



HAL
open science

Fast construction of panoramic images for cystoscopic exploration

Yahir Hernandez-Mier, Walter Blondel, Christian Daul, Didier Wolf, François Guillemin

► **To cite this version:**

Yahir Hernandez-Mier, Walter Blondel, Christian Daul, Didier Wolf, François Guillemin. Fast construction of panoramic images for cystoscopic exploration. *Computerized Medical Imaging and Graphics*, 2010, 34 (7), pp.579-592. 10.1016/j.compmedimag.2010.02.002 . hal-00441663

HAL Id: hal-00441663

<https://hal.science/hal-00441663>

Submitted on 13 Jun 2023

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.

Fast construction of panoramic images for cystoscopic exploration

Y. Hernández-Mier^{a,b,*}, W.C.P.M. Blondel^a, C. Daul^a, D. Wolf^a, François Guillemin^{a,c}

^a Centre de Recherche en Automatique de Nancy (CRAN), UMR 7039 Nancy University - CNRS, 2 Avenue de la Forêt de Haye, F-54516 Vandœuvre-Lès-Nancy, France

^b Polytechnic University of Victoria, Carretera Victoria-Soto la Marina Km. 5.5. Parque Científico y Tecnológico TECNOTAM, 87137 Cd. Victoria, Tamaulipas, Mexico

^c Anticancer centre "Centre Alexis Vautrin" (CAV), Avenue de Bourgogne, 54511 Vandœuvre-Lès-Nancy, France

Cystoscopy is used as a reference clinical examination in the detection and visualization of pathological bladder lesions. Evolution observation and analysis of these lesions is easier when panoramic images from internal bladder walls are used instead of video sequences. This work describes a fast and automatic mosaicing algorithm applied to cystoscopic video sequences, where perspective geometric transformations link successive image pairs. This mosaicing algorithm begins with a fast initialization of translation parameters computed by a cross-correlation of images, followed by an iterative optimization of transformation parameters. Finally, registered images are projected onto a global common coordinate system. A quantifying test protocol applied over a phantom yielded a mosaicing mean error lower than 4 pixels for a 1947×1187 pixels panoramic image. Qualitative evaluation of 10 panoramic images resulting from videos of clinical cystoscopies was performed. An analysis performed over translation values from these clinical sequences (*in vivo*) is used to modify the mosaicing algorithm to be able to do a dynamic selection of image pairs. Construction time of panoramic images takes some minutes. At last, algorithm limits are discussed.

1. Introduction

1.1. Clinical context

Among hollow organs cancers, bladder cancer is the 7th most extended in the world [1]. 95% of bladder tumors originate in the epithelium or urothelium (surface tumors) [2]. One of the main problems concerning bladder tumors is their dangerous recurrence potential. Recurrence, with an important progression risk is often associated to Carcinoma *In Situ* (CIS). CIS in the bladder is a cancerous lesion that can be multifocal and diffuse. Due to its nature (flat, non-papillary tumor, located at the mucous membrane), CIS may not be detected by conventional cystoscopy; therefore, recurrence risk is increased. Some recent research results advice lifetime monitoring of patients after partial bladder surgical removal (cystectomy) to avoid recurrence [3]. Cystoscopy is the reference examination allowing the physician to visually detect and locate pathological lesions, specifically cancer, in the bladder. A cystoscope is a narrow instrument (soft or rigid) that is inserted through the urethra to reach the bladder and visualize internal bladder walls. Due to the reduced size of the instrument,

observation of every region of interest is made through a narrow visualization window. The physician moves this reduced field of view (maximal area of about 1 cm^2) over nearly all the tissue surface and the corresponding images are then recorded on a video system. The study presented in this paper proposes using information delivered by cystoscopes to construct extended views of tissue regions to follow the evolution of diseases and to simplify organ observations. Currently, the only physical media used by clinicians are photographs of a few particular sites, a bladder sketch with annotations and a video of the examination. With a photograph, only a limited part of a region of interest may be visible with enough details and resolution. With a video sequence, extended tissue areas can be analyzed along the cystoscope movements, but the obtained image sequence consists of a large amount of redundant data. In addition, as a complementary step to the anatomo-histopathological examination of tissue samples, the physician must compare *in vivo* the same bladder regions from one examination to another. In this case, finding a particular region of interest at a particular moment in the video sequence could take long and could be annoying for the physician. Urologists expressed the need for a fast-access visual medium to effectively locate potential bladder lesions, to easily observe neighboring regions around suspicious lesions or zones, and to efficiently compare zones of interest through their evolution. Finally, this kind of complementary visual medium could also be applied to recently developed fluorescence endoscopic imaging systems, giving a spatial correspondence between fluo-

* Corresponding author at: Polytechnic University of Victoria, Carretera Victoria-Soto la Marina Km. 5.5. Parque Científico y Tecnológico TECNOTAM, 87137 Cd. Victoria, Tamaulipas, Mexico. Tel.: +52 834 172 03 86.

E-mail address: yhernandezm@upv.edu.mx (Y. Hernández-Mier).

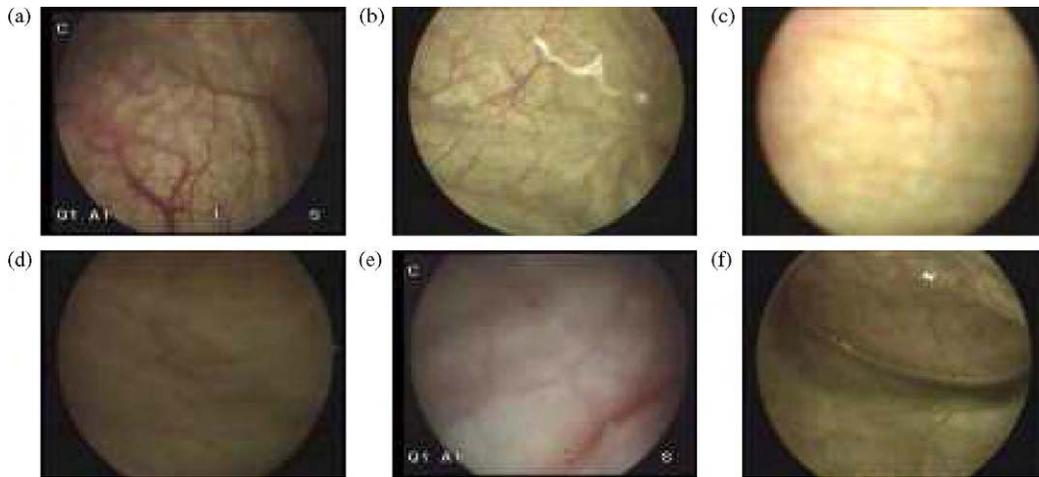


Fig. 1. Examples of images that can be found in a conventional cystoscopic video sequence. (a) Image with many visible blood vessels. (b) Elements in suspension (clearer) crossing the field of view of the cystoscope. (c) Blurred image due to defocusing of the cystoscope. (d) Blurred image due to rapid movements of the cystoscope in which few structures are visible due to poor lighting conditions. (e) Image taken at a short distance. (f) Image presenting light reflections at the air bubble.

rescence and white light images, resulting in higher diagnostic efficiency [4].

1.2. Aim and applications of image mosaicing

Image mosaicing is a process by which a panoramic image is built by superimposing partial views (images) of a scene or object. This panoramic image is a concise representation of data leading to a considerably reduced data storage space. Mosaicing is mainly based on an image registration process for determining geometric transformations $T_{i,i+1}$ between image pairs I_i and I_{i+1} , successively taken from different views. These local transformations are then used to compute a global transformation T_{i+1} , that is used to place the last transformed image I_{i+1} (as well as the preceding ones) in a global common coordinate system (for instance, the coordinate system of the first image).

Image mosaicing is used today in many different areas such as consumer photography [5–7], stabilization of hand-held recorded video (operator movement compensation) [8,9], construction of virtual environments [10–12], submarine exploration [13,14] and construction of images with super-resolution [15–17].

In the medical field, research has been conducted for improving the field of view in ultrasound imaging by registration of images obtained from manual scans [18]. The authors performed image registration using Mutual Information (MI) as similarity measure and an perspective transformation of 8 degrees of freedom. Although images used in that work were very noisy, no image pre-processing was used. In conventional X-ray radiology, Yaniv and Joskowicz [19] have constructed panoramic images for observing long bones that are not visible on a single radiograph. To register two or more radiographic images, they considered that no rotation exists between images and that transformations between images can be represented by movements that are parallel to the projection axis. A ruler was placed next to each image and alignment was manually performed using ruler marks. However, the main applications of image mosaicing in the medical field concern ophthalmology, where constructed panoramic images allow for easier and more efficient diagnosis. The approaches developed in ophthalmology, generally use similarity measures based on the comparison of pixel gray-levels in similar image regions and transformation models up to 6 degrees (affine model) for image registration. For example, Zhou et al. [20,21] registered successive angiogram images of the back of the eye using an inverse compositional-based registration algorithm [22–24] which was modified by introducing a tunable

increasing parameter. Some other approaches, such as that from Can et al. [25], use segmentation of visible blood vessels of the retina to extract intersection points that can be used to register images. Results of this method are highly dependent on the segmentation algorithm ability to find blood vessel intersections with enough robustness. This approach uses a non-linear transformation model of 12 degrees of freedom to represent the deformation between images and the retinal curvature. Up to now, the only referenced application of image mosaicing to hollow organs, and specifically to bladder, is the work of Miranda-Luna et al. [26,27,4]. In their work, extended views of internal bladder walls were constructed by mosaicing images extracted from a video sequence, corresponding to successive views taken by a cystoscope. Their approach uses Mutual Information (MI) as similarity measure to register images through an iterative optimization algorithm (stochastic steepest descent). Although panoramic images so constructed are visually coherent and the used mosaicing algorithm is robust, computation time of a panoramic image is very long (several hours). This is due to the analytical modeling of histograms or probability density functions (Parzen's window) used to compute MI and to the number of iterations needed in the optimization.

