3D Flow Field Estimation and Assessment for Live Cell Fluorescence Microscopy

Sandeep Manandhar, Patrick Bouthemy, Erik Welf, Gaudenz Danuser,
Philippe Roudot, Charles Kervrann

To cite this version:

Sandeep Manandhar, Patrick Bouthemy, Erik Welf, Gaudenz Danuser, Philippe Roudot, et al.. 3D Flow Field Estimation and Assessment for Live Cell Fluorescence Microscopy. Bioinformatics, 2020, 36 (5), pp.1317-1325. 10.1093/bioinformatics/btz780 . hal-02308001

HAL Id: hal-02308001
https://hal.science/hal-02308001v1

Submitted on 8 Oct 2019

Abstract

Motivation: The revolution in light sheet microscopy enables the concurrent observation of thousands of dynamic processes, from single molecules to cellular organelles, with high spatiotemporal resolution. However, challenges in the interpretation of multidimensional data requires the fully automatic measurement of those motions to link local processes to cellular functions. This includes the design and the implementation of image processing pipelines able to deal with diverse motion types, and 3D visualization tools adapted to the human visual system.

Results: Here, we describe a new method for 3D motion estimation that addresses the aforementioned issues. We integrate 3D matching and variational approach to handle a diverse range of motion without any prior on the shape of moving objects. We compare different similarity measures to cope with intensity ambiguities and demonstrate the effectiveness of the Census signature for both stages. Additionally, we present two intuitive visualization approaches to adapt complex 3D measures into an interpretable 2D view, and a novel way to assess the quality of flow estimates in absence of ground truth.

Availability of software: http://serpico.rennes.inria.fr/doku.php?id=data:index

Contact: sandeep.manandhar@inria.fr

3D Flow Field Estimation and Assessment for Live Cell Fluorescence Microscopy

Sandeep Manandhar[1], Patrick Bouthemy[1], Erik Welf[2], Gaudenz Danuser[2],
Philippe Roudot[2], and Charles Kervrann[1]

[1]Inria, Centre Rennes-Bretagne Atlantique, 35042, France
[2]Lyda Hill Department of Bioinformatics, UTSouthwestern, Dallas, Texas, 75390, USA

October 8, 2019

1 Introduction

Motion estimation of cellular and sub-cellular particles is crucial for the understanding of cellular functions. Intracellular dynamics can be complex and exhibits a diverse range of motions. For example, retrograde flow of actin is a slow process to extend lamellopodia, whereas formation of filopodia and subsequent cell migration can show faster and larger motions. In three-dimension (3D), light-sheet microscopy now enables the observation of even more complex behaviours, such as the interaction of a cell in an extracellular matrix (ECM) that provides mechanical support for cells and influences signaling pathways (Davidenko *et al.*, 2016). Those large and heterogenous 3D datasets motivate the design of a 3D motion estimation method to measure a large range of motion of intracellular structures, from molecular motions to cell-wide interactions.

There has been a plethora of works for motion estimation from image sequences in 2D setting. The pioneering work of Horn and Schunck (1981) led to the development of different variational methods for 2D optical flow in natural images (Fortun *et al.*, 2015; Sun *et al.*, 2010). Most of these methods work in coarse-to-fine approach where motion is first estimated with image pair of coarser resolution and warped in subsequent finer resolutions for further refinement. Brox *et al.* (2004) provided a good theoretical explanation of warping. Later Brox and Malik (2011) integrated feature matching into variational energy to deal with large displacements. Revaud *et al.* (2015) proposed a scheme of matching and interpolation as initialization for variational method, which produced good performance in different 2D optical flow benchmarks.The matching is based on a deep predefined convolutional architecture but without learning weights. It correlates patches at multiple image scales. Each patch is broken down into a hierarchy of sub-patches, locally matched up to a certain range of motion, and propagated up the hierarchy via maxpooling and aggregation. Barnes *et al.* (2009) developed a fast patch matching algorithm based on propagation of good matches and random sampling. These methods further inspired other 2D patch-based optical flow estimation methods. Hu *et al.* (2016) defined a coarse-to-fine PatchMatch to obtain semi-dense matches. The method used SIFT features extracted from a coarse grid in each level of image pyramid for matching. In a subsequent paper, Hu *et al.* (2017) proposed super-pixel based sampling for matching and a robust interpolation technique. SIFT and its variants are robust to illumination and rotation invariant to some extent. However, they are computationally demanding. Vogel *et al.* (2013), in their comparative study in optical flow setting, show the robustness of a data term arising from binary descriptor, namely, Census transform. This data term relates to local binary pattern of image gradient distribution. Recently, convolutional neural network (CNN)-based methods have been producing state-of-the-art results in 2D optical flow benchmarks. Ilg *et al.* (2017) used existing 2D dataset from optical flow benchmark as well as synthesized artificial ones for the training purpose. Hui *et al.* (2018) proposed a pyramidal based approach for CNN based feature matching, which they claim is 30 times smaller and 1.36 times faster than the network in (Ilg *et al.*, 2017). Meister *et al.* (2018) proposed an end-to-end unsupervised learning for optical flow estimation but at the cost of reduced accuracy. They use Census-loss instead of brightness constancy assumption as the data term in their loss function. Even though, the CNN-based methods are best performing methods in 2D optical flow, replicating them in 3D setting for fluorescence microscopy image is hampered by the requirements of large GPU memory and large labeled dataset for the training stage.

2D optical flow methods have been reported for fluorescence microscopy in Hubený *et al.* (2007); Delpiano *et al.* (2012); Fortun *et al.* (2013); Liu *et al.* (2014). In 3D setting, Tektonidis and Rohr (2017) used local optical flow to estimate diffeomorphic transformations for non-rigid registration of cell nuclei. Chan and Liebling (2015) measure 3D blood flow by reconstructing 3D divergence-free flow map from multiple 2D flow estimates computed from different views of 2D image pairs. Amat *et al.* (2013) proposed an algorithm involving brightness constancy and smoothness constraint imposed by Markov Random Fields over the graph of presegmented super-voxels in the 3D volume. The method computes a translation vector per super-voxel by assuming that objects do not undergo complex motion. The resulting flow field is sparse which was densified for the purpose of visualization in (Kappe *et al.*, 2016). Boquet-Pujadas *et al.* (2017) took a data assimilation approach. They combine the brightness constancy assumption with an incompressible fluid model of cell to compute various biophysical quantities in cell. Both of these methods build upon the classical constraint of brightness constancy which is in fact a strong assumption in microscopy images with Poisson-Gaussian noise. The former method, though fast, requires a scene containing isolated objects that can be meaningfully segmented into super-voxels and estimates only local translation motion. This limits the prospects of estimating complex motion in cells. The latter method requires a biophysical model of the cell. This also limits the prospects of estimating motion in fluorescence microscopy image sequences. Further, we found that none of the methods present convenient ways for visualization of 3D flow fields. Both methods use, arrows or glpyh-based methods to display sampled flow field which makes it difficult to assess the results in terms of smoothness and discontinuity.

In this work, we present a generic 3D optical flow method that neither requires a model of the cell nor makes any assumption on the motion and distribution of the objects in the volume. We propose a 3D matching framework for handling large displacements and a variational framework for handling small displacements. Both of the frameworks are then combined to handle various range of motion. We start by matching voxels between two volumes. For this, we design a 3D extension of the PatchMatch framework which allows us to deal with large displacements between 3D images. This flow field is used to back warp the target volume resulting in a reconstruction of a volume which is closer to the source volume. Then, we estimate the residual displacements between the source and reconstructed volumes by devising a 3D variational stage. The variational stage can be used independently to handle small displacements. Furthermore, the choice of the similarity measure in the matching stage or in the data term of the variational stage is an issue, especially for biological images. Therefore, we have tested three different similarity measures and chosen the Census signature to tailor both the matching and the variational stages. Further, we propose two convenient ways to visualize 3D flow fields. This is an important step as qualitative evaluation, comparison and communication becomes crucial in lack of 3D optical flow ground truth for real data. Our visualization methods emphasizes smoothness and structural coherence of the flow field which are not obviously rendered with arrow-based plots. We also present a novel error metric for motion accuracy assessment, which quantifies the error between source and warped volumes based on the 3D structure tensor.

