LOCAL AFFINE TEXTURE TRACKING FOR SERIAL REGISTRATION OF ZEBRAFISH IMAGES

Abhir Bhalerao[1] *Luke Pase*[2]
Graham J. Lieschke[2] *Stephen A. Renshaw*[3] *Constantino Carlos Reyes-Aldasoro*[4]

[1] Computer Science, University of Warwick, UK
[2] Australian Regenerative Medicine Institute, Monash University, Australia
[3] MRC Centre for Developmental and Biomedical Genetics, and Department of Infection and Immunity, University of Sheffield, UK
[4] Biomedical Engineering Research Group, University of Sussex, UK

ABSTRACT

The aim of this work is to register serial in-vivo confocal microscopy images of zebrafish to enable accurate cell tracking on corresponding fluorescence images. The following problem arises during acquisition; the zebrafish tail may undergoe a series of movement and non-linear deformations, which if not corrected, adds to the motion of leukocytes being tracked. This makes it difficult to accurately assess their motion. We developed a correlation based, local affine image matching method, which is well suited to the textured DIC images of the anatomy of the zebrafish and enables accurate and efficient tracking of image regions over successive frames. Experimental results of the serial registration and tracking demonstrate its accuracy in estimating local affine motions in zebrafish sequences.

1. INTRODUCTION

Inflammation is critical to life itself, as one of the first reactions of the immune system of a multicellular organism to protect it against harmful stimuli. Neutrophils and macrophages are key cells of the immune system and the rapid arrival of neutrophils to a site of injury or infection is critical for host defence. Zebrafish larvae have emerged in recent years as a key organism for modelling immune responses, with a unique combination of advantages over other systems for the detailed study of biological processes such as inflamation. In particular, their optical transparency allows visualisation of inflammation processes *in vivo*. Genetic manipulations can be easily performed, both to manipulate the inflammatory response, but also to label individual cell populations with fluorescent markers [1, 2]. The combination of near transparency and genetic manipulability allows these cell populations to be observed in high temporal and spatial resolutions, during inflammation resolution, using multiphoton and confocal microscopy.

Imaging of inflammation in zebrafish following tailfin wounding allows the individual cells to be identified and tracked over time. Larvae are anaesthetised and immobilised in agarose, and in many cases they remain static during imaging. However, in some circumstances particularly over prolonged imaging periods, the fish tail can undergo changes in position and shape, not only due to movement of the immobilised sample, but due to deformation associated with the inflammatory process (e.g. figure 3).

When cells are tracked while the tail is undergoing these deformations, the apparent displacement can be due to the movement of the tail itself and not entirely due to motion of the cells [3]. A process of image registration prior to tracking is thus required to compensate for the non-linear deformation of the tail.

Serial image registration is an important topic in biomedical imaging [4, 5] and has been widely studied because it has numerous applications, such as image guided surgery [6]. Image registration methods consist of a number of standard steps: finding features to correspond; choosing a feature similarity metric; estimating the transformations between feature points or groups of feature points.

In this paper, we present a serial image registration and texture tracking method that uses local affine correlation and is particularly suited to textural imagery, such as the anatomy of the zebrafish as shown by differential interference contrast (DIC) microscopy images. The method uses a Fourier domain parameterisation of the local affine matching problem, which enables us to sharpen the image data to emphasise textural content, plus allows us to decouple the displacement from the local linear motion. We detail a new prior weighted, min/max approach that overcomes the need for numerical derivatives and leads to stable and smooth motion estimates. The approach is shown to make the resulting method fast and subpixel accurate. Results of tracking are illustrated on synthetic and real zebrafish imagery, over hundreds of frames, to correct for cell-migration estimates taken from corresponding. We also present a novel way to combine the local affine regions transformation to effect accurate frame-to-frame alignments.

2. LOCAL AFFINE TRACKING/REGISTRATION

Registration can be divided into *rigid*, when there is no change in shape, i.e. translation and rotations, *affine* or *linear* when there is also extension, compression or shear, and *non-linear*, where the changes cannot be characterised as linear or affine. In the case of non-linear deformation, unless the transformation is constrained, the problem becomes underdetermined. The principal difference between tracking

and registration is that the latter requires an accurate correspondence at the pixel level, where in many computer vision problems, for tracking it is often sufficient to only align a relatively sparse set of feature points.