1.3. Objectives and application constraints

The objective of the present study is to develop a fast and automatic mosaicing algorithm to construct panoramic images of the internal wall of bladders, useful for the physician *in situ*. This algorithm must be able to provide panoramic images without affecting the application protocol of the cystoscopic examination. As a result of our previous study [26], urologists have requested that computation time of these panoramic images be reduced, so that they can evaluate and verify their coherence, their precision and their exploitability before the end of a clinical examination. Finally, the proposed mosaicing algorithm must be robust enough to deal with image variability found in cystoscopic video sequences (visible or not visible blood vessels, lighting conditions, blur, texture, and others). Fig. 1 shows examples of images acquired with a cystoscope under these variable conditions.

2. Mosaicing algorithm

The mosaicing algorithm presented here is composed of three main parts: image pre-processing (Section 2.1), registration algorithm and image stitching (Sections 2.2 and 2.3), and dynamic

selection of images to be registered (Section 2.4). The selection stage is based on an analysis of translation information delivered by an initial image correlation step. This stage avoids entering into the iterative part of the registration algorithm for image series where small movements or “static” sequences exist.

2.1. Image pre-processing

In the case of endoscopic images, it is necessary to correct or compensate geometric and photometric aberrations to achieve efficient registration. Three of the most relevant aberrations to the study presented here are: a vignetting effect, due to the wide field of view (FOV) of the endoscope (wide angle lens); a fiber optics pattern appearing over images (when a fiberscope is used) and radial distortion caused by the cystoscope optics (wide FOV). Correction of these effects is essential for a good performance of the subsequent registration stage. The analysis of the effects of these aberrations and their correction are developed in Section 3. The following subsections explain the strategies implemented to solve these problems.

2.1.1. Vignetting and fiber pattern

The vignetting effect corresponds to a lighting inhomogeneity in the image resulting in intensity variations of very low spatial frequency. It can be simply modeled as the application to an image without lighting artifacts I_o of a multiplicative term I_m and an additive term I_a , so that:

$$I_v = I_o \times I_m + I_a \quad (1)$$

where I_v is the image affected by vignetting.

In practice, it is difficult to determine the contribution of each term; therefore, approximations of the lighting gradient are built for correction. Considering, for example, that the image background is essentially composed of the illumination terms, it can be modeled by quadratic or polynomial functions [28]. This approach is suitable when the lighting conditions between images of a sequence do not change, which is not the case of cystoscopic sequences. Another approach implemented in [29–32] applies a high-pass filter to eliminate the corresponding very low frequencies. The characteristics of the filter are determined with the help of the analysis of Fourier spectra of typical images. This approach has been implemented in the present study because it offers a good compromise between efficacy and computation time (using adapted implementation techniques). A Gaussian mask $G_\sigma(x, y)$ was used:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

with σ being the standard deviation of the mask and (x, y) the pixels coordinates.

From the Fourier spectrum of cystoscopic images (Fig. 2(a) and (b)), we experimentally determined a standard deviation value for the Gaussian mask in the frequency domain $\sigma_{ill} = 0.012N$, with $N = 256$ the width of the square region of interest (ROI) of the image. This value corresponds to the interval of spatial frequencies $[0, 0.012N]$, where the very low frequency components related to lighting inhomogeneities are found. Useful information, bladder texture, is found in higher frequency components of the Fourier spectrum. After image background is obtained (Fig. 2(c)), it is subtracted from the initial image, leading to a vignetting-corrected image (Fig. 2(d)).

When a flexible cystoscope (fiberscope) is used, a spatially periodical pattern due to fiber optics bundle is visible on the images (see Fig. 2(a)) and can affect the registration robustness. As indicated by arrows in Fig. 2(e), the presence of this pattern produces high frequency peaks in the Fourier spectrum of the image. In literature,

there are few studies addressing this problem. The method presented in [27] localizes these high-frequency peaks in the Fourier spectrum to define the cut-off frequency of a Gaussian filter that is next applied to every image of the sequence acquired with the same fiberscope. Recently, Winter et al. [33] proposed an automatic method for the detection of the position of these peaks in the Fourier spectrum image followed by the computation of the cut-off frequency of the low-pass filters to eliminate the fiber pattern. The main advantage of their method is its capacity to automatically build adaptive masks. In the application developed in the present study, for the same endoscope, the characteristics of the filter can be computed once, and they will be the same for all images.

Spatial frequencies of the fiber pattern are located in the frequency bandwidth $[0.156N, 0.312N]$ in the fiberscope used in our tests (see Fig. 2(e)). The result of the application of a Gaussian filter with standard deviation in the Fourier domain $\sigma_{patt} = \frac{0.156}{3} = 0.052N$, is shown in Fig. 2(f).

Two low-pass filters were combined to form a band-pass filter that eliminates very-low frequencies due to vignetting and high frequencies due to the fiber pattern. As a result, useful information in the frequency bandwidth $[0.036N, 0.156N]$ is preserved. In spatial domain, sizes of the first and second Gaussian masks are respectively 83×83 pixels and 19×19 pixels. This filtering operation can be efficiently implemented in the frequency domain by masking unwanted frequency regions in the Fourier spectra of images.

2.1.2. Radial distortion

In [30] and [31], an algorithm for the correction of the radial distortion produced by cystoscope optics was proposed. Their approach obtains correction parameters from the cystoscope that will be used to perform the clinical exam. However, access to a cystoscope for calibration is not always simple in day-to-day practice, and the correction operation on every image implies a computation time that cannot be neglected. It was observed in all the analyzed cystoscopic images that radial distortion values at the central part of images are small enough if a reduced region of interest (ROI) is taken ($N^2 = 400 \times 400$ pixels typically). Consequently, radial distortion correction may be avoided on this reduced-size image which can be directly used in our mosaicing algorithm.

2.2. Image registration

2.2.1. Geometric transformation

The generic process of image registration of successive images I_i and I_{i+1} from a cystoscopic video sequence aims to superpose common parts of these images. Mathematically described by Eq. (3), it maximizes a similarity measure between extracted characteristics from target image $I_t = I_i$ and homologue characteristics from source image $I_s = I_{i+1}$ to which geometric transformation $T_{i,i+1}$ is applied (\circ). f_t and f_s are specific feature extraction algorithms which may be applied to these images. Parameter values of transformation $T_{i,i+1}$ are selected in a search space \mathbf{T} by means of an optimization algorithm opt to finally obtain transformation $\tilde{T}_{i,i+1}$ when $f_t(I_i) \approx f_s(I_{i+1} \circ T_{i,i+1})$ (registered images)

$$\tilde{T}_{i,i+1} = \arg \underset{T_{i,i+1} \in \mathbf{T}}{\text{opt}} S\{f_t(I_i), f_s(I_{i+1} \circ T_{i,i+1})\} \quad (3)$$

Transformation type T , similarity measure S , feature extraction algorithms f_t and f_s and optimization method opt must be chosen by taking into account some constraints. These constraints include sensor type, image acquisition modality, available computation time and registration precision required by the application. The type of transformation type T is chosen considering on the one hand, the projection model of the cystoscopic acquisition system, and on the other, the deformable nature of the bladder wall. Consid-

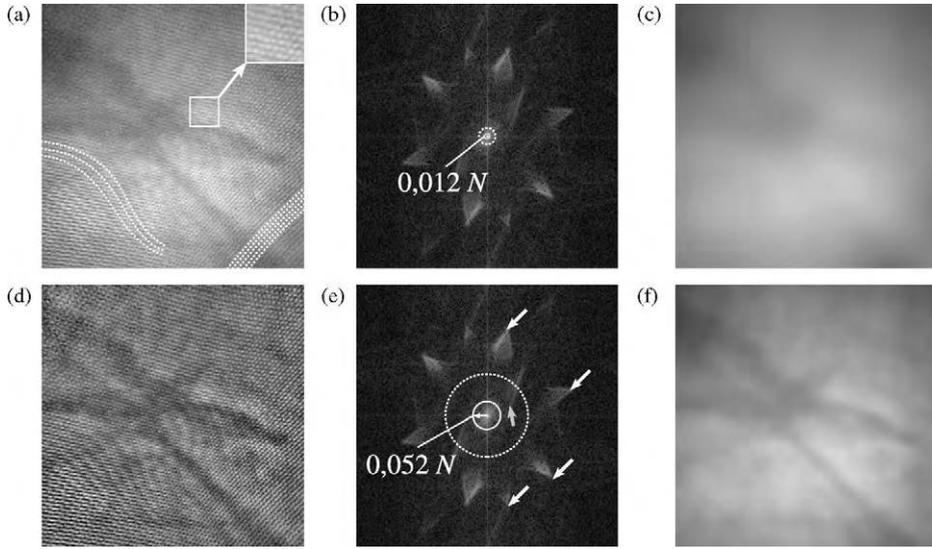


Fig. 2. (a) Original fiberoptic image with enlarged fiber pattern and line shapes patterns. (b) Fourier spectrum of image (a). The continuous circle at the center part of the image indicates the location of spatial frequencies values of $0.012N$ and the dashed circle shows the extension of the Gaussian mask ($3\sigma_M$). (c) Image resulting from low-pass filtering with $\sigma_M = 0.012N$. (d) Normalized image after pixel-wise subtraction of image (c) to image (a). (e) Fourier spectrum of image (a). The white arrows show the frequency components corresponding to the fiber pattern and the gray arrow shows the region of the spectrum corresponding to line patterns. The continuous circle corresponds to the value of $\sigma_{patt} = 0.052N$ and the dashed line shows the extension of the Gaussian mask. (f) Image resulting from low-pass filtering with $0.052N$ on image (d).

ering that radial distortion influence is small when a reduced-size window at the central part of acquired images (as mentioned in previous subsection) is selected, the 3D/2D projection model of the cystoscope is linear. Concerning non-linearities induced by the bladder wall deformation due to movements of surrounding organs, a non-linear transformation model could be considered for describing the geometrical relationship between adjacent pairs of images. However, once the bladder is filled up with physiological serum for the need of the endoscopic examination, the bladder wall can be considered as rigid enough to appear motionless with reference to the video acquisition rate (25 s). In addition to the apparent rigidity of the bladder wall, a relatively low displacement speed of the endoscope (a few mm/s typically) makes it possible to use a linear transformation model. The most complete 2D/2D linear transformation model we can use is a perspective model of 8 degrees of freedom, which can be represented as a 3×3 matrix M defined in block form and homogeneous coordinates as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} w = \underbrace{\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ v_1 & v_2 & v \end{bmatrix}}_M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

where $x = [x, y]^T$ and $x' = [x', y']^T$ are the coordinates of the corresponding homologous points in images I_i and I_{i+1} respectively; $a_{11}, a_{12}, a_{21}, a_{22}$ are dimensionless parameters forming a 2×2 non-singular matrix representing an affine transformation that includes scale and shearing factors, and in-plane rotation; $t = [t_x, t_y]$ is a vector that represents translation in pixels; v_1 and v_2 are parameters that are used to model perspective parameters; and v is a dimensionless parameter which is generally different from zero. $w = v_1x + v_2y + v$ is an homogeneity factor related to perspective. This is the transformation model that we use in the present study.

