Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Efficient Tracking of the Heart Using Texture

Aurélien Noce, Jean Triboulet, Philippe Poignet

LIRMM - UMR 5506 - UM2 CNRS - 161 rue Ada, Montpellier, France
{noce,triboule,poignet}@lirmm.fr

***Abstract*— Performing motion tracking in real-time is an old and recurrent problem in Computer Vision. It has been addressed through a large set of approaches [17], [9], [1], but achieving a high level of robustness is still a challenge, especially with low definition input. In the considered application, tracking the heart motion in endoscopic beating heart sequences, the sensitivity of existing algorithms to visual artifacts and variations in illumination is an issue that calls for improvements. In the prospect of developing a motion compensation architecture for robotically assisted beating heart surgery, we address the problem of visual information retrieval by proposing a new *Composite Tracking Algorithm* using both template matching and texture analysis. As we will show in this paper, the use of texture characterization of the heart surface improves the overall precision and robustness, in comparison with other prior approaches.**

I. Introduction

Driven by the increase of computers power, *computer vision* has been evolving rapidly and an ever-growing number of works have appeared. In the field of robotics, *visual servoing* has shown its enormous potential, and recent advances in visual servoing schemes [3], [2] and pattern tracking techniques [4], [1] settle the bases for a new generation of vision-centered applications. But extracting and interpreting structured information from pixel data is still a very complex challenge that brings in several research areas such as signal processing, computer science and automatic control. Furthermore many open problems are still to be considered in: denoising of the image, selection of accurate regions of interest, robustness of the tracking procedure to changes in illumination… all those issues becoming critical in experimental conditions.

This work takes place in a larger project aiming at developing a robotic platform for motion compensation of the beating heart. This is a very demanding task, especially in terms of robustness and accuracy: 1mm or less precision in the tracking is required to achieve effective virtual stabilization of the heart. Several works have investigated motion compensation for surgical applications, starting with the work of Nakamura in [13] that leaded to further developments in [16], [6], [15]. Our contribution is a *Composite Tracking Algorithm* (CTA) that uses both texture characterization and pattern matching to track specific regions on the heart during motion, as the initial step to perform visual servoing.

The paper is organized as follows. In Section II, we give an overview of our goals and methodology. Sections III and IV introduce the proposed tracking algorithm. In Section V, we present tracking results on artificial and experimental images, and compare the performance with respect to other commonly used tracking methods. Finally, Section VI sums up those results.

II. Motivations

A. Prior Work

In a former paper, we have presented a method for 3D reconstruction [12] based on tracking of artificial markers located on the heart surface. Experiments with this approach pointed out several practical problems, related to the use of endoscopic equipment, the power of the light source and the complexity of placing artificial markers on the heart. We decided to conduct further investigations on imaging aspects, and more precisely on improving the region tracking procedure. Our guess was that with a sufficiently robust tracking algorithm, we might avoid the use of artificial landmarks by tracking directly a pattern on the heart surface (see Fig. 1), solving at one time the question of the positioning of those landmarks.

B. Introducing Texture Analysis

Serveral approaches can be considred to perform visual tracking: finding similarities between patterns in consecutive images [8], measuring the movement of the target from the optical flow [17], using pattern variations during motion [9] or using shape descriptors [11]. As our goal is to acheive precise tracking with as less as possible need for *a priori* knowledge (we want to be patient-independent) we focused on pattern tracking approaches. Evaluation of classical approaches such as *Sum of Squared Differences* (SSD) minimization or *Correlation* maximization, showed insufficient robustness in region tracking on our experimental set of images. To improve those results we looked for a way to extract more information from image data. Preliminary work, in [14], showed the potential of the approach.

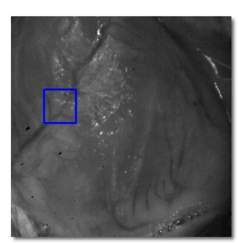

Fig. 1. Sample tracked pattern

1-4244-0788-5/07/$20.00 ©2007 IEEE

Texture characterization is already used in medical applications such as expert systems for automatic diagnosis [10]. However, the real-time constraint was not considered in those applications, while being central in out project. To limit the necessary computing time, we selected a reduced set of texture features (see [14]) recalled in Table I.

III. Theorical Background

A. Region Tracking

The problem of pattern tracking can be stated as a minimization problem. The cost function d represents the distance between images. Let I be a $n \times m$ image, x the tracked pattern and t the current frame, then the motion of the pattern between frame t and $t+1$ is given by the transformation $\mathcal{T}$ that solves:

$$\min_{\mathcal{T}} |d(I(\mathcal{T}(\boldsymbol{x}), t+1), I(\boldsymbol{x}, t)))|, \tag{1}$$

i. e. that projects the pattern in frame t to its new position in frame $t+1$.