One way to tackle non-linear or deformation registration is to approximate local deformations as being rigid (or affine), and blend a set of local region transformation to estimate a global non-linear deformation. Local affine region matching for tracking in video data was proposed by Kruger and Calway [7]. The frames were divided into square regions (blocks), and these are corresponded across successive frames to estimate object motions. In [7], they used a hierarchy of blocks and large-scale motions to constrain and estimate small scale motions. Likar and Pernus similarly use hierarchical local affine block matching with thin-plate spline interpolation to effect elastic registration of images [8].

2.1. Local Region Matching: DFT Cross-Correlation and Simplex Optimization

The local registration/tracking problem is posed as a cost minimization under affine transformation of coordinates. Costs are taken as sum of squared image difference, and localised by a window, W of size B, on the image. Each point x and associated image pixels around x, $I_w(x)$, s.t. $I_w(x) = W(x)I(x)$, is compared with a corresponding region in the next frame, J. Then, a local affine deformation assumes that $J_w(x) = I_w(Ax + t)$. We can determine the displacement t by cross-correlation:

$$\arg\max_{\hat{t}} R_{IJ}(p) = \sum_{W} I_w(x+p)J_w(x), \qquad (1)$$

Furthermore, the peak of the cross correlation, $\max R_{IJ}$, is used as measure of region similarity.

We write the linear transformation A as a product of a scaling, rotation and shear matrices, $A(\Phi) : \Phi = \{s, \theta, h_x, h_y\}$

$$A(\Phi) = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} 1 & h_x \\ h_y & 1 \end{pmatrix} \begin{pmatrix} cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \qquad (2)$$

To estimate Φ, we use a non-linear optimisation by the Nelder-Mead Simplex method. This is similar to a gradient-descent on the cost surface, but has the advantage, as with other min-max approaches, that when the range of the parameters is known then it does not require derivatives of the optimization function with respect to them to be calculated. Numerical derivatives of the cost-function can become unstable. The algorithms requires initialisation with $n + 1$, n-dimensional input points on the cost surface, which in this case is 5 sets of 4-parameters, Φ_0. In Likar et al. [8], they regularise the displacement field across blocks and the global deformation is smoothed by a thin-plate spline. As we are tracking locally, the output of the matching is not a displacement field and thus we cannot impose a global smoothing. Instead, we constrain the local transformation by a prior weighting on the transformation parameters:

$$p(\Phi) = \exp\left[-\frac{(1-s)^2}{2\sigma_s^2} - \frac{\theta^2}{2\sigma_\theta^2} - \frac{h_x^2}{2\sigma_{h_x}^2} - \frac{h_y^2}{2\sigma_{h_y}^2}\right], \qquad (3)$$

where the standard deviations, $\{\sigma_s, \sigma_\theta, \sigma_{h_x}, \sigma_{h_y}\}$, are chosen to be min/max range of the expected local deformation. The algorithm proceeds as follows:

1. Take a window of size B around the source point, x, from the source image I.
2. Take a corresponding image window around x on a subsequent frame, image J.
3. Run Simplex with Φ_0 and estimate $\widehat{\Phi}$ - the objective function is evaluated at each $\widehat{\Phi}$ by first transforming $J_w(x)$ by $A(\Phi)$ correlating with $I_w(x)$ and letting the cost be $-R_{IJ}(\hat{t})p(\Phi)$
4. Simplex is run until the change in objective function falls below some tolerance, ϵ.

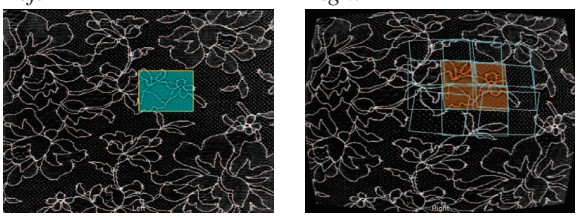


Fig. 1. Example illustrating local affine region matching. The target image (labelled Right) has been synthetically warped with a barrel deformation. The local affine transformation for a single region and its overlapping neighbours, size 128×128 pixels is shown. The pixels in blue on the Left image are taken from corresponding pixels in the target frame (Right), and those in orange on the Right image from the source.

