FAST TIME-SPACE TRACKING OF SMOOTHLY MOVING FINE STRUCTURES IN IMAGE SEQUENCES

D. Tschumperlé ⋆, *Y. Bentolila* †, *J. Martinot* †, *J. Fadili* ⋆

⋆ GREYC IMAGE (CNRS UMR 6072), 6 Bd Maréchal Juin, 14050 Caen Cedex, France
† IFP, 1 et 4 Rue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France

ABSTRACT

We address the problem of temporal tracking fine point-like and filamentary structures exhibiting smooth motions in image sequences. By taking these specific restrictions into account, we propose an original tracking method based on the search for integral lines in time-space structure tensor fields. The method appears to be simple and very efficient regarding computation time and tracking precision, allowing a sub-pixel accuracy in both spatial and temporal domains. We suggest a numerical implementation of the algorithm based on a modified constrained Runge-Kutta scheme. The performance and potential applicability of our algorithm is illustrated with two real cases : the tracking of internal linear features in composite materials observed with X-rays and the tracking of granular-shaped objects moving in a gas flow.

***Index Terms*—** Object tracking, Structure tensors, Integral lines.

1. INTRODUCTION

Automatic tracking of moving objects in image sequences is a very challenging task which has a wide number of applications in digital image processing. This is not surprising then that many different mathematical frameworks have been proposed so far in the literature to tackle the tracking problem, including now popularized Kalman filtering [7, 8, 14] and particle filtering [3, 6], among other techniques. Here, we focus on a specific tracking application aiming at analyzing dynamic phenomena, which can be quantified by the measurement of *smooth motions* of *fine structures* in digitally acquired multispectral image sequences. Basically, we are mainly interested in the motion of pointwise or linear structures, such as fibers. This setting is usually encountered in a wide range of physical studies, including application fields in chemistry [9], applied mechanics [16] or medical imaging [2].

Our application is *not subject to real-time* constraints, and the tracking algorithm we propose in this paper acts as a post-processing method, having full access to forward and backward frame informations. But as the analyzed image sequences are generally composed of a huge amount of frames, we particularly paid attention to the cost of the algorithm in terms of computation time.

The paper is organized as follows : The preliminary section 2 first describes a simple but efficient *spatial tracking* technique allowing to retrieve *filamentary structures* respecting some smoothness properties in a *single multispectral frame* of the sequence. Then, we show that this idea can be extended to the *spatio-temporal* domain in order to track *pointwise* objects subject to smooth motions in the whole sequence (section 3). A modified $4th$-order Runge-Kutta integration method [10] is thereby proposed to deal with the algorithm implementation (section 4). By the way, the combination of both spatial and temporal tracking methods can be used to track filamentary structures throughout an image sequence. We applied the proposed techniques on two real physical applications (section 5). The first one concerns the tracking of internal linear features in composite material observed on X-ray images in order to better understand mechanical forces in action. The second one concerns the tracking of mixed granular-shaped objects moving in a gas flow for refining purposes.

2. SPATIAL TRACKING

We consider a multispectral $2D$ image $\mathbf{I} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^n$ (typically, $n = 1$ for scalar-valued and $n = 3$ for color images), containing a filamentary object such as one fiber of interest we want to analyze. We denote by I_i the ith vector channel of the image $\mathbf{I}$. Given an initial *user-defined* seed point $(x_0, y_0)^T$ belonging to the fiber, the goal of the tracking is to retrieve the function $\mathcal{C} : \mathbb{R} \rightarrow \mathbb{R}^2$ giving the spatial coordinates $\mathcal{C}_{(s)} = (x_{(s)},\ y_{(s)})^T$ of all points composing the filamentary curve. The fiber is supposed smooth enough, and the tracking problem can be thereby formulated as finding the C^1 curve $\mathcal{C}$ such that :

$$\begin{cases} \mathcal{C}_{(0)} = (x_0, y_0)^T \\ \frac{\partial \mathcal{C}}{\partial s} = \mathbf{u}_{(\mathcal{C}(s))} \end{cases} \tag{1}$$

where s stands for the curvilinear abscissa of $\mathcal{C}$ and $\mathbf{u}_{(\mathcal{C}(s))} \in \mathrm{S}^1$ is the unit tangent vector to the fiber, defined at least on each fiber point $\mathcal{C}_{(s)}$. The fiber is said to be tracked *forward* when $s \rightarrow +\infty$ and *backward* when $s \rightarrow -\infty$. A function integration step is obviously needed to retrieve the function $\mathcal{C}$.