In this paper we considered rigid transforms as the T function, translation:

$$\mathcal{U}(t) = \begin{pmatrix} u_x \\ u_y \end{pmatrix}, \tag{2}$$

and homography:

$$\mathcal{H}(t) = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}. \tag{3}$$

With those models, we make the assumption that the geometric transformation is rigid. This is a stong assumption, but the small displacements in the image, resulting from the high frequency of the camera, make this hypothesis reasonable.

Other displacement models can be also be considered, but those approaches proved their efficiency in terms of speed, accuracy, and robustness in the test sequences.

B. Computing Similarities Between Patterns

We use normalized forms of SSD and Correlation to compute template based distance d as a first step in the Composite Cost Function computation.

C. Texture Features and Texture Comparison

Texture features [7], [18] associate quantitative information such as contrast, roughness or homogeneity to a region of interest. Those values form the *texture vector* that characterizes the region. The considered vectors have 8 components, given in Table I:

$$\begin{aligned} \boldsymbol{V}(I) = &\quad (x_1, x_2, \ldots, x_8)^T \\ \boldsymbol{V}(I') = &\quad (y_1, y_2, \ldots, y_8)^T, \end{aligned}$$

and the distance between texture vectors is computed by the way of the euclidean distance d:

$$t(I, I') = \sqrt{\sum_{i=1}^{8} (x_i - y_i)^2}. \tag{4}$$

TABLE I

Considered texture features

Approach	Feature
Coocurrence Matrix	Energy
Coocurrence Matrix	Contrast
Coocurrence Matrix	Cluster Shade
Coocurrence Matrix	Cluster Prominence
Run-length Matrix	Non-Uniformity
Run-length Matrix	Short Low Grey Level Run Emphasis
First Order Statistics	Skewness
First Order Statistics	Kurtosis

Then the vector is normalized, as needed by our algorithm:

$$\hat{t}(I, I') = \frac{d(I, I')}{\max\limits_{\forall I, I'}(d(I, I'))}. \tag{5}$$

Evaluation of texture-based characterization on experimental data showed that it is more robust to changes in shape and illumination than pattern-based approaches.

IV. Composite Cost Function

In order to combine accuracy of pattern-matching and robustness of texture-based characterization, we use a *Composite Cost Function* (CCF) σ:

$$\begin{aligned} \sigma(I, I') &= \lambda.d(I, I') + \gamma.\hat{t}(I, I'), \\ \lambda + \gamma &= 1. \end{aligned} \tag{6}$$

Parameters λ and γ, giving the balance between the two approaches, can be set statically, but dynamic values proved better efficiency during our test phase. We tried several expressions for those parameters, and finally came up with two efficient (in terms of tracking robustness and computational load) solutions:

$$\gamma = d(I, I'), \tag{7}$$

$$\gamma = \exp\left(\frac{-d(I, I')}{1 - d(I, I')}\right), \tag{8}$$

$$\lambda = 1 - \gamma.$$

Solution (7) favors pattern-tracking over texture characterization, while solution (8) is more balanced.

V. Results

The CTA has been evaluated on a set of artificial sequences and on experimental beating heart sequences. Artificial sequences were generated by superimposing a textured pattern and a background then applying a known motion to this pattern, as seen on Fig. 2 where the tracked pattern is represented by a red square.

Then several algorithms were compared on their ability to track the pattern, among which:

- Correlation,
- SSD,
- Optical Flow,
- ESM [1], a recent optimisation scheme and homography-based tracking method.

