CONSTRUCTING PANORAMIC VIEWS OF INTERNAL WALLS OF A BLADDER FROM CYSTOSCOPIC IMAGE SEQUENCES

Y. Hernández–Mier*, W. C. P. M. Blondel*, Ch. Daul* and D. Wolf*

* CNRS UMR 7039 CRAN (Centre de Recherche en Automatique de Nancy)
2, avenue de la Forêt de Haye, 54516 Vandoeuvre–Lès–Nancy Cedex, France.

yahir.hernandez–mier@ensem.inpl–nancy.fr

Abstract: Cystoscopes are devices allowing the observation of internal walls of a bladder. However, their field–of–view is reduced and the clinician can only observe a little part of the zone of interest. A panoramic representation of an area under a cystoscopic exam could represent a visual support to the clinician. This could help in the localization of eventual lesions and in the application of subsequent exams. And if this panoramic view could be constructed in a fast enough way, it would allow the clinician to evaluate the zone of interest while the patient is still under examination. This work describes the methodology involved in the automatic and efficient construction of panoramic views from sequences of images taken from a cystoscopic clinical exam. The proposed solution is divided into three parts. The first part is the preprocessing of input images. The second stage is an efficient image registration algorithm to find projective transformations between adjacent pairs of images. The last part is a stitching procedure that applies transformations found in the second stage to the sequence of images to form a panoramic view of the observed zone. Visual results and timing for registration and stitching processes are presented for three image sequences. Finally, a quantitative evaluation of the construction of a panoramic view is given.

Introduction

Cystoscopy is a clinical standard exam using a thin, lighted instrument (called a cystoscope) inserted into the urethra, allowing the observation of internal walls of the bladder. Cystoscopes can generate and store image sequences in the form of a video. However, the distance from the cystoscope to the walls of the bladder allows the clinician to observe only a reduced part of the investigated tissue area. A panoramic representation of an area under a cystoscopic procedure could represent an innovative visual support in the localization of eventual lesions and post–operative observations, and could be used as a guide in the application of subsequent exams. Moreover, the construction of these panoramas in a fast enough way, would allow the clinician to evaluate the zone of interest while the patient is still under examination.
The use of image mosaicing techniques to obtain a wide field–of–view (FOV) image of a scene is very popular. It authorizes an increase of the FOV while it preserves spatial resolution. It consists in extracting images as a camera moves, finding geometric transformations between images and finally, stitching them to form a larger image. Examples of applications of image mosaicing are found in consumer photography[1][2] as well as in different scientific disciplines [3][4][5][6]. In medicine, it has been used successfully in ophthalmology [7][8] and ultrasonics [9] but none in endoscopy.
This work describes the methodology involved in the automatic and efficient construction of panoramic views from sequences of images taken from a cystoscopic video. The used sequences were acquired with an uncalibrated cystoscope in a routinary clinical exam. The panoramic–view construction process starts with an image adaptation step that consists in a band–pass filtering. Registration of the filtered images is performed using Fourier techniques and minimization of intensity difference between images. Finally, panoramic views are built and visual rendering improved by the application of a weighted average function over overlapping pixels. Visual and timing results are showed for three image sequences as well as a quantitative evaluation of the error in the construction of a panoramic view of a test image sequence.

Materials and Methods

Three panoramic views were constructed from three image sequences containing 80, 162 and 193 images that were extracted from a cystoscopy video. This video was acquired during a standard clinical exam of human bladder. The duration of the sequences varies from 16 s to 20 s. Images in sequences were extracted by considering a minimal overlapping of 70%.
Mosaicing algorithm: Raw images cannot be directly used in the registration process. The sequences of images were band–pass filtered using a pyramidal framework to eliminate illumination inhomogeneities and to reduce blurring caused by the movement of internal walls of the bladder and the fast displacements of the cystoscopic tip. A first filtered image was produced using a three–level Gaussian pyramid. At the same time, a second image was obtained by applying a four–level Gaussian pyramid. Then, these images were subtracted to produce a band–pass filtered image. A central square of 256