2.2.2. Similarity measure

The diversity of cystoscopic image sequences results from variations of illumination conditions, anatomical differences between explored regions inside the bladder and anatomical differences between patients. Refer to Fig. 1 for examples of this variability. Therefore, it is not always possible to extract feature points or to

do region or contour segmentation robustly. In addition, in most sequences, translations are the predominant parameters (see Section 2.4) and their values between image pairs are small enough to consider that viewpoints between consecutive images are little different because of small movements. For these reasons, we selected a registration algorithm based on the direct utilization of gray levels of images and, in consequence, there is no need of segmentation operations. Similarity measure S is then computed directly from gray levels of images. In our application, cystoscopy with monomodality images and limited computation time, the sum of squared differences (SSD) of gray levels of images can be used:

$$\begin{aligned} SSD &= \sum_{x \in I_i \cap I_{i+1}} [I_{i+1}(x) \circ (T_{i,i+1}(M)) - I_i(x)]^2 \\ &= \sum_{x \in I_i \cap I_{i+1}} [I_{i+1}(T_{i,i+1}(x; M)) - I_i(x)]^2 \end{aligned} \quad (5)$$

where $I_{i+1}(T_{i,i+1}(x; M))$ stands for image I_{i+1} after application of transformation $T_{i,i+1}$, where the transformation parameters are found in matrix M . SSD is computed for every pixel of coordinates x that is common to both images I_i and I_{i+1} . This similarity measure has been widely and successfully used in applications such as object tracking [34,24], mosaicing of consumer-grade photographs [6,7], and biomedical images [21,20].

2.2.3. Forward and direct compositional algorithms

The definition of the SSD presented in the previous section was used in the registration algorithm proposed by Lucas & Kanade ($L\&K$) [35] to obtain optical flow between images of a sequence for a stereo vision application. Since then, this algorithm has been widely applied to moving object tracking and image mosaicing. In the original article, only translations were considered, but this algorithm has been extended for considering other transformation parameters, as we did in our application. Starting from known initial transformation parameters M , the $L\&K$ algorithm searches ΔM increments that, when added to the initial parameters, will provide minimal SSD between the transformed source image $I_{i+1}(T_{i,i+1}(x; M))$ and the target image $I_i(x)$. From Eq. (5) and with a known estimation of

M , the algorithm of $L&K$ iteratively minimizes:

$$SSD = \sum_{x \in I_i \cap I_{i+1}} [I_{i+1}(T_{i,i+1}(x; M + \Delta M)) - I_i(x)]^2 \quad (6)$$

for ΔM increments updating M , so that $M \leftarrow M + \Delta M$, and corresponding to the following transformation update:

$$T_{i,i+1}(x; M) \leftarrow T_{i,i+1}(x; M + \Delta M) \approx T_{i,i+1}(x; M) + \frac{\partial T_{i,i+1}}{\partial M} \Delta M \quad (7)$$

ΔM increments are obtained by a second order non-linear optimization (Gauss-Newton). The least-squares minimization problem SSD implies the computation of the first derivative of the cost function in Eq. (6). This is achieved by linearization of $I_{i+1}(T_{i,i+1}(x; M + \Delta M))$ using a first order Taylor series expansion with respect to ΔM , that is:

$$\begin{cases} SSD = \sum_{x \in I_i \cap I_{i+1}} [I_{i+1}(T_{i,i+1}(x; M)) + \nabla I_{i+1}(x) J_T(M) \Delta M - I_i(x)]^2 \\ \frac{\partial SSD}{\partial M} = \sum_{x \in I_i \cap I_{i+1}} [\nabla I_{i+1}(x) J_T(M)]^T [I_{i+1}(T_{i,i+1}(x; M)) \\ + \nabla I_{i+1}(x) J_T(M) \Delta M - I_i(x)] \end{cases}$$

with the Jacobian (the first derivative of transformed coordinates of the pixel located at (x, y) with respect to transformation parameters), the gray-level Gradient and the second derivative matrix (the Hessian) of source image $I_{i+1}(x)$, which are respectively:

$$\begin{cases} J_T(M) = \frac{\partial T_{i,i+1}(x; M)}{\partial M} = \begin{pmatrix} \frac{\partial T_{i,i+1}(x; M)}{\partial m_1} & \dots & \frac{\partial T_{i,i+1}(x; M)}{\partial m_9} \\ \frac{\partial T_{i,i+1}(y; M)}{\partial m_1} & \dots & \frac{\partial T_{i,i+1}(y; M)}{\partial m_9} \end{pmatrix} \\ \nabla I_{i+1}(x) = \begin{pmatrix} \frac{\partial I_{i+1}(x, y)}{\partial x} & \frac{\partial I_{i+1}(x, y)}{\partial y} \end{pmatrix} \\ H_{i+1} = \sum_{x \in I_i \cap I_{i+1}} \left[\nabla I_{i+1}(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right]^T \left[\nabla I_{i+1}(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right] \end{cases} \quad (9)$$

During this minimization process, these derivatives must be computed at every iteration, implying long-time computation operations ($O(n^2N)$ for the Hessian, with n being the number of transformation parameters and N the number of pixels). In [23], the authors observed that when consecutive images being registered are quite similar, as in our application, the gradient of the source image can be replaced by the one of the target image. By taking the target image I_i as source image to which the transformation $T_{i,i+1}$ is applied, and the source image I_{i+1} as target image, the Gradient, the Jacobian and the Hessian of I_i can be computed once and for all, leading to significantly reduced computation time without affecting robustness. In this inverse approach, they develop a composition algorithm that can be applied to homographic transformations, i.e. perspective.

The forward compositional approach, instead of the additive one, developed in [36,37,23] minimizes:

$$SSD = \sum_{x \in I_i \cap I_{i+1}} [I_{i+1}(T_{i,i+1}(T_{i,i+1}(x; \Delta M); M)) - I_i(x)]^2 \quad (10)$$

with respect to ΔM and then updates T at each iteration step, so that:

$$T_{i,i+1}(x; M) \leftarrow T_{i,i+1}(x; M) \circ T_{i,i+1}(x; \Delta M) = T_{i,i+1}(T_{i,i+1}(x; \Delta M); M). \quad (11)$$

The inverse compositional approach detailed in [23] minimizes the following expression:

$$SSD = \sum_{x \in I_i \cap I_{i+1}} [I_i(T_{i,i+1}(x; \Delta M)) - I_{i+1}(T_{i,i+1}(x; M))]^2 \quad (12)$$

with respect to ΔM and iteratively updates:

$$T_{i,i+1}(x; M) \leftarrow T_{i,i+1}(x; M) \circ T_{i,i+1}(x; \Delta M)^{-1} \quad (13)$$

After a first order Taylor series expansion, Eq. (12) becomes:

$$SSD = \sum_{x \in I_i \cap I_{i+1}} [I_i(T_{i,i+1}(x; \Delta M_0)) + \nabla I_i(x) J_T(M) \Delta M - I_{i+1}(T_{i,i+1}(x; M))]^2, \quad (14)$$

with ΔM_0 the initial increments matrix and $\nabla I_i(x)$ the gradient of initial target image over \vec{x} and \vec{y} .

Considering that the transformation matrix composed of initial increments ΔM_0 is an identity matrix, Eq. (14) is minimized so that:

$$\begin{aligned} \frac{\partial SSD}{\partial \Delta M} &= 2 \sum_{x \in I_i \cap I_{i+1}} \left[\nabla I_i(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right]^T \\ &\times \left[I_i(x) + \nabla I_i(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \Delta M - I_{i+1}(T_{i,i+1}(x; M)) \right]^2 = 0 \end{aligned} \quad (15)$$

or

$$\Delta M = H_i^{-1} \sum_{x \in I_i \cap I_{i+1}} \left[\nabla I_i(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right]^T [I_{i+1}(T_{i,i+1}(x; M)) + I_i(x)], \quad (16)$$

with

$$H_i = \sum_{x \in I_i \cap I_{i+1}} \left[\nabla I_i(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right]^T \left[\nabla I_i(x) \frac{\partial T_{i,i+1}(x; M)}{\partial M} \right] \quad (17)$$

being the Hessian matrix for image $I_i(x)$ and H^{-1} its inverse, whose value can be pre-computed once and for all before the iterative stage.

In our application, we experimentally verified that the forward and inverse compositional algorithms yielded very similar results but the latter was effectively faster, without losing robustness or precision. Consequently, we implemented this approach in our cystoscopic image mosaicing algorithm.

2.3. Image stitching

In mosaicing, the stitching of the pairwise registered images from the video sequence is performed by placing image I_{i+1} on a global coordinate system, $(O_g, \vec{x}_g, \vec{y}_g)$, which may correspond to the coordinate system of the first image in the sequence. Optimal local transformations, $\tilde{T}_{i,i+1}$, between images I_i and I_{i+1} obtained after completion of the registration algorithm are used to compute successive global transformations, \tilde{T}_{i+1} , so that:

$$\tilde{T}_{i+1} = \tilde{T}_{i,i+1} \times \tilde{T}_i = \prod_{j=0}^{i-1} \tilde{T}_{i-j,i-j+1} \quad (18)$$

in which $i \in [1; n-1]$ and n is the number of images in the sequence. This sequential way of placing images on the global coordinate system works only if image registration error is minimum. Mosaicing error, directly related to registration error, is evaluated in Section 3.