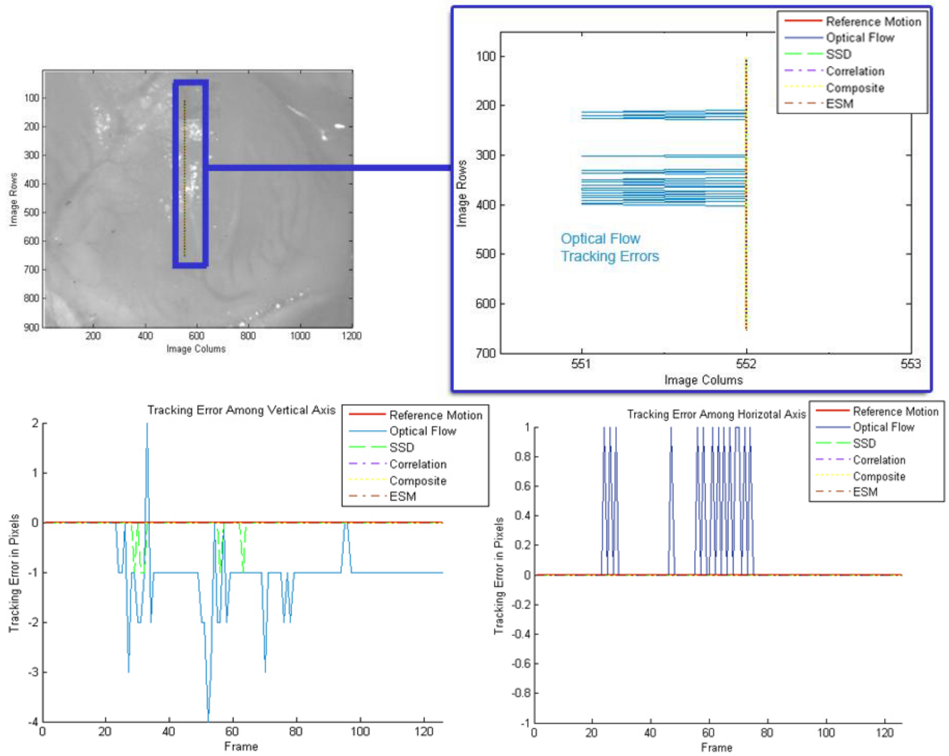

Fig. 3. Sample Trajectories for Translating Motion

Beating heart sequences were used for qualitative evaluation of the tracking. All sequences are 256 grey-levels, with resolution going from 128×128 pixels to 512×512 pixels.

A. Results on Artificial Images

1) Translating Motion: The sequence represents a translation of the pattern, as illustrated on Fig. 2. Most approaches perform very well on those sequences, both in terms of robustness: typical results are presented in Table II and Fig. 3, both concerning the same sequence. In terms of accuracy, CTA performs as well as SSD and Correlation, and better than Optical Flow. The deviation in the trajectory of the optical flow tracking can be explained by the presence of lens flares. Other methods performs very well.

2) Rotation and Translation: Adding rotation to the motion is more demanding and optical flow tracking quickly diverges after a few iterations. Fig. 4 shows an example of circular translation with small rotations of the pattern. It is a very interesting case because the pattern goes through all the problematic regions of the heart. Correlation and SSD method pain to track the motion when the pattern goes through regions of the heart presenting lower contrast or specularities. On the other hand CTA performs well on this sequence, as seen on Fig. 4.

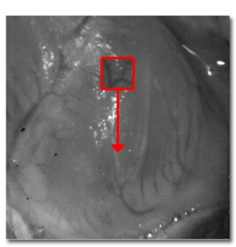

Fig. 2. Artificial Sequence

TABLE II

ERRORS IN SIMULATION: TRANSLATION TRACKING (SEQUENCE OF 100 IMAGES AT 100HZ)

Method	Mean Error in Pixels	Strandard Variation
Correlation	0.00	0.00
SSD	0.04	0.19
Optical Flow	0.90	0.73
Composite	0.00	0.00
ESM	0.00	0.00

B. Beating heart sequences

Interpreting natural sequences is more delicate because the motion is not precisely known and much more complex: deformation of the surface and changes in pose and illumination are very demanding in terms of tracking. Experimental sequences were used to compare the algorithms in desired conditions. Those sequences are adequate for qualitative evaluation of our method.

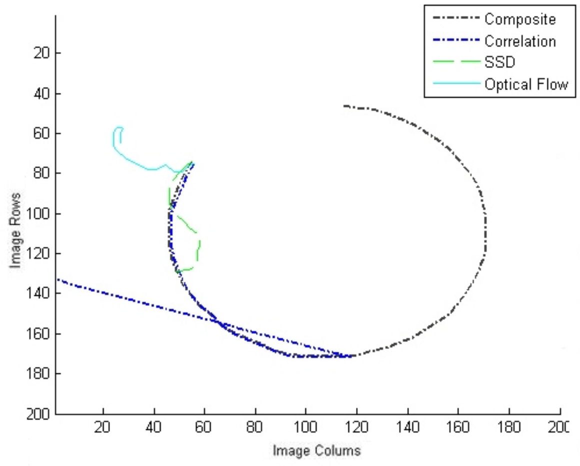

Fig. 4. Sample Trajectory for Circular both Translation and Rotation of the Pattern

As expected with beating heart images, the results of tracking with optical flow are not satisfactory. The tracking using correlation and SSD give good results, but on some specific regions their accuracy were not as good as expected. Finally, the CTA improves precision while keeping good accuracy.