ISSN: 1727-1983

x 256 pixels was taken as region of interest (ROI) in every image to reduce effects caused by a weak radial distortion (not corrected in this case).
These preprocessed images can now be used in the registration procedure. The first step in registration, is the application of the well known phase correlation registration algorithm [10]. This algorithm is based on the fact that a displacement in the spatial domain corresponds to a linear change in the the Fourier domain, as indicated by the shift property of the Fourier transform. This is, if an image $I_0(\mathbf{x})$, where $\mathbf{x} = (x,y)^T$, differ only by a translation $\mathbf{x_0} = (x_0,y_0)^T$ from image $I_1(\mathbf{x})$ [i.e., $I_1(\mathbf{x}) = I_0(\mathbf{x}-\mathbf{x_0})$], their Fourier transform, denoted by $F_0(\mathbf{u})$ and $F_1(\mathbf{u})$, where $\mathbf{u} = (u,v)^T$, are related by:

$$F_1(\mathbf{u}) = e^{-j2\pi(\mathbf{u}\cdot\mathbf{x_0})}F_0(\mathbf{u}). \quad (1)$$

Computing the correlation function between Fourier transforms, we have

$$R(\mathbf{u}) = \frac{F_0(\mathbf{u})F_1^*(\mathbf{u})}{|F_0(\mathbf{u})||F_1(\mathbf{u})|}, \quad (2)$$

where * denotes the complex conjugate.
By taking the inverse Fourier transform of R we obtain a function whose peak value indicates the pixel translation in the spatial domain [11]. This would be:

$$r(\mathbf{x}) = \delta(\mathbf{x}-\mathbf{x_0}). \quad (3)$$

This method delivers, in a fast way, a first approach of the translation parameters values between images at pixel level.
But translation is not the only transformation found in a cystoscopic sequence. Discarding radial distortion in images and considering the degrees of freedom of the cystoscope, a projective transformation model was chosen as the more appropriate model to register adjacent images in the sequence. The displacement obtained by the phase correlation algorithm was used to initialize an iterative process to find the projective transformations, at sub–pixel level, between each adjacent pair of images. The iterative process is based on the efficient inverse composite registration algorithm developed in [12]. This algorithm is a generalization of the Lucas–Kanade algorithm [13] to parametric motion models. Its objective is to align an image $I_1(\mathbf{x})$ to a template image $I_0(\mathbf{x})$. We consider a transformation function $T(\mathbf{x};\mathbf{p})$, that maps a pixel position $\mathbf{x}$ to a pixel position $\mathbf{x}'$ with a set of projective transformation parameters $\mathbf{p} = (p_0, p_1, \ldots, p_8)^T$. This algorithm assumes that the current transformation parameters $\mathbf{p}$ are known and iteratively minimizes

$$\sum_x [I_0(T(\mathbf{x};\Delta\mathbf{p})) - I_1(T(\mathbf{x};\mathbf{p}))]^2, \quad (4)$$

with respect to an increment $\Delta\mathbf{p}$ of the transformation parameters. The summation is done over a defined ROI. At every iteration, the transformation function is updated by a bilinear combination (denoted by the $\circ$ operator) of the parameters of $T(\mathbf{x};\mathbf{p})$ and $T(\mathbf{x};\Delta\mathbf{p})$.

$$T(\mathbf{x};\mathbf{p})_{new} = T(\mathbf{x};\mathbf{p})_{old} \circ T(\mathbf{x};\Delta\mathbf{p})^{-1}. \quad (5)$$

The transformation parameters obtained for every image are stored and then used to stitch images to form a panoramic view. The stitching process is performed by assigning a weighted average value to superposed pixels by a Gaussian function,

$$\begin{aligned} w_1 &= 0.9e^{-\frac{r}{2\sigma^2}} + 0.1; \quad (6)\\ w_0 &= 1 - w_1, \end{aligned}$$

where w_1 is the weight of pixels of the image to stitch and w_0 is the weight of the existing panoramic image; r is the radial distance of a pixel measured from the center of the image and σ is the opening of the Gaussian function. These weights are then multiplied by the intensity values of the pixels in the image being stitched, $I_1(\mathbf{x})$, and the intensity values of the existing panoramic image $I_m(\mathbf{x})$,

$$I_{pano}(\mathbf{x}) = w_1 I_1(\mathbf{x}) + w_0 I_m(\mathbf{x}), \quad (7)$$

where $I_{pano}(\mathbf{x})$ is the new panoramic image, including $I_1(\mathbf{x})$. This was applied $\forall\, I_1(\mathbf{x}), I_m(\mathbf{x}) \neq 0$. This process eliminates or attenuates visual artifacts at image borders. The previous algorithms were programmed in Matlab. The program was divided into three parts. The first part performs the pre–processing of images (filtering). The second one registers images and stores transformation parameters. The last part constructs the panoramic view and applies the weighting function to pixels being stitched. Execution times for the registration and stitching processes were computed on a PC 2GHz 1Gb RAM.