2 Materials and Methods

2.1 Sample Preparation

First, we describe the acquisition of the data used in our study.

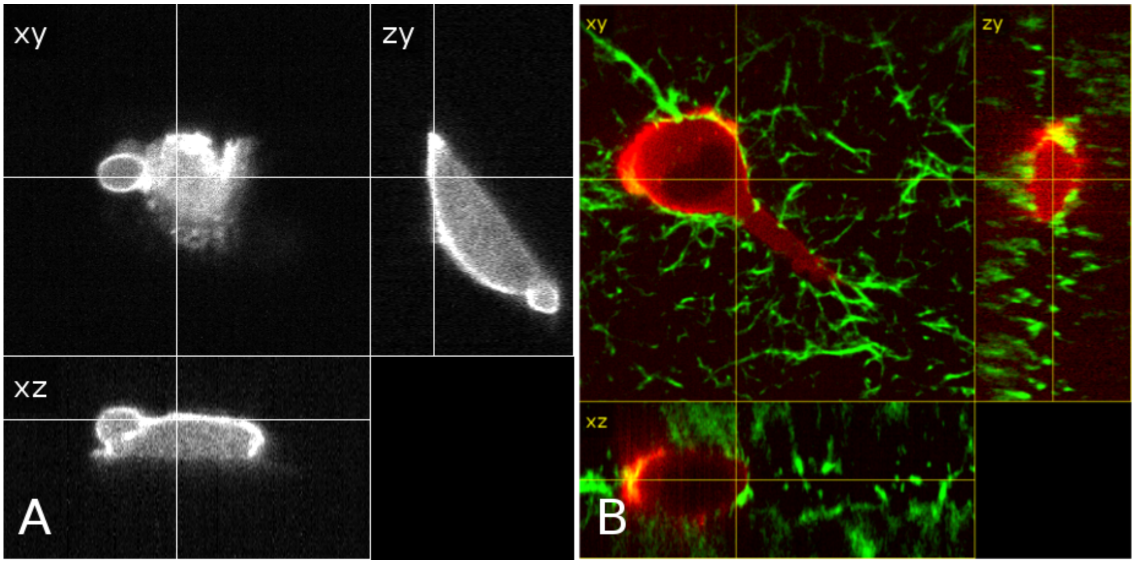

Figure 1: (A) Sequence 1 depicting a MV3 blebbing cell. (B) Sequence 2 depicting a MV3 cell (in red) retracting in collagen environment (in green).

Sequence 1: This sequence contains a MV3 cell undergoing blebbing (Charras, 2008) where small protrusions, termed blebs, slowly move along the cell surface (see Fig. 1A). F-actin was labelled in the cell through lentivirus encoding Tractin-EGFP and imaged by diaSLM microscopy (Dean *et al.*, 2016). The cell specimen was plated on a coverslip inclined at 45°. The full image sequence lasts 175 seconds acquired at 2.86 Hz with each volume of size $50 \times 50 \times 30$ μm^3 ($300 \times 300 \times 83$ voxels).
Sequence 2: This sequence shows an ezrin-labelled MV3 cell undergoing large scale retraction in collagen (see Fig. 1B). The sequence starts with a cell protrusion latched to a collagen network. The filopodia are eventually released, resulting in a motion in the direction of retraction that suggests slingshot-like behaviour. Complex dynamics are captured in both channels containing the cell and the collagen. The GFP-ezrin construct was expressed through lentivirus. The sequence was acquired using ASLM microscopy (Dean *et al.*, 2015). Each volume is $80 \times 80 \times 40$ μm^3 ($512 \times 512 \times 101$ voxels) acquired at 0.04 Hz. Throughout the text, we will refer to the channel containing the cell and the collagen as the cell channel and the collagen channel, respectively.

2.2 Motion Estimation

Our contributions focus on two issues that are specific to fluorescence imaging: a) the difficulty of associating objects from one frame to another in a noisy or a poorly descriptive fluorescence footprint, b) unpredictable displacement with a combination of small and large motions. To tackle those challenges, we propose: *i*) 3D PatchMatch method for matching voxels between two 3D images, *ii*) an extension of Census signature based variational method for 3D image sequence, and finally *iii*) the combination of the 3D PatchMatch method with 3D variational method to address large and small displacements as well.

Let I_1 and I_2 be two consecutive 3D volumes, defined over a volume domain Ω, in a sequence of 3D light microscopy (LM) images. We refer to I_1 and I_2 as the source and the target, respectively. The problem of motion estimation is to compute the 3D flow field $\mathbf{w} : \Omega \to \mathbb{R}^3$ such that any voxel pair $I_1(\mathbf{p})$ and $I_2(\mathbf{p} + \mathbf{w}(\mathbf{p}))$ corresponds to the same physical point. Here, $\mathbf{p}$ is a point defined in 3D space by coordinates (x, y, z). We refer to the three components of $\mathbf{w}(\mathbf{p})$ as u, v and w, defined for x, y and z axes, respectively. A similarity measure $d(\mathbf{p})$ is then required to match a physical point across successive time steps. We consider a 3D patch $\mathbf{P}$ in I_1 centered at voxel $\mathbf{p}$ and its counterpart patch $\mathbf{P}^c$ in I_2 centered at voxel $\mathbf{p}^c$. The best matching pair of $\mathbf{P}$ and $\mathbf{P}^c$ results in the 3D displacement vector $\mathbf{w}(\mathbf{p}) = \mathbf{p}^c - \mathbf{p}$.

2.2.1 Similarity Cost

In the literature of 2D variational optical flow methods, the similarity measure is involved in the so-called data term. With the additional constraint of smooth motion, these methods are able to estimate small motion with good accuracy. To deal with large motion, various methods have integrated pyramidal approaches and matching schemes. So far, the methods with matching schemes have better performance (Revaud *et al.*, 2015; Xu *et al.*, 2017)

Table 1: Formulations of similarity measures

SSD	$\sum_{\mathbf{q}\in\mathbf{P}} (I_1(\mathbf{q}) - I_2(\mathbf{q}+\mathbf{w}(\mathbf{p})))^2$
*CS**	$\sum_{\mathbf{q}\in\mathbf{P}} (H(r_{\mathbf{q}}) \oplus H_c(r_{\mathbf{q}+\mathbf{w}(\mathbf{p})}))$
*TZNCC***	$1 - \frac{1}{\sigma_{\mathbf{P}}\sigma_{\mathbf{P}^c}} \sum_{\mathbf{q}\in\mathbf{P}} (I_1(\mathbf{q}) - \bar{\mathbf{P}})(I_2(\mathbf{q}+\mathbf{w}(\mathbf{p})) - \bar{\mathbf{P}}^c)$

* $H(r)$ and $H_c(r)$ refer to bits of Census signatures of $\mathbf{P}$ and $\mathbf{P}^c$ respectively. The function $H(r)$ is defined in Eq.(1). $\oplus$ is the bit-wise XOR operator.

** $\bar{\mathbf{P}}$, $\bar{\mathbf{P}}^c$, $\sigma_{\mathbf{P}}$ and $\sigma_{\mathbf{P}^c}$ are the mean intensity values and standard deviations of $\mathbf{P}$ and $\mathbf{P}^c$ patches, respectively.