In order to determine the tangent vectors $\mathbf{u}_{(\mathcal{C}(s))}$, two natural assumptions are considered : First, the desired tracked object must be contrasted enough with respect to the image background. Second, we suppose that this filamentary structure is composed of pixels having roughly the same vector-valued intensity (color). For the scalar-valued case ($n = 1$), $\mathbf{u}$ should be typically related to the isophote direction $\frac{\nabla I^T}{\|\nabla I\|}$. In a more general manner, we propose to handle the multispectral case ($n \in \mathbb{N}$) and to gain robustness to noise by choosing $\mathbf{u}$ as the *minimal eigenvector* (unit vector associated with the *smallest* eigenvalue) of the so-called *smoothed structure tensor* $\mathbf{G}_{(\mathcal{C}(s))}$ located as $\mathcal{C}_{(s)}$. $\mathbf{G}$ is defined as the field of 2×2 symmetric and positive-definite matrices :

$$\mathbf{G} = \left(\sum_i \nabla I_i \nabla I_i^T \right) * G_\sigma \tag{2}$$

where $\nabla I_i = \left(\frac{\partial I_i}{\partial x} \quad \frac{\partial I_i}{\partial y} \right)^T$ is the usual gradient vector of the channel I_i, and G_σ is a normalized $2D$ Gaussian kernel. The structure tensor is a natural generalization of the gradient for multispectral images [5] and has shown to have several good properties when used to retrieve *multispectral isophote directions* [15, 17]. $\mathbf{G}$ is defined on the whole domain Ω so as $\mathbf{u}$, which can be thus seen as a vector field $\mathbf{u} : \Omega \rightarrow \mathrm{S}^1$ whose value on a fiber point $\mathcal{C}_{(s)}$ gives the tangent vectors to $\mathcal{C}$. Note that the smoothing parameter σ has a direct impact on the regularity of the eigenvector field $\mathbf{u}$, and thereby implicitely controls the regularity of the tracked fiber $\mathcal{C}$. In particular, higher smoothing may be used to get rid of fiber occlusion problems by reconstructing correct tangent vectors on occluded area.

For simplicity reasons, we assumed here that $\mathbf{u}$ is correctly *aligned*. As an eigenvector represents an orientation without notion of direction, the computed field $\mathbf{u}$ can be in fact possibly discontinuous, independantly from the value of σ. This is anyway not an issue for the function integration step since we need only the orientation information as the tracking direction is generally explicitely known (see Section 4). Also, it is interesting to note that our filamentary structure tracking algorithm shares common ideas with some existing white matter tractography methods proposed in the field of diffusion-tensor MRI medical imaging [2]. Result of our filamentary structure tracking approach is illustrated on Fig.1.

3. EXTENSION TO TEMPORAL TRACKING

Now, let us consider $\mathbf{S} : \Omega \times [0, T] \rightarrow \mathbb{R}^n$ as the multispectral image sequence containing pointwise objects in motion we want to track. $\mathbf{S}$ is a volume $W \times H \times T$ of stacked multispectral image frames. We limit ourselves to pointwise objects tracking since we are able now to detect filamentary structures starting from a single point (see section 2).