2.2. Blending Local Transformations: Frame-to-Frame Registration

Frame-to-frame registration can be achieved by combining a set of local region transforms (as estimated above). In Likar et al. [8] and in others, this is typically achieved by a smoothing functional, such as a spline deformation [9]. Here we blend the locally transformed pixels by using interpolation windows, rather than blending the transformation parameters to estimate a dense displacement field because it is computationally simpler and achieves very good results if the regions are made to overlap by 50%. To register pixels of frame $f^{'}$ to frame f, we estimate local affine transformations $T(x_{ij})$ on a regular grid, x_{ij}, sampled at steps $B/2$ in vertical and horizontal directions. Then inverting the transformations, we bilinearly interpolate regions from image $I_{f'}$ to I_f. Each pixel in registered frame $\hat{I}(f)$ combines 4 overlapping transformed regions from I_{f+1} (figure 1) using cosine squared weights as follows

$$\hat{I}_f(x_{ij}) = \sum_{p=0,q=0}^{1,1} WC_{i+p,j+q} I_{f'}(T^{-}1(x_{i+p,j+q} - t_{i+p,j+p})), \qquad (4)$$

where the blending values in any block of size B is calculated as

$$WC(r,s) = \cos\left[(r - B/2)\frac{\pi}{B}\right]\cos\left[(s - B/2)\frac{\pi}{B}\right]. \quad (5)$$

3. ZEBRAFISH CELL MOTION COMPENSATION

The acquired zebrafish microscopy data consists of multiple channels: one contrast channel in which the zebrafish anatomy can be seen (this presents a grey scale image); one or more fluorescence channels that show only the cells of interest and can be separately tracked. Because of the non-linear motion of the fish, the cell motions contain unwanted displacements. Given an observed cell track,$\{y(f)\}$ over some range of frames $f_0 \leq f \leq f_1$, we can compensate the motions by an additive motion model, giving the recurrence relation:

$$\begin{aligned} y(f+1) &= y(f) + dx_f(f) + dx_c(f) \quad (6) \\ &= \sum_{f'=1}^{f} dx_f(f') + \sum_{f'=f_0}^{f} dx_c(f'),\ f \leq f_1 \end{aligned}$$

where the true cell motion from frame f to $f+1$, $dx_c(f)$, is biased by the $dx_f(f)$ at the pixel location $x(f)$. Then, an estimate of the true cell location on frame f, denoted by $x(f)$, is given simply by subtracting the sum of motion estimates up to f, i.e.

$$x(f) = y(f+1) - \sum_{f'=1}^{f} dx_f(f'). \quad (7)$$

4. RESULTS AND DISCUSSION

Images of anaesthetized zebrafish larvae and neutrophils (Tg(mpx:EGFP)^{i114}) were taken with a scanning confocal microscope (LSM 5 Live, Zeiss, Germany) using a Plan-Apochromat 20x/0.8NA objective with 0.5X Zoom and wide aperture (pinhole - 100 μm).

The images were high-pass filtered to emphasize textures. The cross-correlation was implemented in the Fourier domain for speed. A Gaussian window of the form $W(p) = \exp(0.5||x-p||/\sigma^2)$, where $\sigma = 2B/10$, was chosen. The window size was selected so that the maximum expected motion was less: $B > t_{max} = 64$. The standard deviations of the prior on the transformation parameters, $p(\Phi)$ are set to be $\{0.1, \pi/8, 0.1, 0.1\}$. The tolerance for the Simplex, $\epsilon = 1e-3$, was set to give a precision of 2 or more decimal places to the transformation parameters, $\widehat{\Phi}$. Parameters were initialised by choosing a starting simplex with points at the expected likely maxima: $\Phi_0 = \{1 \pm 0.3, \pm\pi/4, \pm 0.1, \pm 0.1\}$. Multiple simulations showed that the initial values were not critical (results not shown).