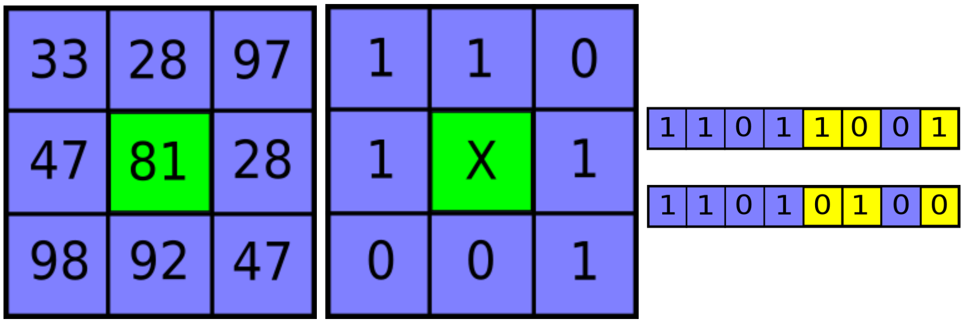

Figure 2: Left: Illustration of a 3×3 patch with intensity values. The interest point is marked in green and rest of the points are its connected neighbours. Middle : Census signature of $\mathbf{P}$. Here the Census string is 11011001. Right: Census string of two different patches. Yellow cells depict the places where bits are different. Here, the Hamming distance is 3.

In pursuit of appropriate data term for 3D fluorescence images, we compare three different similarity measures: (i) Sum of squared difference (SSD), (ii) Truncated Zero normalized cross-correlation (TZNCC), (iii) Hamming distance of Census signature (CS). These are common measures used in image registration, disparity computation in stereo-vision, and optical flow problems. SSD, also known as L^2 norm, remains a popular choice (Ourselin *et al.*, 2002) for its simplicity. TZNCC is known for its robustness to Gaussian noise but complex in computation (Di Stefano and Mattoccia, 2003). CS is recommended for its robustness to noise and intensity variation over time (Zabih and Woodfill, 1994). It is defined as:

$$H(r_q) = \mathbf{1}[r < 0], \tag{1}$$

where $r_q = I_1(\mathbf{q}) - I_1(\mathbf{p})$ and $q \in \mathbf{P}$ (see Fig. 2) and $\mathbf{1}[.]$ is the indicator function.

Formulations of these measures are given in Table 1. For comparison, we embed these three appearance based criteria within our 3D PatchMatch framework. We apply them to 3D volumes depicting mid-range to long-range motion.

2.2.2 3D PatchMatch

Our goal is to find the best matching pair $\mathbf{P}$ and $\mathbf{P}^c$ resulting in the 3D displacement vector $\mathbf{w}(\mathbf{p}) = \mathbf{p}^c - \mathbf{p}$. To this end, we extend the well-known 2D PatchMatch method (Barnes *et al.*, 2009) which proves efficient and accurate in many 2D image registration problems. It is an iterative method where displacements having best matching score are propagated from neighbours, while using random search allows for update (Barnes *et al.*, 2009). The resulting flow field involves integer displacements. The idea is to avoid brute force matching and relying on natural coherence in the images to propagate good matches. As a result, the method performs faster than other matching algorithms (Muja and Lowe, 2009), while keeping a low memory footprint. The algorithm proceeds in two stages: 1) initialization , and 2) iteration between displacement propagation and random search.

Initialization: In this stage, we initialize the displacement field either with random vectors or vectors approximated with coarse matching.

Propagation: In this stage, for any iteration i, the current displacement vectors are propagated. The scanning of volume is done slice by slice and in raster scan manner for each slice. In the scanning process, each current matching score for patch of a given voxel is compared with those of its causal neighbours (i.e., neighbours just previously scanned). If a neighbour has the best matching score, its displacement vector is copied to the current voxel. This enforces smoothness in the resulting flow field. The matching score is determined by one of the similarity measures discussed in 2.2.1.

Random Search: In this stage, we try to get a better match by first randomly selecting patches in I_2 around the current best estimate. Iterations of random sampling are done in a cuboid region having side length of R. If the randomly selected patch yields a better matching score, it becomes the new best corresponding patch. At next iteration, a new random sample will be drawn from a region with side length of $R/2$ around the current best estimate. These iterations of random sampling continue by halving the side length until the cuboid region has unit volume.

We implement 3D PatchMatch in a coarse-to-fine manner. We use three resolution levels, each one being half the resolution of the preceding one. The top or the coarsest level starts with random initialization. After few iterations of propagation and random search, a coarse flow field is obtained. It is then upscaled and upsampled, and used as the initialization for the next level. This scheme reduces the number of iterations required in the lower levels by providing a good guess for the flow field. Also, the coarse-to-fine strategy guarantees that the large displacements are captured at the top level with low ambiguity and are refined at lower levels. The final 3D flow field might have false matches which are removed by forward-backward consistency check. The final flow field is semi-dense.

2.2.3 Residual Matching Cost

We define the residual matching cost as the similarity measure estimated for the matched pair in the final iteration of 3D PatchMatch. We use this cost to compare the three similarity measures introduced in Section 2.2.1. The cost is of importance as the flow field estimated by 3D PatchMatch is due to the propagation of displacements precisely having low residual matching cost. Trustworthy measurements are expected in the labeled regions, which are the ones of key interest. Residual matching cost helps us evaluate matches in such regions.

Figure 3 summarizes the evaluation. The range of the matching costs are as follows: SSD range is $[0, \infty)$, CS range is $[0, |P|]$ where $|P|$ is the number of neighbors, TZNCC range is $[0, 2]$. The lower the cost, the better the confidence in matching. Even though the costs are not directly comparable, their juxtaposition in Fig.3 reveals how confidently the final matches were made when estimating motion depicted in Fig. 4.

The raw data display the labeled region with high intensity and the unlabeled cytosol as noisy interior. TZNCC and CS exhibit good confidence in the labeled region in contrast to the cytosolic region. Surprisingly, SSD is assertive in unlabeled nucleus and cytosolic region, and unreliable in the labeled region. This means that the displacements computed using SSD are propagated from unlabeled region rather than trustworthy labeled region. Figure 4 juxtaposes 3D flow fields computed using the state-of-the-art method by Amat *et al.* (2013), and our 3D PatchMatch using SSD, CS and TZNCC measures, respectively. Clearly, Amat *et al.* (2013)'s method fails to capture large displacement of the protrusion. Our method proves to be effective here. Further, the flow fields computed using CS and TZNCC are less noisier than that of SSD.

Robustness of TZNCC comes from the normalization of the intensities with the local average and local standard deviation. Robustness of CS comes from the composition of the Census signature which captures the underlying structure of the patch by the arrangement of bits representing the binarized directional derivatives of the intensity. Granularity of the intensity distribution of patches from these regions can be visualized by the rank transform (Zabih and Woodfill, 1994). Rank of a voxel is the total number of voxels in its neighbourhood which have smaller intensity values than itself. Figure 5 illustrates that the labeled regions have high and uniform rank, whereas the background and unlabeled region inside the cell have non-uniform rank distribution. The uniformity and the high rank in patches of labeled region suggest less number of unexpected bit flips when comparing two labeled patches, and hence, smaller Hamming distances in these regions. Consequently, the propagation is driven by the displacements in trustworthy labeled region rather than those in the unlabeled regions. In conclusion, we found CS and TZNCC to be more discriminative than SSD. Taking the computational complexity into account, we choose CS over TZNCC to carry further with the variational refinement method. Additionally, the over-all confidence of CS appears to be more coherent within the cell than that of TZNCC.