To smooth the gray-level discontinuities between the side pixels of image I_{i+1} and those co-localized in the panoramic image

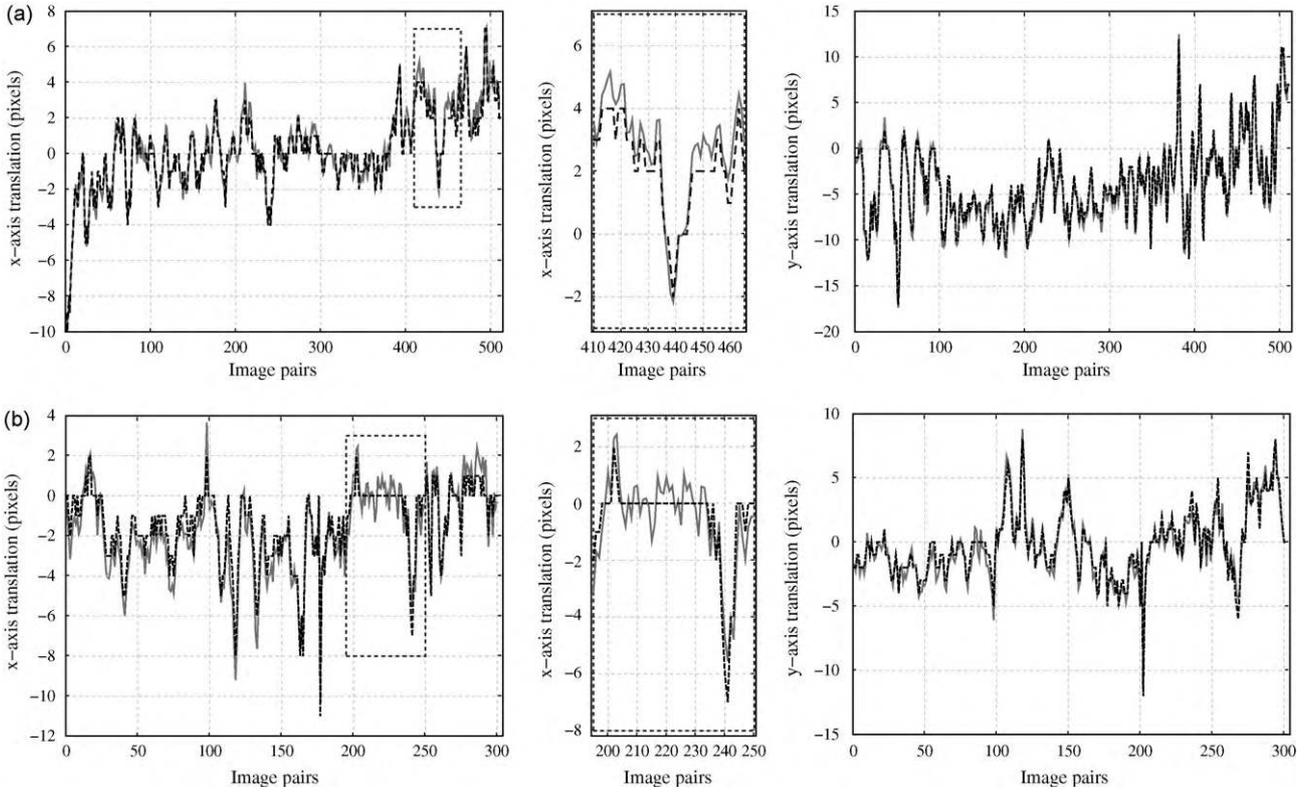


Fig. 3. Translation parameter values as a function of consecutive pairs of images obtained for a 511 image sequence (a) and a 301 image sequence (b) from two clinical cystoscopies (*in vivo*). Left: \bar{x} translations. Middle: close-up of the dashed rectangle on the corresponding left graphics. Right: \bar{y} translations. Black-dotted and gray lines represent translation values obtained from the cross-correlation algorithm (present study), and values from the registration algorithm based on Mutual Information (MI) [26].

already built, a spatial weighting function is applied. This is done by assigning pixel gray-level values that are weighted by a Gaussian function so that:

$$\begin{cases} lp_{i+1} = w_1 I_{i+1} + w_0 lp_i \\ w_1 = 0.9e^{r/\sigma^2} + 0.1 = 1 - w_0 \end{cases} \quad (19)$$

where lp_i and lp_{i+1} are respectively, the panoramic image before and after stitching of image I_{i+1} ; w_1 and w_0 are the weights assigned respectively to the pixels of the image to be stitched and to the ones of the panoramic image already constructed; r is the radial distance of a pixel, measured with reference to the stitched image center and σ is the standard deviation of the Gaussian function.

Based on human observation (qualitative) of the cystoscopic images, the value of σ was fixed at $0.25N$ ($N = 400$ pixels). As a result, gray-level continuity on image borders and also light gradient diffusion on the map are ensured.

2.4. Speed improvements

The speed of convergence of the registration process (related to the number of iterations) described in the previous section strongly depends on the initialization values of the transformation parameters. To rapidly initialize these parameters with values closer to an “optimal” solution, it is proposed to first compute the cross-correlation of images using their Fourier transforms [38], so that:

$$CC(x, y) = \mathcal{F}^{-1}[\mathcal{I}_i(u, v)\mathcal{I}_{i+1}^*(u, v)] \quad (20)$$

in which $\mathcal{I}_i(u, v)$ and $\mathcal{I}_{i+1}(u, v)$ are the Fourier transform of target and source images I_i and I_{i+1} respectively; (u, v) are the spatial frequencies; and $*$ indicates the complex conjugate.

The matrix $CC(x, y)$ thus computed has a correlation peak whose coordinates constitute a good approach of the exact translation

values existing between images being registered. However, this information from cross-correlation is only workable if variations of the other transformation parameters stay in the same limits, few degrees in rotation and up to 10% for scale factor.

In order to demonstrate the usefulness of the cross-correlation operation and to ensure its exploitability for the application presented here, a test was performed over 3860 pairs of images extracted from ten clinical cystoscopic sequences acquired *in vivo*. We compared the translation values obtained from cross-correlation implemented in the algorithm of the present study to those obtained from a registration algorithm based on MI mentioned earlier (Section 1.2) which had already been validated [26]. Results for two of these representative sequences are given in Fig. 3. The translation values obtained by the two approaches are globally similar. Mean value of the difference (\pm standard-deviation) between translations obtained for ten sequences is $0.04 (\pm 2.33)$ pixels over \bar{x} , and $0.02 (\pm 2.54)$ pixels over \bar{y} . The maximum difference is 34.22 pixels over \bar{x} and 34.44 pixels over \bar{y} . These results show that cross-correlation applied to cystoscopic images allows a simple and fast estimation of initial translation values close to the optimal solution. Results of efficient application of correlation highlight the fact that translations are the predominant varying transformation parameters for these clinical cystoscopic sequences. The image registration method described in Section 2.2 works well if the superposed areas between images to register is higher than 80% of image size, as it will be detailed in Section 3. When this is not the case, the registration process may be slowed down or even fail. Thus, for evaluating the superposition rate between all consecutive images of our test clinical sequences, the distribution of the absolute values of \bar{x} and \bar{y} translations (Fig. 4), obtained by registering the 3860 image pairs previously mentioned, was represented. These distributions indicate that the translation

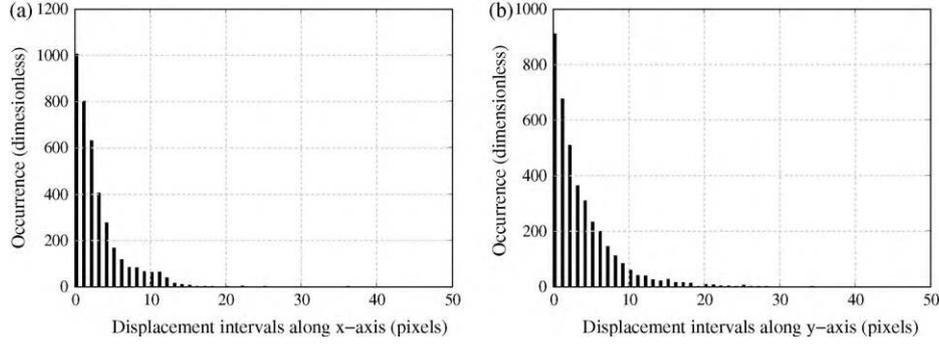


Fig. 4. Distribution of absolute values of translations for 10 image sequences (3868 image pairs) over \bar{x} -axis (a) and \bar{y} -axis (b).

values found are always lower than 50 pixels. Around 90% of the translation values are lower than 10 pixels. Mean translation values (\pm standard-deviation) over \bar{x} and \bar{y} are respectively, 3.20 (± 3.34) and 4.04 (± 4.37) pixels. For the tested sequences, the maximum translation value was 43.56 pixels (\bar{y} -axis). This high value occurred only once in a very specific sequence having the largest translation values among all tested sequences. Considering the size of images and the maximum values of translation obtained for the tested sequences, the superposition rate between two adjacent images is always greater than 80%. In addition, cystoscopic video recorded in clinics is characterized by many sequences where the cystoscope moves slowly or stays almost static. In these cases, a significant time is spent in registering images very close to each other, this is, barely transformed images. Consequently, the use of the translation estimations delivered by the cross-correlation stage to perform a dynamic selection of image pairs to be registered, is proposed. Based on the characterization of \bar{x} and \bar{y} translation distributions over the ten cystoscopic sequences, a conditional step has been added to the algorithm presented in this study. Thus, the iterative image registration and the computation of global transformation matrices are only performed when translations between images are greater than 9 pixels, that is 85% of the translation values in our test sequences. Once this threshold is reached, iterative registration of the current image and the last one registered, as well as update of the global transformation matrix are performed. Since the iterative part of the registration algorithm takes most of computation time, a substantial time saving is obtained by avoiding registration of images that are too close to each other.