Similarly to the spatial case, the tracking problem is formulated as follows : considering an user-defined input point $\mathcal{P}_0 = (x0, y0, t0)$ in one frame $t0$ of the sequence $\mathbf{S}$, find the positions $\mathcal{P}_{(s)} = (x_{(s)}, y_{(s)}, t_{(s)})$ of this point in all frames, i.e at least when $t_{(s)} \in \mathbb{N}$. We assume that the tracked pointwise object is contrasted enough and exhibits a smooth motion throughout the sequence. Then, the trajectory $\mathcal{P} : [0, T] \rightarrow \mathbb{R}^3$ can be computed as the integration of a tangent vector field $\mathbf{w} : \Omega \times [0, T] \rightarrow \mathbb{R}^3$:

$$\begin{cases} \mathcal{P}_{(0)} = (x_0, y_0, t_0)^T \\ \frac{\partial \mathcal{P}}{\partial s} = \mathbf{w}_{(s)} \end{cases}$$

This kind of tracking can be actually seen as a spatio-temporal tractography. The main challenge here consists in the choice of the tangent vector field $\mathbf{w}$. Unfortunaly, the temporal extension of the tracking method *is not as simple* as choosing $\mathbf{w}$ to be the minimal eigenvector of the structure tensor field (2) of $\mathbf{S}$ (which is now a 3×3 matrix). Intuitively, it would correspond to compute the direction of smallest intensity variations in the image sequence viewed as a volumetric image. But nothing ensures that this $3D$ direction would have a non null temporal component. In fact, most of the smallest variations are often encountered in the same frame as the current tracked point, since objects we want to track are not perfectly pointwise but are usually composed of regions with pixels having similar intensities. This is for instance the case for fibers in Fig.1 and balls in Fig.2. As a result, the minimal eigenvector of the 3×3 structure tensor is generally not representative of the object motion only, but also of its spatial structure. Thus, mixing variations of all spatial and temporal axes (x, y, t) in the same structure tensor is definitely not a good idea.

The solution relies on the simultaneous but separated tracking of the considered point in the (x, t) and (y, t) spaces instead of considering the whole $3D$ domain (x, y, t). We propose to deduce the spatio-temporal tangent vector $\mathbf{w}$ by analyzing the 2×2 smoothed structure tensor fields $\mathbf{G}_{xt}$, $\mathbf{G}_{yt}$ both defined on the domain $\Omega \times [0, T]$ and computed as

$$\mathbf{G}_{xt} = \left(\nabla S_{i_{xt}} \nabla S_{i_{xt}}^T \right) * G_\sigma \text{ and } \mathbf{G}_{yt} = \left(\nabla S_{i_{yt}} \nabla S_{i_{yt}}^T \right) * G_\sigma$$

where $\nabla S_{i_{xt}} = \left(\frac{\partial S_i}{\partial x} \ \frac{\partial S_i}{\partial t} \right)^T$, $\nabla S_{i_{yt}} = \left(\frac{\partial S_i}{\partial y} \ \frac{\partial S_i}{\partial t} \right)^T$ are the two spatio-temporal gradients respectively in (x, t) and (y, t) spaces. For each point (x, y, t) of the sequence, the minimal eigenvectors $\mathbf{u}_{xt} = (u_{xt}, v_{xt})^T$ and $\mathbf{u}_{yt} = (u_{yt}, v_{yt})^T$ of $\mathbf{G}_{xt}$ and $\mathbf{G}_{yt}$ give coherent measures on the spatio-temporal variation directions of $\mathbf{S}$. Mixing the directions $\mathbf{u}_{xt}$ and $\mathbf{u}_{yt}$ allows to retrieve a good approximation of the point motion, and ensures that the estimated motion tangent vector $\mathbf{w}$ goes through time :

$$\mathbf{w} = \begin{cases} (u_{xt}, \frac{u_{yt}|v_{xt}|}{|v_{yt}|}, v_{xt})^T & \text{if } |v_{xt}| < |v_{yt}| \text{ and } v_{xt} \neq 0 \\ (\frac{u_{xt}|v_{yt}|}{|v_{xt}|}, u_{yt}, v_{yt})^T & \text{if } |v_{xt}| > |v_{yt}| \text{ and } v_{yt} \neq 0 \end{cases}$$