2.2.4 Variational Method

3D PatchMatch is efficient in computing semi-dense integer displacements. To compute dense sub-voxelic displacements, we design a 3D variational method. A generic form of variational function based on the minimization of the energy is:

$$E(\mathbf{w}) = \int_{\Omega} D(I, \mathbf{w}) + \alpha \mathcal{R}(\mathbf{w}) d\Omega, \tag{2}$$

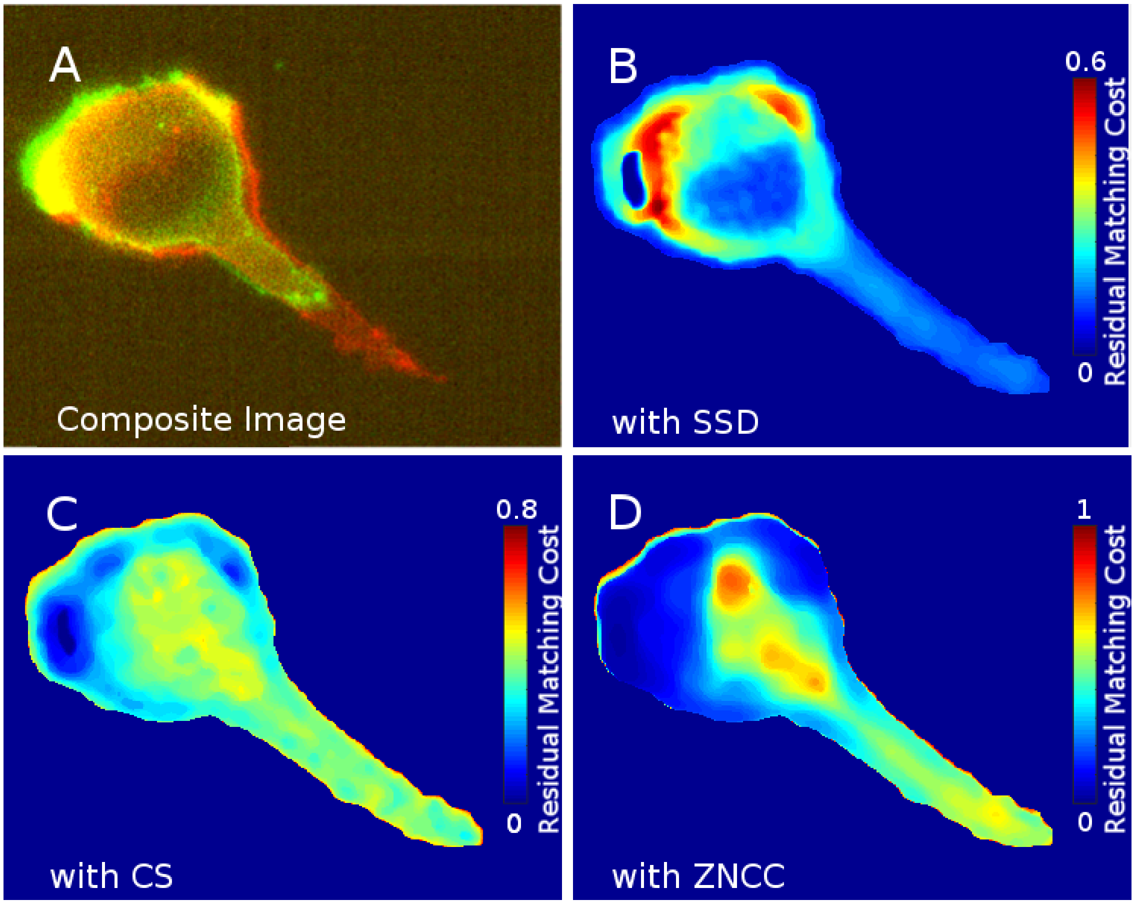

Figure 3: Residual matching cost. (A) Overlapped consecutive frames in sequence **2** where source is in red and target in green. Residual matching cost using (B) SSD, (C) CS, and (D) TZNCC based measure, respectively. The color code for the residual matching cost is given in the right side of the respective figures and normalized to the maximum of computed values of the respective measures in the given slice.

where $D(I, \mathbf{w})$ is the data term that penalizes variation of intensities, $\mathcal{R}(\mathbf{w})$ is a regularizer for the 3D flow field $\mathbf{w}$ and α is the balancing parameter for the two terms. Horn and Schunck (1981) and Amat *et al.* (2013) used the brightness constancy assumption in the data term as L^2 norm and Huber norm (Huber, 1964), respectively. In our work, we exploit the Census signature constancy assumption as we have done with 3D PatchMatch in subsection 2.2.2.

In what follows, we consider the use of quadratic regularizer, $\mathcal{R}(\mathbf{w}) = \|\nabla \mathbf{w}\|_2^2$ (Horn and Schunck, 1981), considering a smooth 3D flow field, where ∇ is the spatial gradient operator and the function D is the data term based on the Census constancy assumption. The Census signature defined in Eq. 1 cannot be linearized. As a consequence, we take the approximation of the Census signature in continuous form proposed in (Hafner *et al.*, 2013). To simplify notations, we combine spatial and temporal terms to denote $\mathbf{p} = (x, y, z, t)^\top$ in I_1 and $\mathbf{p}' = (x + u, y + v, z + w, t + 1)^\top$ in I_2. We write the spatio-temporal gradient operator as ∇_4 and the motion vector to be estimated as $\mathbf{w}(\mathbf{p}) = (u, v, w, 1)^\top$. Then D can be defined in a local spherical coordinate system around $\mathbf{p}$ as :

$$D(I(\mathbf{p}), \mathbf{w}) = \int_0^{2\pi} \int_0^{\pi} H_\epsilon'^2(\partial_\phi^\theta I(\mathbf{p}))(\mathbf{w}(\mathbf{p})^\top \nabla_4(\partial_\phi^\theta I(\mathbf{p})))^2 d\theta d\phi, \tag{3}$$

where ∂_ϕ^θ acts as a spatial directional derivative operator with direction given by the spherical angles (ϕ, θ) in a unit sphere around central voxel at $\mathbf{p}$, with $\phi \in [0, 2\pi)$, $\theta \in [0, \pi]$. Practically, we discretize ϕ and θ to obtain a 26-connected neighbourhood around $\mathbf{p}$. Please refer to the supplementary document for the further elaboration on the continuous Census signature. Here, H_ϵ and H_ϵ' are the approximated Census signature (Hafner *et al.*, 2013) and its first derivative respectively, which are given by :

$$H_\epsilon(r) = \frac{1}{2}\left(1 + \frac{r}{\sqrt{r^2 + \epsilon^2}}\right), \qquad H_\epsilon'(r) = \frac{\epsilon^2}{2(r^2 + \epsilon^2)^{3/2}}. \tag{4}$$