3. Results

3.1. Quantitative evaluation of the mosaicing precision on phantom

Quantitative evaluation of mosaicing precision is difficult to be done with video sequences recorded in clinical conditions (*in vivo*), where the position coordinates of the viewpoint of the cystoscope into the bladder are unknown. To quantify the mosaicing precision, it is necessary to best simulate realistic conditions of acquisition while having a fine control of the cystoscope movements. To do so, an acquisition protocol on a phantom with precise control of 3D movements of the cystoscope was established. A multi-axis motorized micrometric positioning system (translation stages *UMR8.51* and linear actuators *LTA-HS*, *Newport*, 0.1 μm precision) was mounted for applying in plane translations and scale factor variations, that is, along \bar{x} , \bar{y} and \bar{z} axis, as well as in plane rotations (rotation stage *PR50CC*, *Newport*, 0.01° precision) and out of image plane rotations (rotation stage *URS75CC*, *Newport*, with 0.01° precision). The sequence of movements was controlled by a motor drive controller (*XPS*, *Newport*) with a LabView programmed

interface. Due to the size and weight of the acquisition system, the movements of the cystoscope were simulated by movements of the phantom on which the image acquisitions were performed. The phantom is a high-resolution printed photograph (10 \times 12 cm) of the internal surface of a pig bladder, previously excised and “wide opened”. A pig bladder is very similar to a human one in texture and anatomy.

The acquisition of an image sequence was performed following the predefined path sketched in Fig. 5(a). The course of the cystoscope involved movements along every axis composing the positioning system (2 rotations and 3 translations). Movements over the 1st acquisition path included a 42 mm translation in \bar{x} , combined to a 42 mm translation along \bar{z} (scale-factor variation). Movement over the 2nd path involved a 42 mm translation along \bar{y} coupled to in plane rotations from 0 to 20°. The 3rd path of movement is composed of a -42 mm translation along \bar{x} , returning to initial coordinate $x = 0$, with an off-plane rotation from 0 to 16° (perspective modification). The last set of movements, 4th acquisition path, involved translations along \bar{y} and \bar{z} , and in- and off-plane rotations simultaneously, so that the device finally goes back to its starting position. A total of 169 images were acquired (42 images for each acquisition path). The first (I_{001}) and last (I_{169}) images were obtained from the same position.

In order to obtain a measure of the construction error in the panoramic image, a grid of black dots regularly spaced was printed over the photograph of the bladder (see Fig. 5(b)). The panoramic image obtained after applying our mosaicing algorithm to the acquired image sequence is shown in Fig. 5(c). Mosaicing error computation was carried out in three steps. Firstly, four dots manually selected in the first image of the panoramic image are registered with their counterparts in the digitized photograph. To perform this registration, a rigid transformation model is used (2D translations, rotation and scale factor). Secondly, the segmentation of every visible dots (31 dots) in the panoramic image (Fig. 5(d)) and the automatic computation of centroid coordinates of these segmented dots are performed using morphological operators. Thirdly, these coordinates are compared to the known coordinates of the center of the dots printed on the photograph. Mosaicing precision given by the distance 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 dot centroids segmented in the photograph, so that:

$$e_m = \frac{1}{N_p} \sum_{j=1}^{N_p} \sqrt{(x_m^j - x_r^j)^2 + (y_m^j - y_r^j)^2}. \quad (21)$$

Fig. 5(d) shows the 31 segmented dots in the panoramic image 5(c) and their 31 counterpart dots in the printed photograph used as reference. Fig. 5(e) is an enlarged image of one of these dots. The

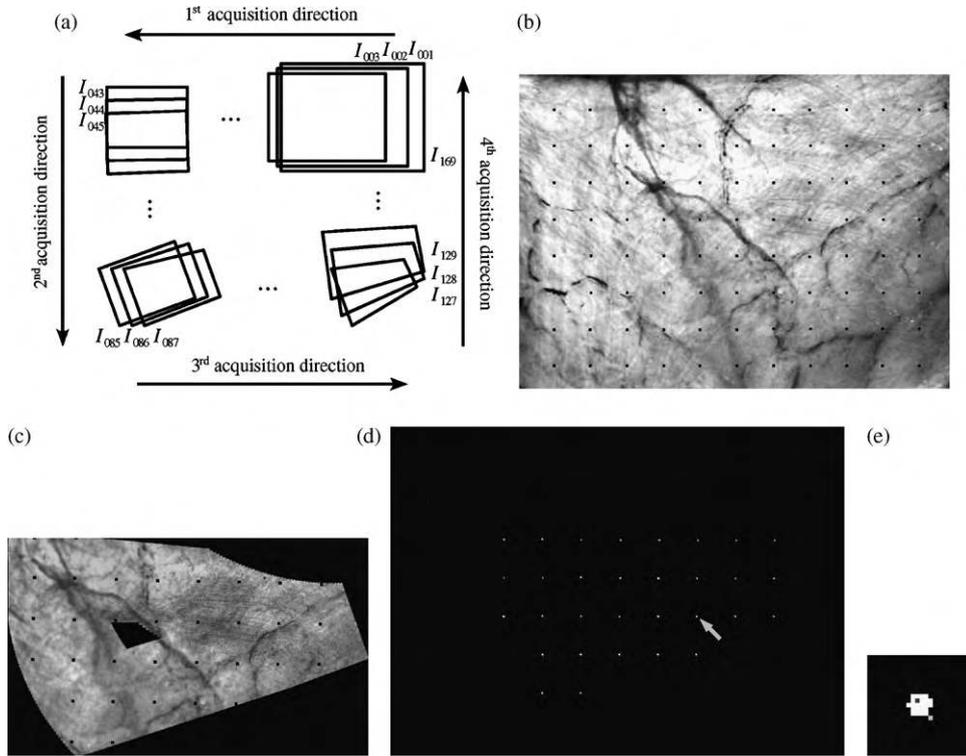


Fig. 5. (a) Illustration of the acquisition path of images $I_{001}, I_{002}, \dots, I_{169}$ applied onto the bladder photograph phantom for quantifying mosaicing precision. (b) Pig bladder photograph with a regular grid of black dots printed on it. These dots have been enlarged to facilitate their observation in this figure. (c) Panoramic image (1947×1187 pixels) constructed from 169 images acquired along the path in (a). Black dots have been enlarged. (d) Image of the dots (centroid of detected dots) segmented in the panoramic image (c) with reference dots located in the photograph (b) in the same coordinate system. (e) Enlarged image of one dot in image (d) (pointed by the gray arrow) showing the centroid of the detected dot (in light gray) and the segmented dot in the reference image (in dark gray). Mean distance between homologous dots is 3.82 pixels.

centroids (barycenters) of dots detected in the panoramic image are shown in light-gray, whereas the centroids of dots in the reference printed photograph are shown in dark-gray. At last, the grid of dots with regular vertical and horizontal spacing can be identified in the panoramic image, with a mean error of 3.82 pixels and a maximum error of 6.73 pixels. Considering that the size of the panoramic image is 1947×1187 pixels, the maximum error represents 0.57% of the image height (1187 pixels). This error percentage is very small and low enough for providing the physician a good quality panoramic image efficiently exploited in the case of the presented application. This point is further discussed in the following section.

3.2. Mosaicing of clinical data

Our mosaicing algorithm was applied to cystoscopic sequences acquired during conventional clinical procedures in order to demonstrate its applicability to real conditions (*in vivo*) and to perform a qualitative validation on clinical data. Among 6 video recordings, lasting from 4 to 11 min, obtained from 6 different patients, 10 image sequences lasting from several seconds to several dozens of seconds, were selected. We considered as particularly interesting those sequences where the air bubble, the ureters, the trigone, some scars or polyps may be observed. The primary interest of these anatomical structures is not to give exploitable elements to the registration algorithm, but for the physician to more easily associate the panoramic image to a particular anatomical place in the bladder. The number of images in the sequences vary from 150 images (6 s) to 1300 images (52 s). Representative results for 3 of these sequences are given in Figs. 6–8. It is worth noticing that for these sequences, where lighting conditions and texture features vary significantly, our approach is robust enough to handle variations observed in images. The urologists who analyzed these

panoramic images confirmed their visual coherence. The mosaicing errors, which were not measured here, were estimated small enough to allow the clinician to properly interpret or analyze the image qualitatively. The mosaicing algorithm was programmed in C language using the OpenCV vision library. Panoramic images were constructed using a Pentium IV 3.2 GHz, 2 Gb RAM computer. A mean number of 12 iterations was needed by the optimization algorithm to register a pair of images. Registration time for an image pair varied between 1.2 and 1.5 s. Total construction time, registration and stitching, was 9.31 min for image in Fig. 6 (301 images), 9.39 min for image in Fig. 7 (530 images) and 12.69 min for image in Fig. 8 (450 images). These computation times makes it possible to construct a partial panoramic image of the bladder during the standard cystoscopic examination procedure. The next section gives details on how to further reduce these computation times by not registering the totality of images in a sequence but only those separated by significant variations of translations.