Intuitively, the tracking procedure reads as follows. For each curve point $(x_{(s)}, y_{(s)}, t_{(s)})$, we get two motion vectors $\mathbf{u}_{xt}$ and $\mathbf{u}_{yt}$ in the (x,t) and (y,t) spaces, characteristic of how important is the estimated temporal motion along x and y. The final motion vector $\mathbf{w}$ is then chosen mainly from the one that varies the less, and is extended to the $3D$ domain using the data provided by the other.
It appears that this relatively simple motion vector computation works surprisingly well in order to estimate the coherent motion appearing in the image sequence.

4. IMPLEMENTATION CONSIDERATIONS

The main issues for the numerical implementation of the tracking algorithm are twofold.

Computation of the 2×2 structure tensors : A classical way of estimating the structure tensor field consists in computing a smoothed version of $\mathbf{G} = \sum_{i=1}^{n} \nabla I_i^c \nabla I_i^{c^T}$ where $\nabla I_i^c = \frac{1}{2}(I_{i(x+1,y)} - I_{i(x-1,y)} \;\; , \;\; I_{i(x,y+1)} - I_{i(x,y-1)})^T$ is the classical central finite difference scheme of the image gradient. Here, we rather propose a slightly modified scheme which advantagely uses the more precise forward and backward finite difference schemes, while remaining symmetric :

$$\mathbf{G} = \sum_{i=1}^{n} \left(\frac{\nabla I_i^f \nabla I_i^{f^T} + \nabla I_i^f \nabla I_i^{b^T} + \nabla I_i^b \nabla I_i^{f^T} + \nabla I_i^b \nabla I_i^{b^T}}{4} \right)$$

where ∇I_i^f and ∇I_i^b correspond to the classical forward and backward differences :

$$\nabla I_i^f = \begin{pmatrix} I_{i(x+1,y)} - I_{i(x,y)} \\ I_{i(x,y+1)} - I_{i(x,y)} \end{pmatrix} \quad \text{and}$$

$$\nabla I_i^b = \begin{pmatrix} I_{i(x,y)} - I_{i(x-1,y)} \\ I_{i(x,y)} - I_{i(x,y-1)} \end{pmatrix}$$

The same kind of formulae is used to compute the spatio-temporal structure tensors $\mathbf{G}_{xt}$ and $\mathbf{G}_{yt}$.

Curve integration : The problem consists in numerically integrating the equation (1). A classical Euler scheme would consist in iterating :

$$\mathcal{C}_{(s+ds)} = \mathcal{C}_{(s)} + ds \; \mathbf{w}_{(\mathcal{C}(s))}$$

with ds, an user-defined integration step close to zero. A more precise integration algorithm has been obtained here, using the so-called $4th$-order Runge Kutta scheme [10]. It consists of introducing several interpolations of the motion vector $\mathbf{w}$ for one iteration :

$$\mathbf{k}_1 = ds \; \mathbf{u}(\mathcal{C}_{(s)}), \qquad \mathbf{k}_2 = ds \; \mathbf{u}(\mathcal{C}_{(s)} + \tfrac{\mathbf{k}_1}{2}),$$

$$\mathbf{k}_3 = ds \; \mathbf{u}(\mathcal{C}_{(s)} + \tfrac{\mathbf{k}_2}{2}), \qquad \mathbf{k}_4 = ds \; \mathbf{u}(\mathcal{C}_{(s)} + \tfrac{\mathbf{k}_3}{2})$$

Then, the next point $\mathcal{C}_{(s+ds)}$ is obtained by

$$\mathcal{C}_{(s+ds)} = \mathcal{C}_{(s)} + \tfrac{\mathbf{k}_1}{6} + \tfrac{\mathbf{k}_2}{3} + \tfrac{\mathbf{k}_3}{3} + \tfrac{\mathbf{k}_4}{6}$$

Linear interpolation has been used successfully to estimate the value of the vector $\mathbf{w}$ on inter-pixel coordinates.
As noticed in section 2, only the orientation of $\mathbf{w}$ is significant, not its direction. To get rid of this direction, we introduce a realignement step before each tracking iteration, such that the estimated vectors $\mathbf{w}$ in the 3×3 neighborhood of the current point $\mathcal{C}_{(ds)}$ are aligned with the one used for the previous iteration. It ensures that the tracking is done in a coherent direction without half-turns.