Quantitative evaluation: For this evaluation, we use an endoscope instead of a cystoscope because of availability issues. However, the image quality of endoscopically acquired images is lower compared to cystoscopic images. Moreover, radial distortion in endoscopic images is greater. Thus, a radial distortion correction is performed over acquired images using the algorithm described in [14]. This was the only extra step in the evaluation process.
To evaluate the quality of the constructed panoramic view, two sequences of 27 images were acquired with an endoscope using the positioning system showed in figure 1. This system allowed us to control, with sub–millimetric displacements, the translations as well as rotations and perspective movements between images. The first sequence consisted in 27 images taken from a high–quality digital photography of the internal surface of a pig bladder. The acquisition involved three directions of displacement: a translation movement (15 images) over $\vec{x}$, a diagonal translation (7 images) over $\vec{x}$ and $\vec{y}$, and a diagonal translation with 2^o rotations and 2^o perspective rotations of the endoscope between images (5 images). The same process (same positions) was repeated over a grid pattern (10 mm equally spaced black crosses) to obtain the second sequence allowing for error computation. A schematic representation of the process is shown in figures 2 and 3.
The registration and stitching algorithm was then applied to images acquired from the bladder photography

ISSN: 1727-1983

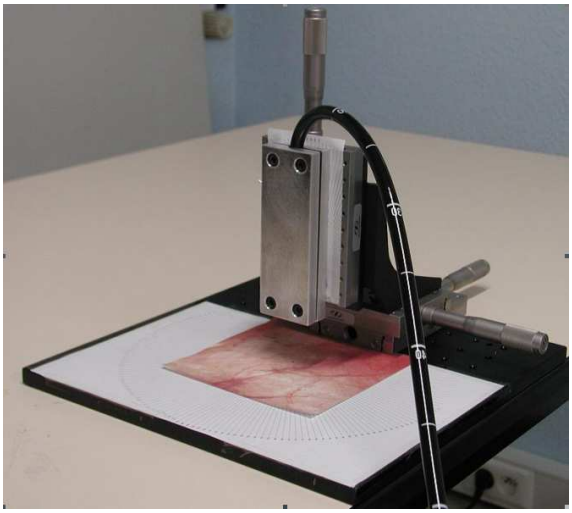

Figure 1: Positioning testing set–up used to acquire the test sequence. This device allows a millimetric control over translation movements.

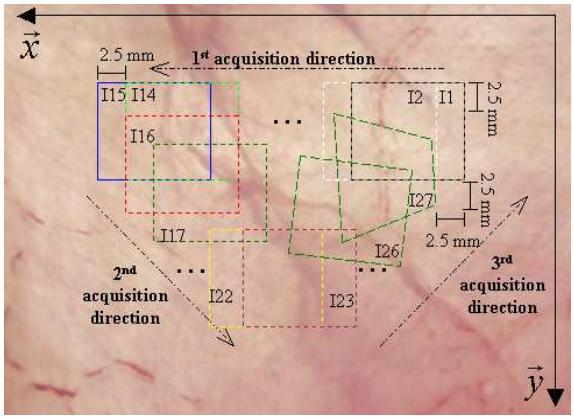

Figure 2: Schematic representation of the image acquisition path performed over a photography of pig bladder (internal surface). I1, I2,. . . , I27 are the first, second,. . . , twenty-seventh acquired image.

to find the transformations between them to construct the panoramic image. Then, these transformations are applied over the set of images acquired from the grid pattern and a second panoramic image was constructed. The first image of the grid sequence was considered as the reference image and therefore, without distortion. In order to compute mean, minimum and maximum errors in the constructed panoramic image, distances between adjacent points of this first image were measured and the mean distance computed. Then, distances between points in the panoramic image were obtained and compared against the mean distance in the first image.