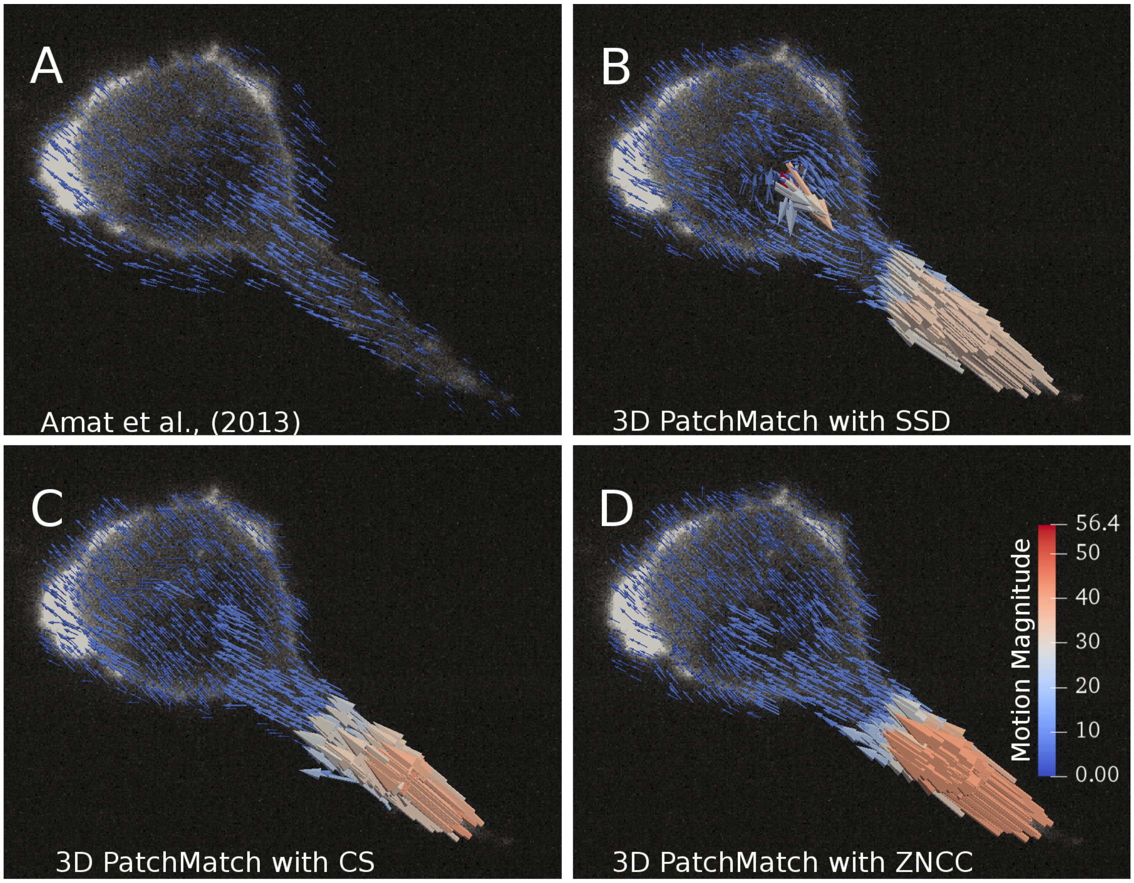

Figure 4: Motion magnitude estimated using (A) Amat *et al.* (2013) method, and 3D PatchMatch with (B) SSD, (C) CS, and (D) TZNCC measures, respectively. The 3D flow field is plotted with arrows. The color code for the magnitude of the motion is given on the right side of sub-figure D.

Minimization of Eq.(2) with the aforementioned data term and regularizer can be solved through associated Euler-Lagrange equations (see Section 1 in the supplementary document).

2.3 Combination of the two methods

Often, a sequence contains large and small motions. To demonstrate this scenario, we use sequence 2 of MV3 cell retraction in collagen (see figs. 3, 9). 3D PatchMatch is able to handle large displacements but often fails where sub-voxelic motion is present. Conversely, the variational method on its own is only robust for small displacements. It can also be employed to deal with medium motion using coarse-to-fine approach. The idea is to compute motion in coarser levels where larger displacements are scaled down to smaller ones. However, the number of levels to effectively scale down large motion is not easy to set a priori. Further, this approach fails to capture motion of smaller objects as they get lost at coarsest levels and subsequently in warping stages. In contrast, 3D PatchMatch does not face such issues. To deal with all possible large displacements, initial search region at the coarsest level for 3D PatchMatch is set to the whole volume domain. Hence, the number of coarser levels are not critical. There is no warping performed between levels, which preserves smaller objects to be present in successive finer levels.

To deal with mixed motion, we adopt an incremental approach as in Revaud *et al.* (2015). For this, we first compute large integer displacements using 3D PatchMatch. To smoothen and densify the flow field, Gaussian interpolation is applied. The remaining smaller displacements can be dealt with using our variational method. By warping I_2 backward with the interpolated flow field $\mathbf{w}'$, we obtain a warped volume $\tilde{I}_2$ which is similar to I_1. We then proceed to the variational step using I_1 and $\tilde{I}_2$ as input. The flow field $\tilde{\mathbf{w}}$ obtained from this step possesses smaller displacements, which are added to the displacements obtained from 3D PatchMatch. Thus, the final flow field is computed by vector summation of flow fields obtained from the both stages. Algorithm 1 presents the steps involved.

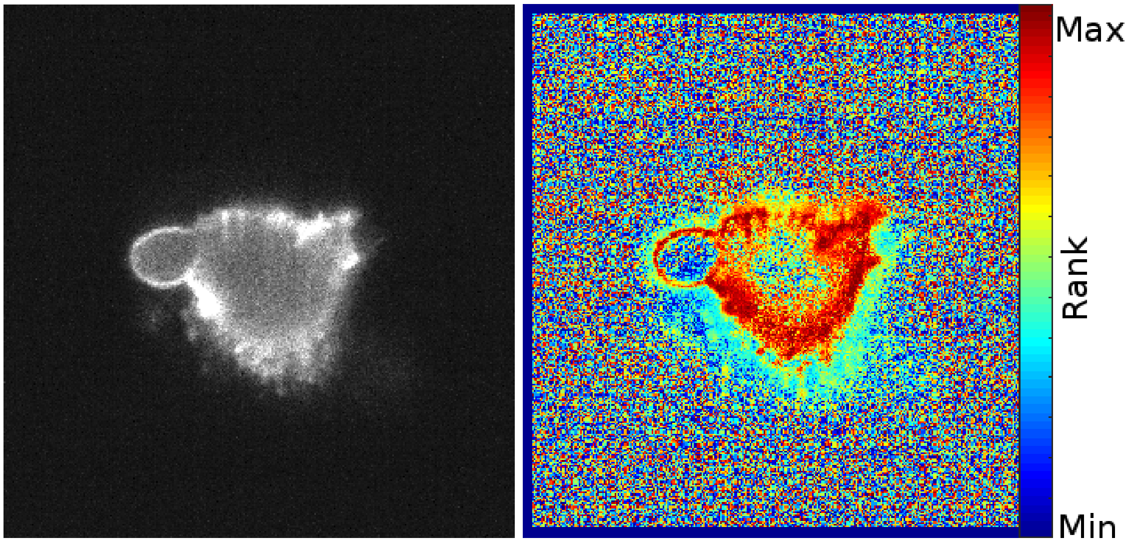

Figure 5: Left: A raw slice of a frame in sequence **2**. Right: The rank transform. The color code for the patch rank is given on the right side, where Min = 0, Max = $|\mathbf{P}|$.

Algorithm 1 Combining 3D PatchMatch and Variational method

```
INPUT: A pair of 3D images $I_1$ (source) and $I_2$ (target)
OUTPUT: 3D flow field $\mathbf{w}$
$\mathbf{w}' \leftarrow 3DPatchMatch(I_1, I_2)$        ▷ $\mathbf{w}'$ is the semi dense flowfield
$\mathbf{w}' \leftarrow GaussianInterpolation(\mathbf{w}')$        ▷ densify $\mathbf{w}'$
$\tilde{I}_2(\mathbf{p}) \leftarrow I_2(\mathbf{p} + \mathbf{w}'(\mathbf{p})), \forall \mathbf{p} \in \Omega$        ▷ Warp $I_2$
$\tilde{\mathbf{w}} \leftarrow Variational(I_1, \tilde{I}_2)$        ▷ compute flow field $\tilde{\mathbf{w}}$ with Eq. 2
$\mathbf{w} \leftarrow \tilde{\mathbf{w}} + \mathbf{w}'$        ▷ sum up $\mathbf{w}'$ and $\tilde{\mathbf{w}}$
```

2.4 Assessment

Due to the lack of ground truth, to assess the accuracy of the computed 3D flow field, we mostly rely on visual inspection. We also propose a quantitative approach to validate the results based on the reconstruction of volume from the computed 3D flow field.

Visualization: A straight-forward way to visualize a 3D flow field is to use 3D glyphs or arrows to represent direction and magnitude. To avoid too many overlapping vectors, the plot is subsampled. It then becomes difficult to establish visual consistency between such representation and raw image data which is typically viewed slice by slice. Furthermore, such rendering is difficult for the purpose of comparison and communication.

Color-coded flow map is now a common practice for viewing 2D optical flow fields (Baker *et al.*, 2011). Color mapping of 2D vector field was described in the work of Abramoff *et al.* (2000), where they proposed to map the Hue subspace to the direction and the Saturation subspace to the magnitude of the vector, while keeping the Value of the map to 1. However, creating a 3D color map for 3D flow fields is not straightforward. One might presume that Value space could be used to represent the additional dimension. This type of mapping for 3D fluorescence volume sequences is not trivial for two reasons:

- Even when the Value space is used for the 3D vector representation, there are not enough colors discernible to human eyes to accommodate 3D vectors of all orientations and amplitudes.
- Viewing 3D flow field for fluorescence data in 2D planes has to be done smartly as not all portions of the flow field are informative.