5. APPLICATIONS AND RESULTS

Some results aof our proposed spatial and temporal tracking algorithm is illustrated with two different applications, one for mechanical purposes with a sequence containing fibers, and the other for refining purposes, with a sequence of moving pointwise granular-shaped objects.

Fiber Tracking : In [16], we already described a complete image processing pipeline in order to separate and display internal structures of a composite material submitted to external stresses. Fibers composing this material are moving and must be tracked to understand their mechanical behavior. Image sequences of these moving fibers are acquired thanks to a specific X-ray video camera, then filtered using techniques presented in [16], including a denoising step [15], a wavelet-based image contrast enhancement step [12], followed by a Morphological Component Analysis (MCA) [13], enabling to distinguish between fiber components of the image. The spatial and temporal tracking algorithms described in section 2 and 3 have been succesfully applied to extract the geometry of the fiber (Fig.1).

Granular-shape tracking : This example illustrates a refinery process, which consists in injecting a gas across a catalyst moving bed (complete description of the experiment is detailed in [9]). It involves a complex medium : a mix of granular material with a flowing gas inside. We are interested in a phenomena appearing inside the refinery reactor when a horizontal flow goes through a moving granular bed. Thanks to image analysis applied to experimental equipment, a criterion is used to characterize the cavitation phenomena; it is based upon the kinematics of the grains. Here, our algorithm enables the tracking of a granular-shape object through a gas flow. The results are illustrated in the image sequence in Fig.2.

6. CONCLUSION & PERSPECTIVES

We proposed an original spatio-temporal tracking algorithm that is very simple to implement, does not need expensive

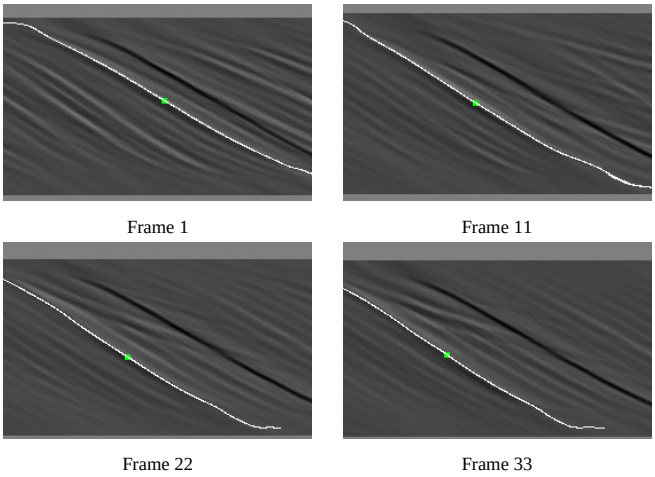

Fig. 1. Tracking of one fiber in a X-Ray sequence.

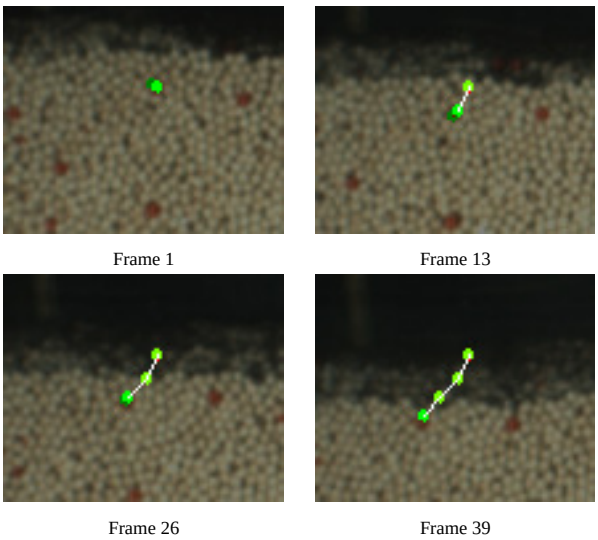

Fig. 2. Tracking of one granular-shaped ball in a color sequence.