3.3. Improvements and comparisons

According to the study of the occurrence of translation values between images in the 10 cystoscopic sequences used (see Section 2.4), we implemented the proposed method of conditional (selective) registration for translations lower than 9 pixels. Results show a significant reduction of computation time with this method of image selection. The registration and stitching times decreased by 40% compared to the algorithm without selection. This means that, computation times are significantly reduced, from 9.31 down to 4.74 min, from 9.39 to 6.91 min and from 12.69 to 6.70 min for images shown in Figs. 6, 7 and 8. With this decrease in computation time, constructing panoramic images of the internal wall of bladders in a time duration lower than the one of the cystoscopic examination is possible. In addition, we observed that panoramic

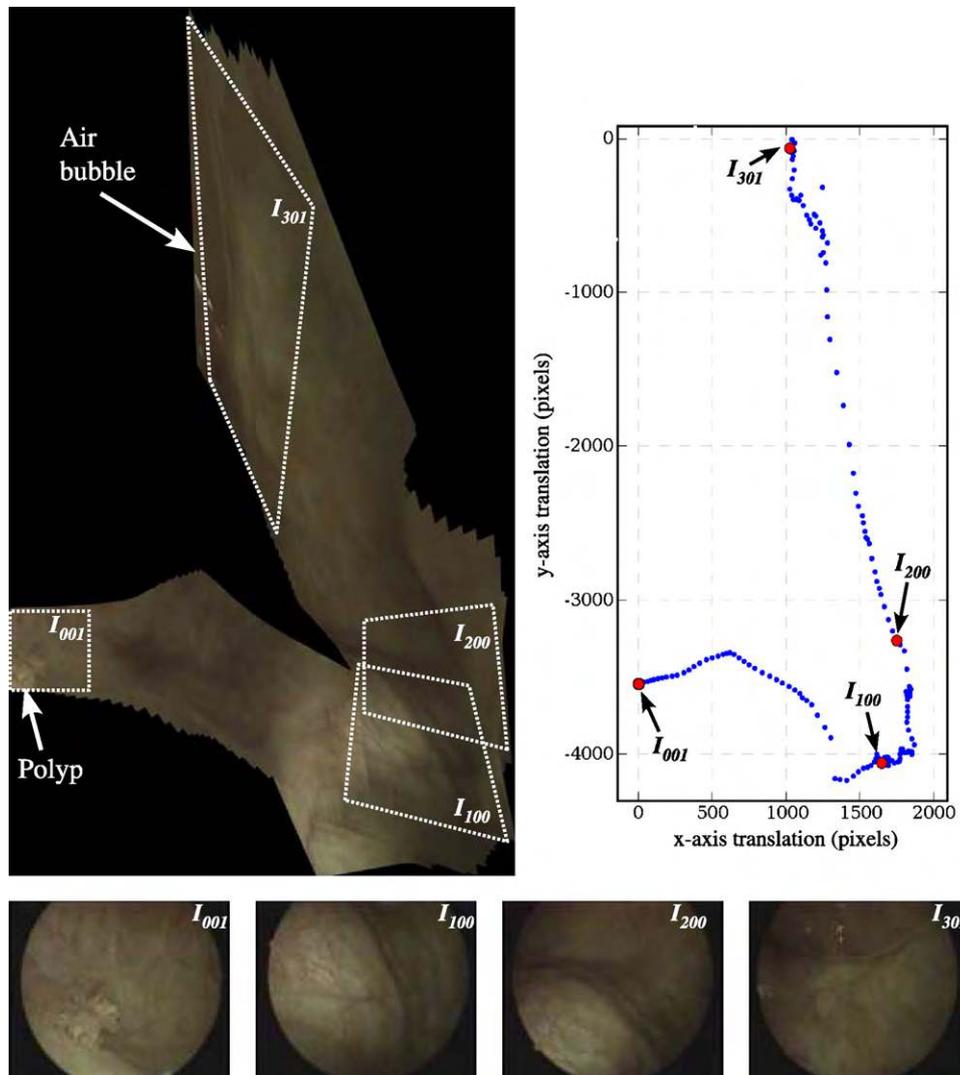


Fig. 6. 2544 × 4362 pixels panoramic image constructed from a 301 images sequence. In this image, a polyp can be located (at the bottom left) in relation to the air bubble visible at the top. The graphic on the right shows the course of the endoscope during the acquisition of this sequence, represented thanks to the computed transformation parameters.

images constructed with and without image selection are visually quite similar.

To date, the only published work on bladder mosaicing was presented in [26]. Using a same sequence of cystoscopic images (*in vivo*), we compared the panoramic image constructed with the mosaicing algorithm presented in [26] and the one developed in the present study. As seen in Fig. 9, a great visual similarity can be observed. The mosaicing mean error using the algorithm developed in [26] is 3.35 pixels after accumulated error correction (43.14 pixels without correction). Using our algorithm, a mean error of 3.82 pixels and a maximum error of 6.73 pixels were obtained without accumulated error correction. Construction of the same panoramic image takes some minutes using our method, while the MI-based algorithm used in [26] takes several hours.

In Section 2.1, we mentioned that image pre-processing is essential in our application for the registration algorithm to properly work. To highlight the effect of the fiber pattern and lighting inhomogeneities on the registration process, the two similarity measures (SSD and MI) were computed for two regions extracted from two images acquired with a fiberscope. The translations between images are known *a priori*: $t_x = 8$ pixels (x -axis) and $t_y = -47$ pixels (y -axis). No other geometric transformation between images is involved. Without applying image pre-processing, the

coordinates of the two parameter space extremum found are wrong: $(t_x, t_y) = (-1, -50)$ for MI and $(t_x, t_y) = (0, -46)$ for SSD. Fig. 10(a) and (b) illustrates the effect of these artifacts on MI and SSD respectively. These cost functions or similarity measures, are affected by many local extrema that make it harder for the optimization algorithm to converge towards a correct value of transformation parameters. When pre-processing is applied to these images, correct results are obtained: $(t_x, t_y) = (8, -47)$ for the two similarity measures.

4. Discussion: limits and applicability of the presented algorithm

The mosaicing method presented in this paper was specifically developed for building panoramic images of the internal wall of human bladders. This approach can be also applied to other hollow organs, such as the stomach. However, its applicability to tubular hollow organs, like the colon and the esophagus, is more difficult because of the endoscope orientation, where severe perspective deformations may occur. As for any other registration algorithms, the two main factors affecting the registration process in our algorithm are image quality (color contrast and deformations produced by the cystoscope orientation) and *intra* and *inter*-patient image

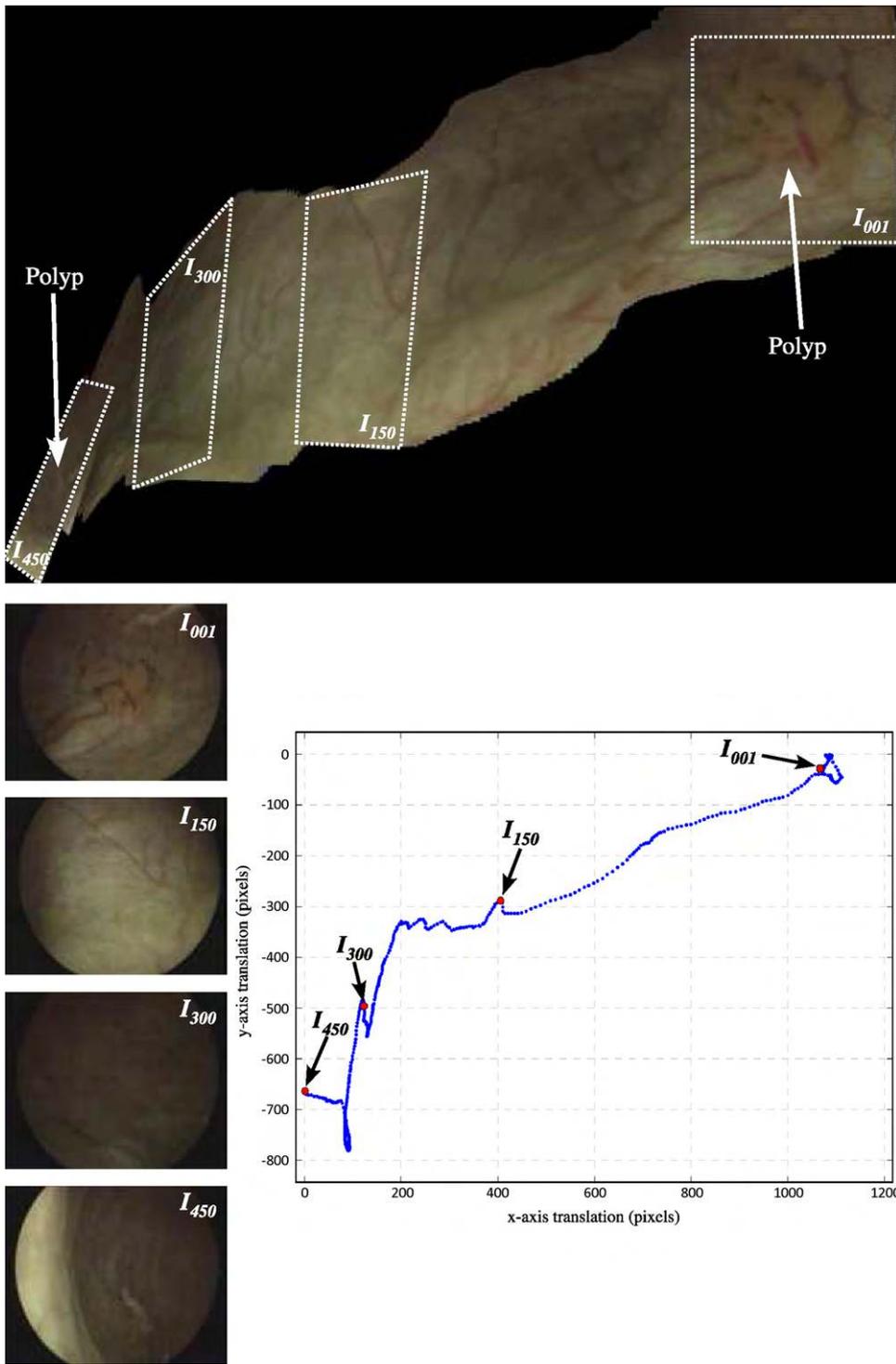


Fig. 7. 1479×1049 pixels panoramic image constructed from a 450 image sequence. Two polyps are visible on the top-right and at the bottom left of the image. In this panoramic image, both polyps can be accurately located in relation to each other. The 4 images shown on the left, below the larger image were extracted from the sequence. They highlight the great variations in lighting conditions.

variability. Although a great variability exists between characteristics of cystoscopic images in a same sequence, generally, lighting and texture conditions do not change drastically from one image to the next in the sequence. The bladder photograph used to evaluate registration precision was also used to quantify the limits of the registration method. So, we applied elementary transformation combinations controlled by the positioning system described in Section 3.1. Considering realistic limits related to our endoscopic application, the maximum values of transformation parameters