One of our contributions lies in resolving these issues by introducing two different techniques that makes color mapping of 3D flow fields possible.

The first method takes a slice-by-slice approach, which we call 3DHSV map. 3D flow field of each slice is used to create a mapping in HSV color space. The direction of the motion is mapped to hue, the magnitude is mapped to saturation and the off-the-plane motion is mapped to value. We use only x and y components to map the direction to hue. The magnitude is computed using all three components.

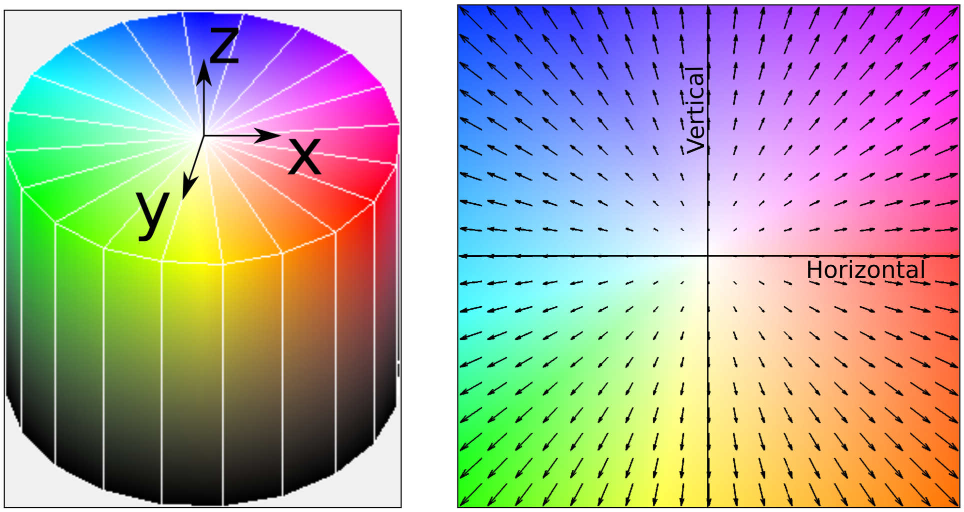

Figure 6: Color code for 3DHSV (left) and 3PHS (right) maps.

The value-components are given by the z-component of the flow field such that voxels in the slice with downward and upward trajectories in z-axis are brighter and darker respectively. One can select the minimum and maximum amplitude of off-the-plane motion so as to create a discernible range of colors pertaining to amplitude of off-the-plane motion. The color code is depicted as HSV cylinder in Fig. 6 (left). Now, it becomes possible to view the motion field in the manner similar to slice-by-slice viewing of volumetric images. However, slice-by-slice viewing of cellular dynamics might not always be revealing. This brings us to the next visualization method.

The interesting motion might be lying in any arbitrary plane that might not be obvious in a slice-by-slice view. To this end, we take a multi-view approach. Multi-view of 3D flow field may be more informative than the previous slice-by-slice approach. To generate multi-view of the 3D flow field, we slice the 3D flow field with a 2D plane and project the sliced 3D vectors (i.e., vectors whose origin belongs to the 2D plane) to the selected 2D plane. The slicing plane can be any arbitrary plane. We have proposed the simpler orthogonal planes of the 3D volumetric data. We call it 3PHS mapping. The choice of planes remains in the hand of user. Usually, the planes that provide the best cut-off of background voxels and the best intersection of specimen signal should be chosen. The projected 2D flow fields can now be visualized with the usual 2D Hue-and-Saturation (HS) mapping. The color code is depicted as HS map in Fig. 6 (right).

3DHSV map can be rendered by processing the entire 3D flow field in one pass, whereas 3PHS requires each plane to be defined before rendering. The result of former method is portable as it can be saved in 3-channel image stack. For the latter method, it is not practical to save the map for all planes. Therefore, we propose to have the user select the desired point in the volume from where three planes can be defined just before viewing. The normalization of value in 3DHSV map may be required according to the user's ability to discern light and dark pixels. However, both methods facilitate the perception of discontinuity and smoothness of motion.

Accuracy estimation: The most known measures from the literature of 2D optical flow are End-Point-Error (EPE) and Average Angle Error (AAE), which are the Euclidean norm and cosine of angle between the estimated and the ground-truth flow vectors, respectively. When no ground truth is available, it is challenging to define measures with the same objective. Zhu and Milanfar (2010) proposed an image-content metric for image restoration, when no reference image is available. This no-reference metric is based upon singular values of local image gradient matrix. They use the global average of the metric to assess the performance of denoising algorithms and automatically set the parameters. In a similar spirit, we propose a novel metric for assessing the performance of optical flow algorithms, when no ground truth is available. Our metric is based on the eigenvectors of local structure tensors.

To objectively quantify the accuracy of the computed 3D flow field, the proposed metric relies on the comparison between the warped and the source volume. We compute a reconstructed volume I_r which should be close to I_1 by backward warping I_2 using the computed 3D flow field. Then, we compare the structures I_r with the real source volume I_1. Comparing structure is better than the commonly used displaced frame difference (DFD), given as $DFD(\mathbf{p}) = |I_2(\mathbf{p} + \mathbf{w}(\mathbf{p})) - I_1(\mathbf{p})|$, as it is robust to noise and independent of the type of interpolation used for warping. DFD tends to be lower where the intensity gradient is relatively low, even though the flow vector is erroneous. It tends to be high where the intensity gradient is strong (i.e., around a sharp edge) even when the flow vectors are close to the ground truth. The structure-tensor-based comparison, however, does not manifest such disproportionate behaviour. Our idea is to rely on the underlying image structure rather than on the image intensity itself as in DFD. The former is both more constraining and differentiating for the flow. Under this premise, we propose a new error measure called Structural Angular Error (SAE). Let us emphasize that it is still a quantitative error measure, not a qualitative one, in the absence of ground truth.

We compute and compare the principal orientation of 3D structure in the volume using the smallest order eigenvector of the structure tensor (Ahmad *et al.*, 2015). Let $V_r(\mathbf{p})$ and $V_1(\mathbf{p})$ be the smallest order eigenvectors of the structure tensors at $\mathbf{p}$ in I_r and I_1, respectively. We then compute the structural angular error (SAE) between $V_r(\mathbf{p})$ and $V_1(\mathbf{p})$ according to :

$$SAE(\mathbf{p}) = \arccos\left(\frac{V_r(\mathbf{p})}{|V_r(\mathbf{p})|} \cdot \frac{V_1(\mathbf{p})}{|V_1(\mathbf{p})|}\right) \tag{5}$$

A map of SAE serves as an error map for the given 3D flow field. On the other hand, as a global metric, the average of SAE over the foreground region provides a quantitative means to assess the performance of optical flow algorithms. It can also serve to find the right parameterization of the method.

3 Experiments

Depending on the nature of the specimen under study, one might run the variational method alone when the specimen exhibits small motion and the combination of 3D PatchMatch and variational stage when the specimen exhibits mixed motion. For sequence **1**, which involves a slow process of blebbing, we have used the variational stage with the energy form of Eq.(2). For sequence **2**, which displays a large retraction, we have applied the combination of both methods.

3D PatchMatch: Given the anisotropy factor of 3 in the z-axis for the given data, we choose a patch of size $9 \times 9 \times 3$. A patch should be small enough to capture the moving structure while not increasing the computation time. From the coarsest to the finest levels, we apply $7, 3, 1$ iterations, respectively. Any further iterations did not improve the results. Median filter of size 3×3 was applied to the 3D flow fields computed in each level, before upscaling and upsampling for outlier removal.