VI. Conclusion and Perspectives

The use of texture characterization proved to be an interesting option for robustness improvement in medical vision-based applications, the main limitation of the approach being the real-time constraint. The CTA shows a way to use texture information while coping with time constraints effectively, with execution time of about 10ms.

The next step in our work is to perform the tracking online (not only on recorded sequences) to validate our method in experimental conditions, and to experiment with visual servoing based on this data. Another objective is to integrate temporal information in the tracking (as in [5]) using specificities of the heart movement.

References

[1] S. Benhimane and E. Malis. Real-time image-based tracking of planes using efficient second-order minimization. In *IEEE/RSJ International Conference on Intelligent Robots Systems*, volume 1, pages 943–948, September 2004.

[2] S. Benhimane and E. Malis. A unified approach to visual tracking and servoing. In *Robotic and Autonomous Systems*, volume 52, pages 39–52, July 2005.

[3] F. Chaumette and E. Malis. 2 1/2 d visual servoing: a possible solution to improve image-based and position-based visual servoings. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 630–635, San Francisco, USA, April 2000.

[4] A. I. Comport, E. Marchand, M. Pressigout, and F Chaumette. Real-time markerless tracking for augmented reality: The virtual visual servoing framework. In *IEEE Transactions on Visualization and Computer Graphics*, volume 12, pages 289–298, July 2006.

[5] R. Ginhoux, J. Gangloff, M. de Mathelin, L. Soler, M. M. Arenas Sanchez, and Jacques Marescaux. Active filtering of physiological motion in robotized surgery using predictive control. In *EEE Trans. Robotics*, 2003.

[6] R. Ginhoux, J. A. Gangloff, M. F. de Mathelin, L. Soler, M. M. A. Sanchez, and J. Marescaux. Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, pages 274–279, 2004.

[7] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. In *SMC*, volume 3, pages 610–621, November 1973.

[8] S. A. Hutchinson, G. D. Hager, and P. I. Corke. A tutorial on visual servo control. *IEEE Trans. Robotics and Automation*, 12(5):651–670, October 1996.

[9] Frederic Jurie and M. Dhome. Hyperplane approximation for template matching. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 24(7):996–1000, 2002.

[10] S. A. Karkanis, Iakovidis D. K., D. E. Maroulis, D. A. Karras, and M. Tzivras. Computer-aided tumor detection in endoscopic video using color wavelet features. In *IEEE Transactions on Information Technology in Biomedicine*, volume 7, pages 141–152, September 2003.

[11] P. Li, O. Tahri, and F. Chaumette. A shape tracking algorithm for visual servoing. In *IEEE Int. Conf. on Robotics and Automation*, pages 2858–2863, 2005.

[12] Sauvée M., Poignet P., Triboulet J., Dombre E., Malis E., and Demaria R. 3d heart motion estimation using endoscopic monocular vision system. In *MCBMS'06: IFAC Symposium on Modeling and Control in Biomedical Systems*, Spetember 2006.

[13] Yoshihiko Nakamura, Kousuke Kishi, and Hiro Kawakami. Heartbeat synchronization for robotic cardiac surgery. In *ICRA*, pages 2014–2019, 2001.

[14] A. Noce, J. Triboulet, P. Poignet, and E. Dombre. Texture features selection for visual servoing of the beating heart. In *IEEE / RAS-EMBS BIOROB'06*, number 183, Febrary 2006.

[15] T. Ortmaier, M. Groger, D.H. Boehm, V. Falk, and G. Hirzinger. Motion estimation in beating heart surgery. In *IEEE Transactions on Biomedical Engineering*, volume 52, pages 1729–1740, October 2005.

[16] T. Ortmaier, M. Groger, and G. Hirzinger. Robust motion estimation in robotic surgery on the beating heart. In *Proc. Computer Assisted Radiology and Surgery (CARS)*, pages 206–211, 2002.

[17] T. Suzuki and T. Kanade. Measurement of vehicle motion and orientation using optical flow. In *IEEE International Conference on Intelligent Transportation Systems*, pages 25–30, 1999.

[18] J. S. Weszka, C. R. Dyer, and A Rosenfield. A comparative study of texture measures for terrain classification. In SMC, editor, *Trans. Systems, Man, and Cybernetics*, volume 6, pages 269–285, 1976.