computation time and gives tracking results with a sub-pixel precision. Detailled numerical schemes have been proposed in order to get rid of the most common implementation issues. We illustrated its interest with two various applications in the fields of chemistry and applied mechanics.

7. REFERENCES

[1] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: PDE's and the Calculus of Variations*, vol. 147 of *Applied Mathematical Sciences*.Springer-Verlag, January 2002.

[2] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda and A. Aldroubi *In Vivo Fiber Tractography Using DT-MRI Data*, Journal of Magnetic Resonance in Medicine, Vol. 44, pp.625–632, 2000.

[3] A. Blake and M. Isard. *Active Contours: The Application of Techniques from Graphics,Vision,Control Theory and Statistics to Visual Tracking of Shapes in Motion* Springer-Verlag Ed.,ISBN:3540762175, 1998.

[4] B. Cabral, L.C. Leedom. *Imaging vector fields using line integral convolution.* Computer Graphics Proceedings, p.263–270, 1993.

[5] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, 1986.

[6] N. Gordon, D. Salmond and A. Smith. *Novel approach to nonlinear/non-gaussian Bayesian state estimation.* IEE Proc. F, Radar and Signal Processing, 140(2), pp.107–113, 1993.

[7] J. Jackson, A. Yezzi and S. Soatto. *Tracking deformable moving objects under severe occlusions.* IEEE Conference on Decision and Control, 2004.

[8] N. Peterfreund. *Robust tracking of position and velocity with Kalman snakes.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 21, No 6, pp 564–569, 1999.

[9] F. Pradel, C. Lanaud and Y. Meimon. *Interaction between granular and gas flows: the pinning effect.* ed. Y. Kishino, Powders and Grains 2001, 585-588, Balkema Publishers, 2001.

[10] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *"Runge-Kutta Method"* in Numerical Recipes in FORTRAN: The Art of Scientific Computing, Cambridge University Press, pp. 704-716, 1992.

[11] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis.* Cambridge University Press, 2001.

[12] J.-L. Starck, F. Murtagh, E. Candes, and D. Donoho. *Gray and Color Image Contrast Enhancement by the Curvelet Transform.* IEEE Trans. Im. Proc., vol. 12, No 6, pp. 706-717, June 2003.

[13] J.-L. Starck, M. Elad, and D. Donoho. *Image decomposition via the combination of sparse representations and variational approach.* IEEE Trans. Im. Proc., vol. 14, No 10, pp. 1351-1382, October 2005.

[14] D. Terzopoulos and R. Szeliski. *Active Vision, Tracking with Kalman Snakes.* MIT Press, pp. 3–20, 1992.

[15] D. Tschumperlé. *Fast Anisotropic Smoothing of Multi-Valued Images using Curvature-Preserving PDE's.* IJCV (International Journal of Computer Vision), 2006.

[16] D. Tschumperlé, J. Fadili and Y. Bentolila. *Wire Structure Pattern Extraction and Tracking from X-Ray Images for Riser Mechanisms.* CVPR (Computer Vision & Pattern Recognition), IEEE Conference, New-York, June 2006.

[17] J. Weickert. *Anisotropic Diffusion in Image Processing.* Teubner-Verlag, Stuttgart, 1998.

[18] A. Yezzi and S. Soatto. *Deformotion : Deforming motion, shape average and the joint registration and approximation of structures in images.* International Journal of Computer Vision, Vol 53, No 2, pp 153–167, 2003.