Variational stage: We warped the target volume with the flow field obtained from 3D PatchMatch. We applied Gaussian interpolation on the flow field and then, warped the target volume using trilinear interpolation. The Census approximation factor ϵ was set to 50. The regularization parameter α was set to 0.11 for the channel containing cell data and 0.05 for the collagen channel. As the later parameter controls the smoothness of the flow field, we recommend values in the range $(0.1, 2]$ for the cell channel where smooth results are expected, and lower values for the collagen channel where the flow field is not as smooth. Successive over relaxation (SOR) (David M. Young, 1971) with 10 iterations was employed for solving the related Euler-Lagrange equations.

Post Processing: In order to produce clear visualization, we only take relevant parts of the flow field into account. For this, we segment the source volume using Huang and Wang (1995)'s method, and use the obtained binary mask to restrict the flow field to the foreground region.

For comparison purpose, we implemented a straight-forward 3D version of (Horn and Schunck, 1981) (HS3D). We also compare with the Amat's method. We have also generated a set of volume pairs with ground-truth flow fields. The goal is two-fold: validating the ASAE behaviour by comparison to EPE measure, and augmenting the quantitative assessment of our 3D optical flow method.

Instead of employing arbitrary and unrealistic synthetic flow fields to generate sequences, we simply apply HS3D method to pairs of real volumes to get flow fields. These flow fields are not accurate enough, but this is not the matter here; what matters is that they are closer to realistic motion fields than any arbitrary fields. It then allows us to make quantitative comparison between our method and Amat's method, as if the ground-truth flow was available. More specifically, we use the HS3D computed flow fields for backward warping the target volumes. A new source volume is generated for every target volume. The flow field then serves as the ground truth for this new volume pair.

We utilize Sequence 1 for the generation, as it was acquired at the highest sampling frequency (3.5Hz) among other sequences at our disposal. The sequence is characterized by a dominant bleb with slow and presumably translational motion. We choose a single volume as the source and 10 consecutive target volumes at incrementing time step. This would mean that the computed motion should be increasing with each volume pair in the dominant bleb region. Finally, we generate 10 pairs of volumes for each time steps of the original sequence by backward warping using the HS3D computed flow field. Each of them is of size 144x150x83. We then compare our method against Amat's method. As seen in Fig. 10 (left), clearly, our method outperforms Amat's method with respect to the average EPE. We also compute our proposed Average Structural Angular Error (ASAE) for the flow fields computed with the two methods (see Fig. 10 (right)). The ASAE not only reveals the best performing method but also the trend in error at different time steps for each method, which closely follows that of the average EPE. This proves that the ASAE can be used to assess optical flow accuracy when no ground truth is present. The percentile and average values of EPE and SAE for volume pair at each time step are given in Tables 2 and 3, respectively. The visualization of the ground-truth flow fields and the computed flow fields are provided in the supplementary document.

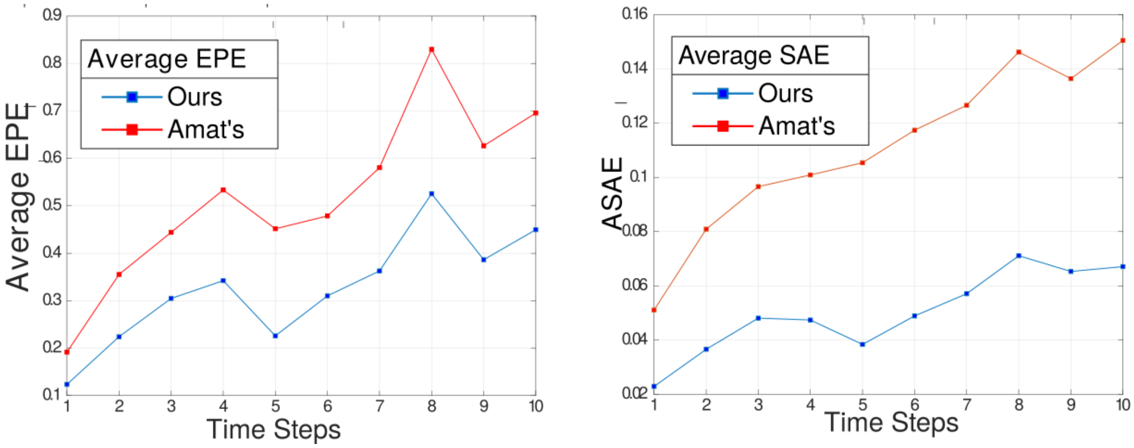

Figure 10: Left: Average EPE for each pair in time steps. Right: Average SAE (ASAE) for each pair in time steps.

Real sequence: Figures 7 and 8 show the results of the three methods. Visually (see Section 2 in the supplementary document), the belb follows the bottom-right direction in xy-plane. The flow field computed by our method and Amat's method reflect the direction. Meanwhile, the result of HS3D shows incorrect motion to the right. Amat's method computes piecewise constant motion for super-voxels. As a consequence, it produces block-like artifacts in the flow field. In contrast, our approach estimates a smooth motion. The regularization outside of the cell body is limited, since we do not perform optimization with a large number of iterations. In any case, as a post-processing stage, we use the foreground segmentation mask for our results and HS3D results to clip off the expansion. Amat's method applies the segmentation mask during the optimization at each iteration, thus producing flow field only in the foreground region.

Now, we elaborate on the results obtained while using the 3D PatchMatch and the variational method separately, and emphasize on the combination of the methods. 3D PatchMatch by design only computes integer displacements. Smoothness of the computed flow field arises from propagation only, and does not provide much regularization in practice. To illustrate this, we use collagen channel in sequence **2**, which depicts the motion of collagen, while the cell undergoes large retraction motion. Figure 9A displays source and target frames in red and green channels, respectively. The overlapping voxels appear in yellow. Figure 9B displays source and reconstructed volume using the flow field obtained from 3D PatchMatch in red and green channels, respectively. For the most part of the reconstructed volume, the overlapping is significant except where structures undergo complex deformations or exhibit small displacements. The next two figures show 3PHS flow map of flow field computed by 3D PatchMatch only and by subsequent refinement using the proposed variational method. The refinement step captures smaller displacements which are missing from the flow field obtained in 3D PatchMatch stage. The red rectangle in the figs. 9C and 9D highlights one instance (best viewed in electronic form). Table 3 reports the average SAE (ASAE) computed using HS3D, Amat's method and our method. Clearly, our method outperforms the two other ones. It is interesting to note that, for the cell channel of the Sequence 1, there is not a large error margin between our method and Amat's method. Despite the obvious failure of Amat's method to compute large displacement of the Filopodium, and visually better performance of our method (see Fig. 4A), the ASAE shows a small overall improvement only. Upon the inspection of the SAE maps (see Fig. 11), we found that both of the methods exhibit high SAE values in the cytosol region which does not have any (fluorescence) labeling. In the more biologically interesting filopodium region, where Amat's method failed, our method prevails and this is precisely shown by the SAE values. Because of higher errors in the larger cytosol region for both methods, the two global ASAE increase, and are finally close to each other (gap of 0.04), hiding the difference of performance between our method and Amat's method on the Filopodium part. We report the mean and the median SAE values in the two regions for both methods in the supplementary document.