(see Eq. (4)) applied to images were: ± 50 pixels for t_x and t_y translations and $\pm 10^\circ$ for rotation in the image plane (ϕ angle). Isometric variations in scale factor (related to parameters a_{11} and a_{22}) were fixed to $\pm 10\%$. Variations in viewpoint, or perspective, were simulated by $\pm 10^\circ$ rotations of angles θ and ψ , around x - and y -axis respectively. Note that, considering the acquisition video rate (25 s) and the limited speed of the cystoscope movements during a clinical examination (few mm/s), the values of these parameters in a standard sequence are much lower than those considered in our

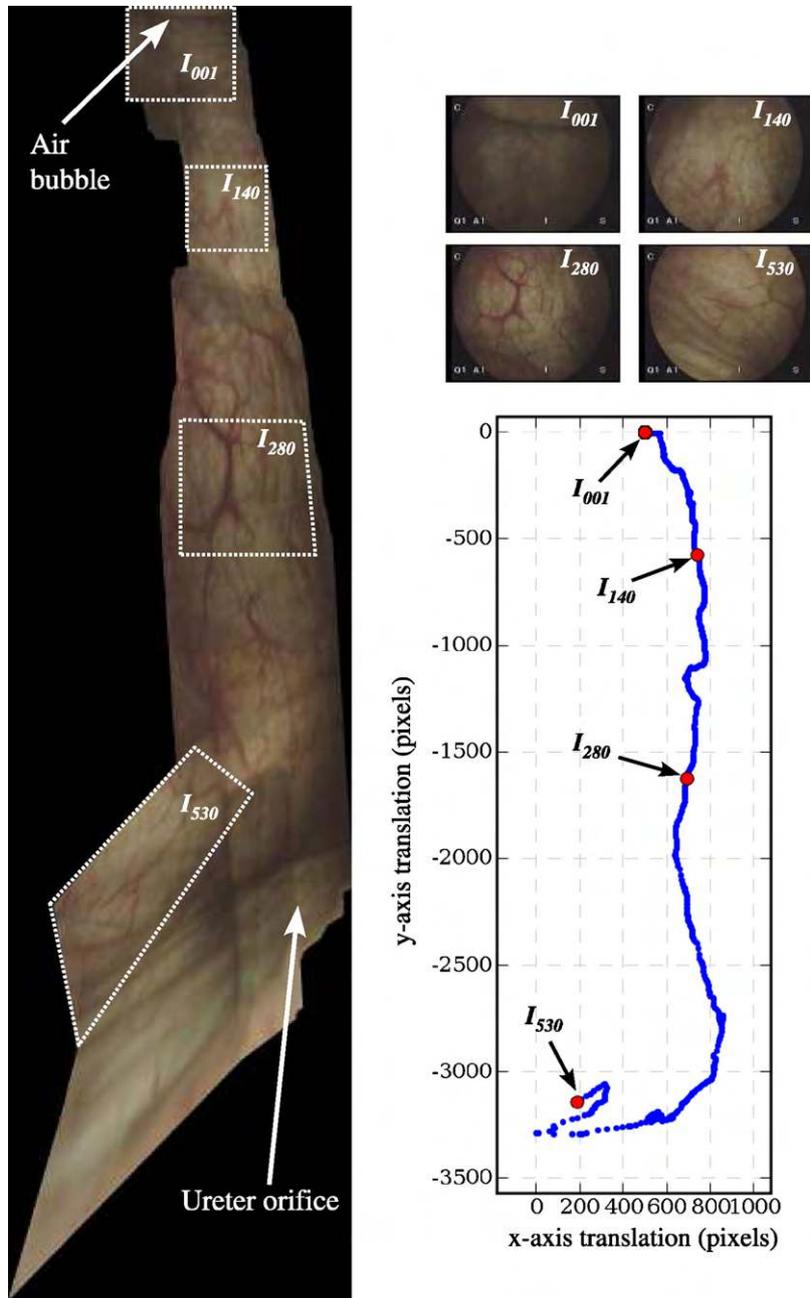


Fig. 8. 1458 × 5471 pixels panoramic image constructed from 530 images. The air bubble at the top and an ureter orifice at the opposite side of the bladder (bottom right) may be observed simultaneously on this panoramic image.

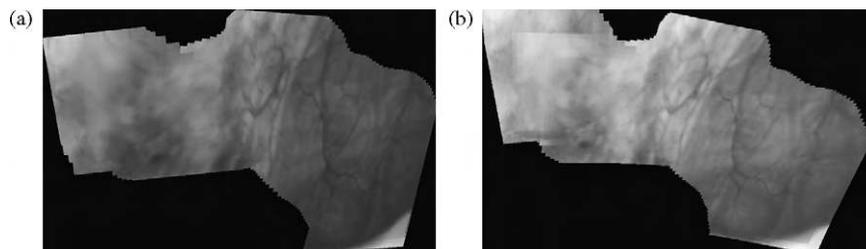


Fig. 9. (a) 1947 × 1187 pixels panoramic image from a sequence of cystoscopic images constructed using the algorithm presented in this study. (b) Same panoramic image constructed using the MI-based mosaicing algorithm presented in [26].

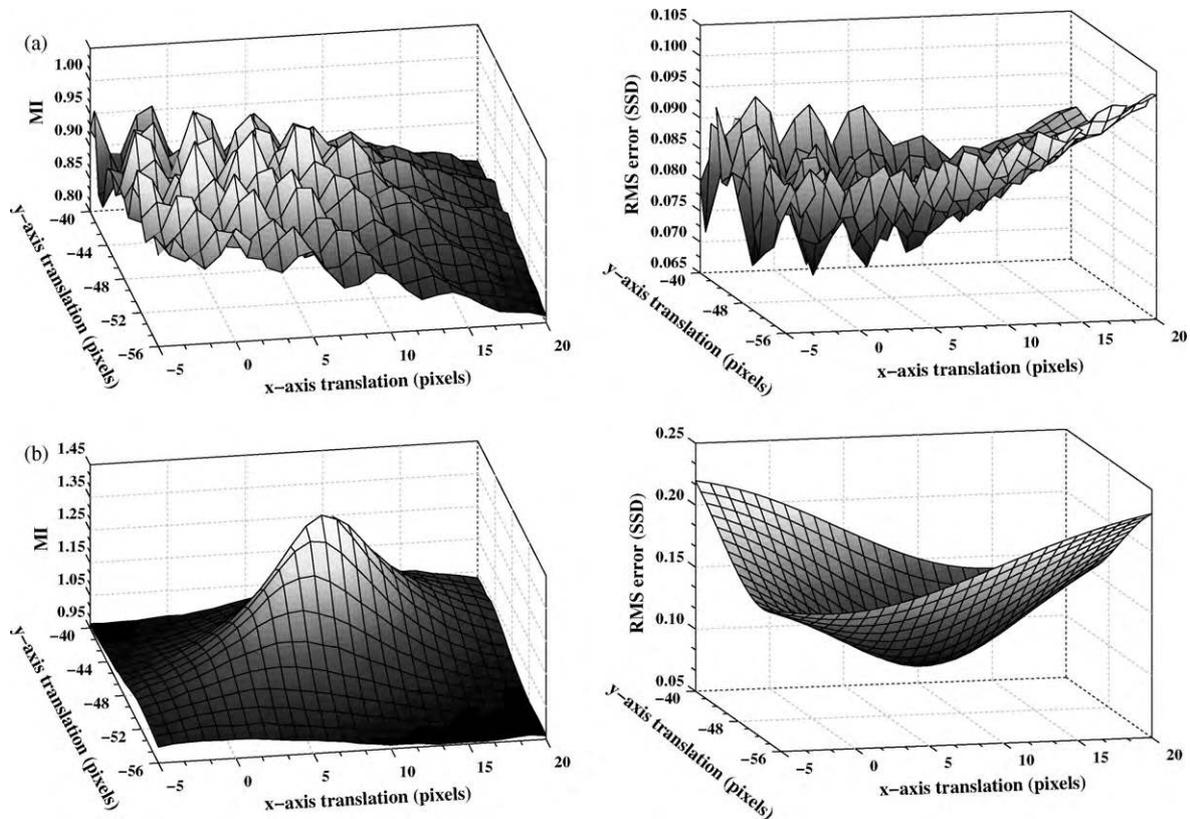


Fig. 10. Effect produced by the fiber pattern and lighting inhomogeneities on MI (top-left) and on the absolute value of SSD (top-right) computed using images without pre-processing (a) and with pre-processing (b). After pre-processing, the cost functions are convex and a unique global extremum exist that corresponds to the correct value of translations (8, -47)

tests. The results of these tests indicate that the impact of rotation angle ϕ (around z-axis) on the algorithm performance is important. If $\phi \approx 10^\circ$ and other transformations are also significant, registration fails. However, we found that for a superposition rate of 75% of image size and for in-plane rotations lower than 8° , correct registration of images with a maximum error less than 1 pixel is systematically obtained. In practice, these rotations are smaller than 8° . This error is very small indeed, because the obtained deviation from the true position represents 0.4% of the height of images (400×400 pixels). In addition, the conditions of the video acquisitions on our phantom – pig bladder excised then “wide opened” – and on a human bladders – clinical *in vivo* – are relatively different. More precisely, pig bladder images as acquired are globally less contrasted, due to weak lighting conditions, non-immersed in a liquid media, light reflections and mechanical strain of the bladder wall, than human bladder images acquired in clinics. The differences between images acquired on the pig bladder photograph and on human bladders *in vivo* are more important than the *intra-patient* differences. However, our algorithm was successfully applied to both types of sequences. The good results obtained for various sequences acquired for different patients show the ability of our approach to deal with image variability due to *intra-* and *inter-patient* differences.