Figure 2 shows the result of texture tracking a region of size 32 over an example zebrafish sequence consisting of 180 frames. The image error is plotted for taking corresponding regions from successive frames with and without the tracking estimate (shown in the middle plot). Note that the transformation parameters show little scale change and small amounts of rotation, frame to frame. The motion compensated image differences remain small and the residual reflects the block nature of the registration as only a local affine of the 32 pixel region is modelled. To show that the method can be used to register the frames, we took two frames (40 and 45) in the part of the sequence where motion is significant. Figure 3 shows the alignments results using overlapping 64 pixel square regions. Frame 45 is registered to frame 40 and the edges are overlaid onto frame 40 to show the accuracy. Overall, the results are good, given that relatively large regions are used and that 5 time points separate the two frames.

Finally, we used the local affine region tracking to compensate for separately estimated cell tracks (using the method in [3]). Figure 4 shows the track positions before and after motion compensation. For instance, track 2, exhibits a bias of over ± 6 pixel. The frame-to-frame motion estimates are estimated at the track position only, but as shown, the results could be overlaid onto any frame of the sequence, or if desired, the contrast images could be warped back to a representative time point. So although the cell tracking does not require linear transform parameters to be known, they would be needed for any pixel-wise registration. Also, it should be noted that translation results are dependent on linear parameters.

5. CONCLUSIONS

A method for accurately tracking regions in serial microscopy images of zebrafish larvae has been presented. It uses a direct, correlation based approach to estimate local affine motions of a windowed image region and exploits the textural distinctiveness of the regions to achieve a robust estimate. Simplex multidimensional optimization with a Fourier cross-correlation is effective for this. We demonstrated its use to track regions in *in vivo* images containing small but abrupt motions that bias cell tracking results. This registration process forms the first step of a larger analysis software: the neutrophils have to be segmented in 3D, their displacements tracked together with the interactions between cells. These are then further analysed to draw biological conclusions depending on the conditions of the experiments: for example, different zebrafish, genetical and pharmacological manipulations, and different initiating stimuli. However, if the initial deformation is not taken into account, the measurements may be biased. The local affine region matcher may find use in other applications where deformable registration is necessary.

6. REFERENCES

[1] S. A. Renshaw, C. A. Loynes, D. M. Trushell, S. Elworthy, P. W. Ingham, and M. K. Whyte, "A transgenic zebrafish model of neutrophilic inflammation," *Blood*, vol. 108, no. 13, pp. 3976–8, 2006.

[2] C. Gray, C. A. Loynes, M. K. Whyte, D. C. Crossman, S. A. Renshaw, and T. J. Chico, "Simultaneous intravital imaging of macrophage and neutrophil behaviour during inflammation using a novel transgenic zebrafish," *Thromb Haemost*, vol. 105, no. 5, pp. 811–9, 2011.

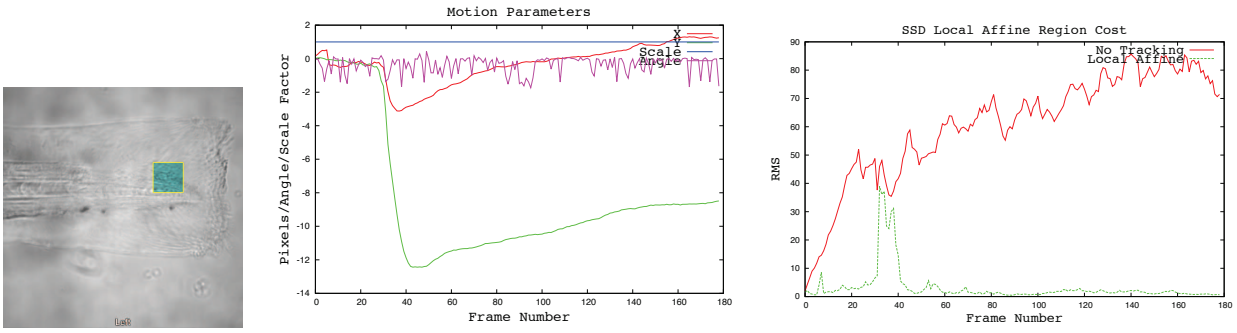


Fig. 2. Analysis of frame-to-frame local affine texture tracking errors. The region marked on the left image (on Frame 1), was tracked over the entire sequence of 180 frames. The middle plot shows the transformation parameters: cumulative shift (x - red, y - green), differential scale (blue) and angle (magenta). The right plot shows the mean squared error between the first frame and subsequent frames with and without compensating for the block motion. The registration error (green line) remains small and fairly constant other than for a small period between frames 35 and 40 where a violent motion is non-linear within the tracked region (of size 32). The error without motion compensation (red line) continues to grow.