Results and Discussion

Figures 4, 5 and 6 show the panoramic views constructed from 80, 162 and 193 images respectively. The algorithm is robust enough to find a good visual

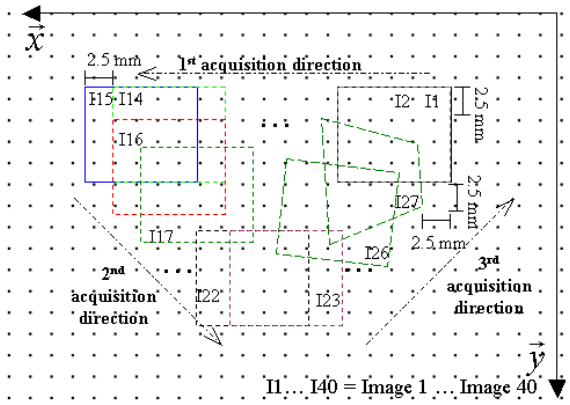

Figure 3: Image acquisition path performed over a grid test pattern for quality evaluation of the constructed panoramic view.

correspondence between images even if blurring and illumination inhomogeneities are present. After application of the weighting average function, visual artifacts between images persist, but only when strong illumination changes between images are involved. Table 1 shows computation times for each constructed panoramic view and for each pair of images processed. Computation times were measured separately for registration and stitching processes because stitching time varies with the transformation complexity and image size after transformation. The maximum displacement recovered for the sequences presented here was 96 pixels.

Table 1: Timing results in seconds for each process and for each constructed panoramic view.

Seq.	Imgs.	Regist.	Stitch.	Total	p/pair
1	80	155.62	25.85	181.47	2.29
2	162	310.23	30.70	340.93	2.11
3	193	376.52	426.54	803.06	4.18

Table 1 shows that the required computation time allows the panoramic image to be constructed in a few minutes and this computation time could be compatible with the duration of a cystoscopic exam (15–20 min).
Stitching time in sequence 3 shows a considerable increase. This is due to the complexity of the transformations obtained for the last images and their larger size after transformation.

Quantitative evaluation: Figure 7 shows the constructed panoramic image from images acquired over the bladder photography. The panoramic image constructed from images acquired over the grid pattern is shown in figure 8. It can be seen that the points from the last superposed images diverge. This is caused by errors in finding the correct transformations when large perspective and rotations are involved. Distance errors between points of the panoramic image and the mean distance of image 1

ISSN: 1727-1983

Figure 4: Panoramic view (959 x 671 pixels) from a sequence of 80 images. The vessels and the suspicious structure (bottom right) in the panoramic view are easily identifiable.

Figure 5: Panoramic view (1651 x 482 pixels) from a sequence of 162 images (vertical scan of the bladder).

Figure 6: Panoramic view (4883 x 1697 pixels) from a sequence of 193 images (vertical scan of the bladder).

are summarized in table 2.
1st MD is the mean distance between points of the first image of the sequence. **Pano MD** is the mean distance between points of the constructed panoramic image. **ME** is the mean error between distances in the panoramic image and the mean of distances in the first image. **Max** and **Min** are the maximum and minimum errors, respectively, between point distances of the panoramic image and the mean distance between points in the first image. Although a maximum error of 25.85 pixels, the visual quality of the constructed panoramic image (figure 8) is enough to authorize a good localization of structures and vessels.

Table 2: Computed errors, in pixels, for distances between points in the constructed panoramic image and the mean distance between points of the first image of the sequence.

1st MD	Pano MD	ME	Max	Min
106.16±4.33	107.03±12.04	10.14±6.49	25.85	0.056

Conclusion

Visual characteristics of the constructed panoramic views are very good and authorize an easier observation of the zone of interest by providing a wider field-of-view.

ISSN: 1727-1983

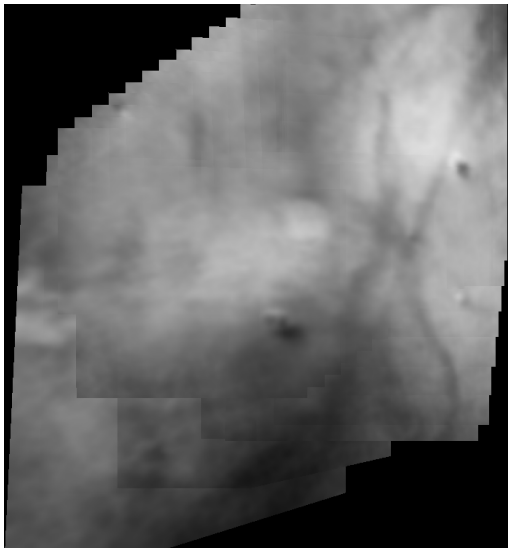

Figure 7: Panoramic view constructed from images acquired over a bladder photography using the endoscopic set–up shown in Figure 1.