Finally, from the point of view of clinical application, visual coherence corresponding to continuity and preservation of the anatomical information is a parameter at least as important as mosaicing error. Therefore, the panoramic images constructed and their corresponding video sequences were compared by the physicians to analyze the presence and continuity of blood vessels, scars and other anatomical features visible on the bladder wall images. In all panoramic images, the clinicians found that they were visually

fully coherent. Thanks to this easy-to-use and easy-to-store media, they were able to rapidly observe specific zones of interest, such as lesions, and to precisely locate them in the bladder.

5. Conclusion and perspectives

We developed a fast and automatic mosaicing algorithm applicable to cystoscopic video sequences. Geometric transformations between image pairs are modeled by a perspective transformation. The presented mosaicing algorithm consists of a fast initialization of translation parameters by an image cross-correlation step, followed by a limited but efficient optimization of all the transformation parameters using an inverse composition-based registration scheme, and finally a projection of successively registered images onto a global common coordinate system. A test protocol performed on a phantom allowed us to measure a mosaicing error lower than 0.3% of the panoramic image size, considering its width. We performed a qualitative evaluation of the panoramic images constructed from 10 different image sequences obtained from standard cystoscopic examinations. An analysis of the distribution of translation values existing between images for these clinical sequences was carried out and exploited to modify our algorithm for dynamic and automatic selection of the image pairs to be registered. In this way, construction of panoramic images of regions of a bladder takes only some minutes. Even if image mosaicing in real time was not reached (with video rate of 25 s), the computation time taken by our algorithm, allows the physician to obtain an exploitable panoramic image before ending the cystoscopic clinical examination. All panoramic images built from clinical cystoscopies constitute easy-to-use and easy-to-store media but are also visually fully coherent for the clinicians who are able to rapidly observe

specific zones of interest and to precisely locate them in the bladder. Complementary work is being carried out to further reduce this computation time using image position prediction techniques based on the translation information between previous images of the sequence.

Acknowledgments

The authors express their gratitude to the “Region Lorraine” the “Ligue Contre le Cancer (CD 52, 54)” and the CONACYT (grant) for their financial support. Authors thank Dr. M.-A. D’Hallewin from the Anticancer center “Centre Alexis Vautrin”, Nancy, France, for her useful comments, and the Experimental Surgery Laboratory at the Faculty of Medicine in Nancy for the fresh pig bladders excisions.

References

- [1] Tyczynski J, Parkin D. Bladder cancer in Europe, ENCR Cancer Fact Sheets. European Network of Cancer Registries. International Agency for Research on Cancer 3.
- [2] Chopin DK, Gattengo B. Résumé du rapport 2001 de l'association française d'urologie (AFU) sur les tumeurs superficielles de la vessie; 2001.
- [3] Holzbeierlein J, Lopez-Corona E, Bochner B, Herr H, Donat S, Russo P, et al. Partial cystectomy: a contemporary review of the memorial sloan-kettering cancer center experience and recommendations for patient selection. *Journal of Urology* 2004;172(3):878–81.
- [4] Blondel W, Daul C, Villette S, Miranda-Luna R, Hernandez-Mier Y, Bourg-Heckly G, et al. Autofluorescence endoscopic spectro-imaging and 2D-cartography for in situ localisation and diagnosis of cancerous lesions. *Traitement du signal* 2005;23:213–33.
- [5] Brown M, Lowe DG, p. 1218–27 Recognising panoramas. In: *Proceedings of the 9th IEEE international conference on computer vision (ICCV2003)*, vol. 2. 2003.
- [6] Baudisch P, Tan D, Steedly D, Rudolph E, Uyttendaele M, Pal C, et al. p. 1–10 Panoramic viewfinder: providing a real-time preview to help users avoid flaws in panoramic pictures. In: *OZCHI '05: proceedings of the 19th conference of the computer-human interaction special interest group (CHISIG) of Australia on computer-human interaction*. 2005.
- [7] Steedly D, Pal C, Szeliski R, p. 15–21 Efficiently registering video into panoramic mosaics. In: *Proceedings of the IEEE international conference on computer vision (ICCV 2005)*. 2005.
- [8] Litvin A, Konrad J, Karl WC, p. 663–74 Probabilistic video stabilization using kalman filtering and mosaicking. In: *Proceedings of the IS&T/SPIE symposium on electronic imaging, image and video*, vol. 5022. 2003.
- [9] Matsushita Y, Ofek E, Tang X, Shum H, p. 50–7 Full-frame video stabilization. In: *Proceedings of 2005 IEEE international conference on computer vision and pattern recognition (CVPR)*, vol. 1. 2005.
- [10] Szeliski R. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications* 1996;16(2):22–30.
- [11] Szeliski R, Shum H, p. 251–58 Creating full view panoramic image mosaics and environment maps. In: *SIGGRAPH '97: Proceedings of the 24th annual conference on computer graphics and interactive techniques*, vol. 31. 1997.
- [12] Douze M. Estimation d'homographies inter-images. cas des mosaïques et du suivi en temps réel. applications en réalité augmentée, Ph.D. thesis, Institut National Polytechnique de Toulouse; 2004.
- [13] Rzhanov Y, Huff L, Cutter Jr R, p. 4 Seafloor video mapping: modeling, algorithms, apparatus. In: *Proceedings of the international conference on image processing (ICIP2002)*. 2002.
- [14] Gracias N, Negahdaripour S, p. 1234–40 Underwater mosaic creation using video sequences from different altitudes. In: *Proceedings of the MTS/IEEE OCEANS'05*, vol. 1. 2005.
- [15] Zomet A, Peleg S, p. 579–83 Efficient super-resolution and applications to mosaics. In: *Proceedings of the 15th international conference on pattern recognition (ICPR'00)*, vol. 1. 2000.
- [16] Capel DP. Image mosaicing and super-resolution, Ph.D. thesis, Robotics Research Group. Department of Engineering Science. University of Oxford; 2001.
- [17] Capel D, Zisserman A. Computer vision applied to super-resolution. *IEEE Signal Processing Magazine* 2003;20(3):75–86.
- [18] Gee AH, Treece GM, Prager RW, Cash C, Berman L. Rapid registration for wide field-of-view freehand three-dimensional ultrasound. *IEEE Transactions on Medical Imaging* 2003;22(11):1344–57.
- [19] Yaniv Z, Joskowicz L. Long bone panoramas from fluoroscopic X-ray images. *IEEE Transactions on Medical Imaging* 2004;23(1):26–35.
- [20] Zhou Y, Wan M, Xue H. A tunable incremental factor augmented inverse image alignment method in fundus angiogram registration and mosaicing. *Computerized Medical Imaging and Graphics* 2004;28:219–24.
- [21] Zhou Y, Xue H, Wan M. Inverse image alignment method for image mosaicing and video stabilization in fundus indocyanine green angiography under confocal scanning laser ophthalmoscope. *Computerized Medical Imaging and Graphics* 2003;27:513–23.
- [22] Baker S, Gross R, Matthews I, Ishikawa T. Lucas-Kanade 20 years on: a unifying framework: Part 2, Tech. Re CMU-RI-TR-03-01, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (Février 2003).
- [23] Baker S, Matthews I. Lucas-Kanade 20 years on: a unifying framework: Part 1, Tech. Re CMU-RI-TR-02-16, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA; July 2002.
- [24] Baker S, Matthews I, Xiao J, Gross R, Ishikawa T, Kanade T. Real-time non-rigid driver head tracking for driver mental state estimation, Tech. Re CMU-RI-TR-04-10, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA; October 2004.
- [25] Can A, Stewart CV, Roysam B, Tanenbaum HL. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002;24(3):347–64.
- [26] Miranda-Luna R, Daul C, Blondel W, Hernandez-Mier Y, Wolf D, Guillemin F. Mosaicing of bladder endoscopic image sequences: distortion calibration and registration algorithm. *IEEE Transactions on Biomedical Engineering* 2008;55(2):541–53.
- [27] Miranda Luna R. Contribution au mosaïquage d'images endoscopiques d'organes creux pour l'aide au diagnostic du cancer en phase précoce, Ph.D. thesis, Institut National Polytechnique de Lorraine (2005).
- [28] Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in mri. *IEEE Transactions on Medical Imaging* 2007;26(3):405–21.
- [29] Leong F, Brady M, McGee J. Correction of uneven illumination (vignetting) in digital microscopy images. *Journal of Clinical Pathology* 2003;56:619–21.
- [30] Miranda-Luna R, Hernandez-Mier Y, Daul C, Blondel W, Wolf D, p. 530–5 Mosaicing of medical video-endoscopic images: data quality improvement and algorithm testing. In: *IEEE Proceedings of the 1st international conference on electrical and electronics engineering*. 2004 (ICEEE). 2004.
- [31] Miranda-Luna R, Blondel W, Daul C, Hernandez-Mier Y, Posada R, Wolf D, p. 3383–6 A simplified method of endoscopic image distortion correction based on grey level registration. In: *IEEE proceedings of the 2004 international conference on image processing (ICIP '04)*, vol. 5. 2004.
- [32] Pirrone EAR, Gambino O. Illumination correction on mr images. *Journal of Clinical Monitoring and Computing* 2006;20:391–8.
- [33] Winter C, Rupp S, Elter M, Münzenmayer C, Gerhäuser H, Wittenberg T. Technique of distortion correction in endoscopic images using a polynomial expansion. *IEEE Transactions on Biomedical Engineering* 2006;53(10):2035–46.
- [34] Ng KC, Ishiguro H, Trivedi M, Sogo T. An integrated surveillance system-human tracking and view synthesis using multiple omni-directional vision sensors. *Image and Vision Computing* 2004;22(7):551–61.
- [35] Lucas BD, Kanade T, p. 121–30 An iterative image registration technique with an application to stereo vision (darpa). In: *Proceedings of the 1981 DARPA image understanding workshop*. 1981.
- [36] Shum H, Szeliski R. Panoramic image mosaics, Tech. Re MSR-TR-97-23, Microsoft Research, Microsoft Corporation, Redmond, WA 98052; December 1997.
- [37] Shum H, Szeliski R. Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision* 2000;16(1):63–84.
- [38] Kuglin CD, Hines DC, p. 163–5 The phase correlation image alignment method. In: *Proceedings of the 1975 IEEE international conference cybernetics society*. 1975.