheme to interpolate accurately the 3D discrete points. The 2D endoscopic images can then be projected on the interpolated points to obtain a continuous surface with natural color and texture information.

## ACKNOWLEDGEMENT

This work is partially sponsored by the CNRS (PEPS INSIS 2010). R. Miranda-Luna was sponsored by the CONACYT. The authors address their grateful thanks to urologist M.-A. D'Hallewin from Cancer Institute CAV in Nancy (France) for her clinical experience and for providing video sequences of various cystoscopic examination. The authors also thank the surgeons from the Experimental Surgery Laboratory (Faculty of Medicine, Nancy) for providing fresh pig bladders.

## REFERENCES

- [1] A. Jalink, J. McAdoo, G. Halama, and Hong Liu. CCD mosaic technique for large-field digital mammography. *IEEE Trans. on Med. Imag.*, 15(3):260–267, June 1996.
- [2] G. Yang and C. Stewart. Covariance-driven mosaic formation from sparsely-overlapping image sets with application to retinal image mosaicing. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)*, volume 1, pages 804–810, June 2004.
- [3] T. Vercauteren, A. Perchant, X. Pennec, and N. Ayache. Mosaicing of confocal microscopic in vivo soft tissue video sequences. In *Lecture Notes in Computer Science, Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI'05)*, pages 804–810, Palm Springs, CA, Oct. 2005.
- [4] J. S. Chou, J. Qian, Z. Wu, and H. F. Schramm. Automatic mosaic and display from a sequence of peripheral angiographic images. In K. M. Hanson, editor, *SPIE Medical Imaging: Image Processing*, volume 3034, pages 1077–1087, Apr. 1997.
- [5] A. Behrens. Creating panoramic images for bladder fluorescence endoscopy. *Acta Polytechnica Journal of Advanced Engineering*, 48(3):50–54, June 2008.
- [6] A. Behrens, T. Stehle, S. Gross, and T. Aach. Local and global panoramic imaging for fluorescence bladder endoscopy. In *Engineering in Medicine and Biology Society (EMBC), 31th Annu. Int. Conf. of the IEEE*, pages 6990–6993, Minneapolis, Sep. 2009.
- [7] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imaging*, 22(8):986–1004, Aug. 2003.
- [8] R. Miranda-Luna, W. C. P. M. Blondel, C. Daul, Y. Hernández-Mier, R. Posada, and D. Wolf. A simplified method of endoscopic image distortion correction based on grey level registration. In *Proc. IEEE ICIP*, volume 5, pages 3383–3386, Singapore, Oct. 2004.
- [9] R. Miranda-Luna, Y. Hernandez-Mier, C. Daul, W. P. C. M. Blondel, and D. Wolf. Mosaicing of medical video-endoscopic images: data quality improvement and algorithm testing. In *Proc. Int. Conf. on Electrical and Electronics Engineering*, Acapulco Guerrero, Mexico, Sep. 2004.
- [10] R. Miranda-Luna, C. Daul, W. C. P. M. Blondel, Y. Hernandez-Mier, D. Wolf, and F. Guillemin. Mosaicing of bladder endoscopic image sequences: Distortion calibration and registration algorithm. *IEEE Trans. Biomed. Eng.*, 55(2):541–553, Feb. 2008.
- [11] W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.
- [12] C. Manhong, L. Wumei, Z. Changhe, and Y. Jianan. Miniaturized three-dimensional endoscopic imaging system based on active stereovision. *Applied Optics*, 42(10):1888–1898, 2003.
- [13] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, November 2000.
- [14] A. Ben Hamadou, C. Soussen, C. Daul, W. Blondel, and D. Wolf. Flexible projector calibration for active stereoscopic systems. In *Proc. IEEE ICIP*, Hong Kong, September 26–29 2010.
- [15] Y. Hernandez-Mier, W. C. P. M. Blondel, C. Daul, D. Wolf, and F. Guillemin. Fast construction of panoramic images for cystoscopic exploration. *Computerized Medical Imaging and Graphics*, 2010.