Figure 8: Panoramic view of images acquired over the grid pattern.

Timing results showed the possibility of obtaining a panoramic image in a few minutes. In addition, the presented algorithm is robust enough to deal with blurring, illumination inhomogeneities and remaining radial distortion in images. When evaluating quantitatively the algorithm we have found that if a strong perspective transformation is involved in the construction of the panoramic view, error in the final panoramic image increases. But we have observed that large perspective changes between images in a cystoscopic sequence do not occur frequently. The quality of images is also important in building a good quality panoramic image. We think that the obtained errors are also a function of the lower quality used images in the evaluation process.

Acknowledgments

We thank the *Centre Alexis Vautrin (CAV)* for proportioning the cystoscopic sequences and *La Région de Lorraine* and *La ligue Régionale (CD54) de Lutte Contre le Cancer* for their financial support.

References

[1] SAWHNEY, H., KUMAR, R., GENDEL, G., BERGEN, J., DIXON, D., and PARAGANOET, V. VideoBrushTM: Experiences with consumer video mosaicing. In *4th IEEE Workshop on Applications of Computer Vision (WACV'98)*, page 56, 1998.

[2] REALVIZ S. A. *Stitcher 3.x main tutorial*. HTTP://www.realviz.com/, 2002.

[3] VASAVADA, A. R. et al. Galileo imaging of jupiters atmosphere: The great red spot, equatorialregion, and whiteovals. *Icarus*, 135:265–275, September 1998.

[4] GRACIAS, N. and SANTOS-VICTOR, J. Underwater mosaicing and trajectory reconstruction using global alignment. In *OCEANS, 2001. MTS/IEEE Conference and Exhibition*, volume 4, pages 2557–2563, 2001.

[5] HOGE, W. S., MAMATA, H., and MAIER, S. E. Efficient construction of histology slide mosaics via phase correlation registration of high resolution tiles. In *Proceedings of the IEEE International Conference on Image Processing ICIP 2003*, pages 1117–1120, September 2003.

[6] TEGOLO, D. and VALENTI, C. A naive approach to compose aerial images in a mosaic fashion. In *11th International Conference on Image Analysis and Processing (ICIAP'01)*, page 512, 2001.

[7] CAN, A., STEWART, C. V., ROYSAM, B., and TANENBAUM,H. L. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):347–364, 2002.

[8] ASMUTH, J., MADJAROV, B., SAJDA, P., and BERGER, J. W. Mosaicking and enhacement of slit lamp biomicroscopic fundus images. *British Journal of Ophtalmology*, 0(85):563–565, 2001.

[9] AIGER, D. and COHEN-OR, D. Mosaicing ultrasonic volumes for visual simulation. *IEEE Computers Graphics and Applications*, 20(2):53–61, 2000.

[10] KUGLIN, C. D. and HINES, D. C. The phase correlation image alignment method. In *Procceedings of the 1975 IEEE International Conference Cybernetics Society*, pages 163–165, 1975.

[11] REDDY, B. S. and CHATTERJI, B. N. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, August 1996.

[12] BAKER, S. and MATTHEWS, I. Lucas–Kanade 20

ISSN: 1727-1983

years on: A unifying framework: Part 1. Technical Report CMU-RI-TR-02-16, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, July 2002.

[13] LUCAS, B. D. and KANADE, T. An iterative image registration technique with an application to stereo vision (darpa). In *Proceedings of the 1981 DARPA Image Understanding Workshop*, pages 121–130, April 1981.

[14] MIRANDA-LUNA, R., BLONDEL, W. C. P. M., DAUL, CH., HERNANDEZ-MIER, Y., POSADA, R., and WOLF, D. A simplified method for endoscopic image distortion correction based on grey level registration. In *2004 International Conference on Image Processing, ICIP2004*, volume 5, pages 3383–3386, October 2004.

ISSN: 1727-1983