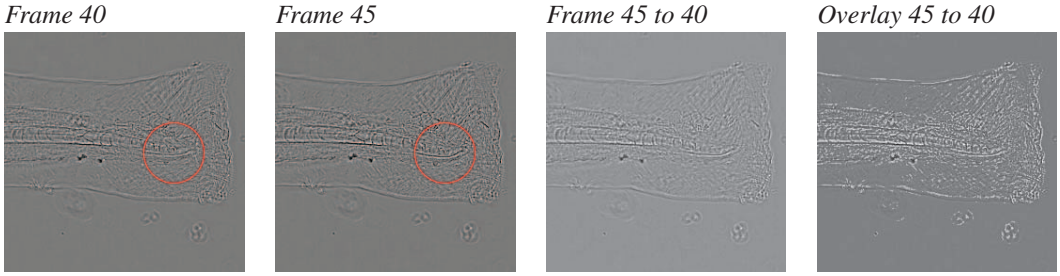


Fig. 3. Example of local affine registration applied to two frames (40 and 45) from zebrafish sequence. The local affine block matcher is applied block-by-block ($B = 64$) across the entire image, moving and interpolating frame 45 blocks onto coordinates of frame 40.

[3] C.C. Reyes-Aldasoro, Y. Zhao, and et al., "Analysis of immune cell function using in vivo cell shape analysis and tracking," in *4th IAPR Intl. Conf. on Pattern Recognition in Bioinformatics*, V. Kadirkamanathan, Ed., Sheffield, UK, 2009.

[4] D.L.G. Hill, P.G. Batchelor, M. Holden, and D.J. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, no. 3, pp. R1–R45, 2001.

[5] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[6] M. Ferrant, A. Nabavi, B. Macq, P. McL. Black, F. A. Jolesz, R. Kikinis, and S. K. Warfield, "Serial registration of intraoperative mr images of the brain," *Medical Image Analysis*, vol. 6, pp. 337–359, 2002.

[7] S.A. Krüger and A. D. Calway, "A multiresolution frequency domain method for estimating affine motion parameters," in *In Proceedings of the IEEE International Conference on Image Processing*, 1996, pp. 113–116.

[8] B. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," *Image Vision Comput.*, vol. 19, no. 1-2, pp. 33–44, 2001.

[9] J. A. Schnabel, D. Rueckert, and et al., "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Proc. of the 4th Int. Conf. on MICCAI*, London, UK, UK, 2001, MICCAI '01, pp. 573–581, Springer-Verlag.

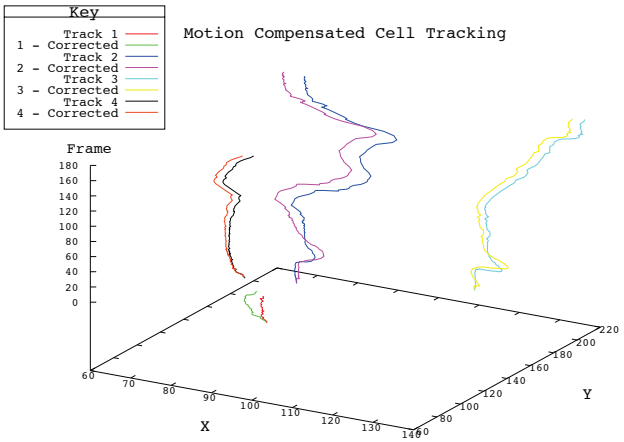


Fig. 4. Motion compensated tracking results. 3D plot showing pre- and post- motion compensated cell tracks 1-4 (z-axis is the frame sequence number).