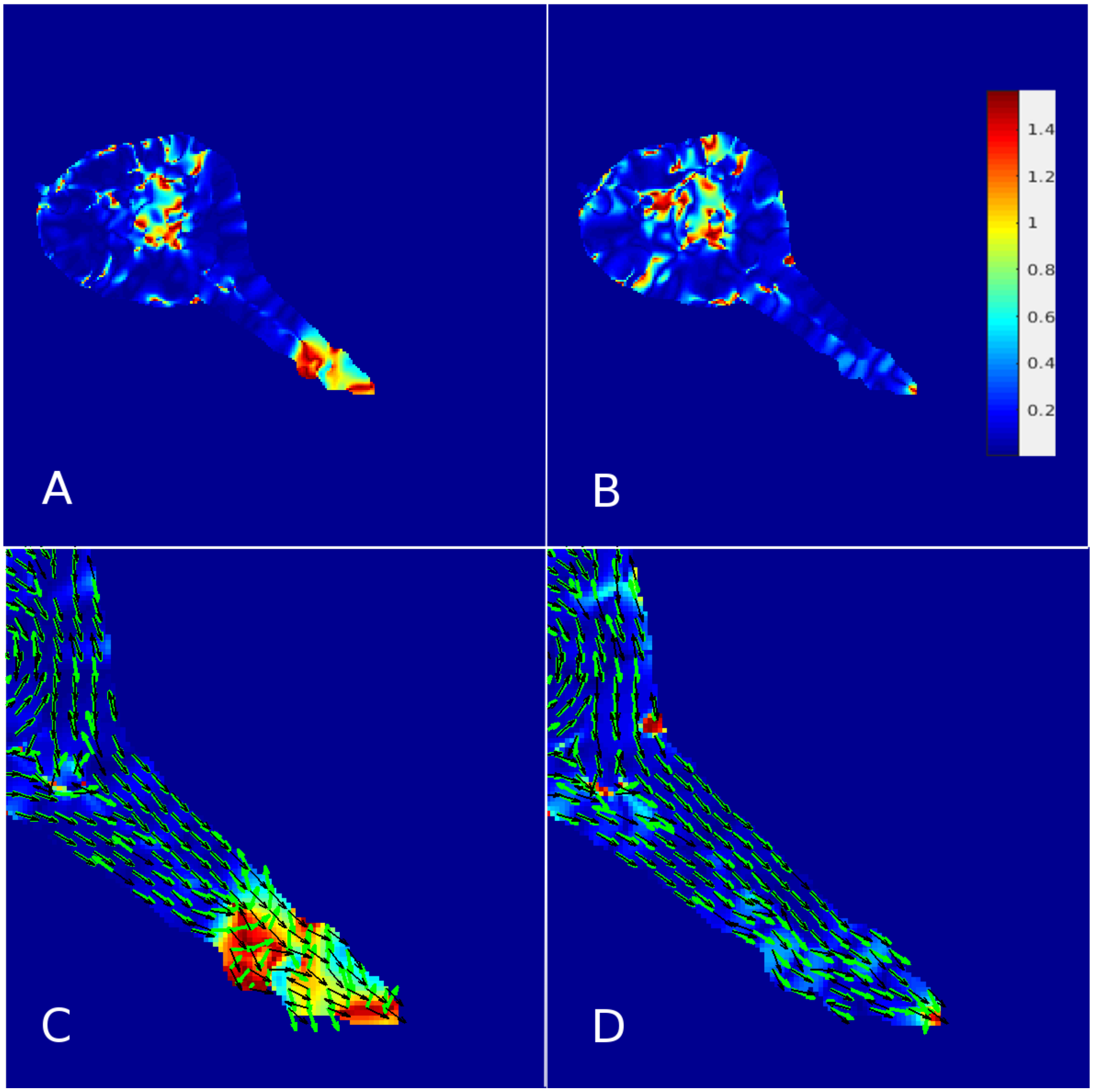

Figure 11: SAE map for the optical flow computed with (A) Amat's method (B) Our method. The color code is given in the right side of (B). Principal orientation of the structures denoted by arrows in black for I_1, and in green for I_r by the optical flow computed using (C) Amat's method and (D) Our method.

We point out that SAE is not applicable for 3D PatchMatch alone because a semi-dense flow field is not sufficient for faithful reconstruction of I_1.

Time complexity of our algorithm is very similar to HS3D as we use SOR for both implementations. The time complexity of SOR method is $O(kN)$, where N is the size of the data and k is the number of iterations. However, additional complexity of $O(nN)$ arises from the computation of the Census Signature, where n is the size of the discrete neighborhood. The aspects that can be parallelized are the computation of the Census Signature and its derivatives. For this, each element of the Census Signature has to be stored as a 3D matrix. Then, we can compute Eq. 4 for each voxel in a parallel manner. We can also parallelize the minimization stage, if we use a splitting method instead of SOR algorithm, which could be a prospect for our work.

These experiments were conducted on a 16 GB machine running with 2.80 GHz Intel core i7 processor. The methods were implemented with C++ language. For a volume of size 512x512x101, the 3D PatchMatch with CS takes 8, 60 and 578 seconds per iterations for three coarse-to-fine levels, respectively. Each iteration for SOR takes around 50 seconds. To reduce the computation time, we can divide the image resolution by two. This requires 3D PatchMatch to run only for the first two coarse levels. The variational method in this case takes 7 seconds per iteration. Finally, the 3D flow field is upsampled and upscaled.

Table 2: Percentile and average EPE for the generated pairs.

Time Step	Amat's method				Ours			
	90%	95%	99%	Mean	90%	95%	99%	Mean
1	0.4382	0.6264	1.1269	0.1917	0.2532	0.3587	0.6799	**0.1237**
2	0.7761	1.1291	2.7410	0.3553	0.3977	0.5491	1.3267	**0.2241**
3	0.9444	1.3608	3.3245	0.4441	0.5534	0.7377	1.7847	**0.3048**
4	1.1617	1.6334	4.0799	0.5335	0.6222	0.8556	2.1611	**0.3423**
5	1.0408	1.4602	2.7193	0.4517	0.5024	0.6944	1.2492	**0.2260**
6	1.1278	1.7025	3.3711	0.4786	0.6358	0.8779	1.8720	**0.3102**
7	1.4597	2.1316	4.2987	0.5804	0.7422	1.0578	2.3148	**0.3628**
8	1.9364	2.6369	5.9004	0.8297	0.9450	1.3820	3.9761	**0.5254**
9	1.5825	2.4362	4.7762	0.6265	0.8043	1.1341	2.3198	**0.3864**
10	1.7317	2.6051	5.7586	0.6955	0.8781	1.2143	3.0211	**0.4497**

Table 3: Percentile and average SAE for the generated pairs.

Time Step	Amat's method				Ours			
	90%	95%	99%	Mean	90%	95%	99%	Mean
1	0.1169	0.1898	0.4695	0.0511	0.0493	0.0768	0.1907	**0.0229**
2	0.1855	0.3086	0.7945	0.0809	0.0802	0.1216	0.2889	**0.0366**
3	0.2263	0.3849	0.9691	0.0966	0.1035	0.1590	0.4025	**0.0481**
4	0.2397	0.3814	0.8915	0.1009	0.1011	0.1534	0.3715	**0.0474**
5	0.2466	0.4136	0.9964	0.1054	0.0876	0.1326	0.3153	**0.0384**
6	0.2808	0.4679	1.1921	0.1174	0.1076	0.1689	0.4440	**0.0489**
7	0.3119	0.5285	1.2221	0.1266	0.1275	0.2021	0.5201	**0.0571**
8	0.3510	0.5813	1.2449	0.1462	0.1572	0.2501	0.6489	**0.0711**
9	0.3496	0.5862	1.3142	0.1364	0.1469	0.2327	0.6522	**0.0653**
10	0.3873	0.6601	1.3369	0.1505	0.1488	0.2440	0.6614	**0.0671**

4 Conclusion

We have introduced a generic framework for 3D motion estimation in 3D LM images sequences. It combines matching and variational method while using Census signature in both stages. We first demonstrated the effectiveness of the Census signature in the matching stage. Drawing on this, we formulated a 3D variational method using the Census constancy assumption. We tested our methods on different sequences containing not only the cell but also the collagen. Our approach aims to be general enough, that is, to handle specimens with different shapes and motion types. Cell and collagen are typical specimens, which are structurally different and exhibit complex motion. We demonstrated that the variational method with the Census signature in the data term outperforms the brightness constancy term used in HS3D. While the matching stage is not necessary for sequence exhibiting small motion, combination of the methods is effective in motion estimation where mixed motion types are expected. Further, we defined two new visualization techniques for the assessment of 3D motion fields. We also presented a novel error (ASAE) measure for the purpose of quantitative validation of 3D optical flow when the ground truth is absent.

Acknowledgements

This work has been supported by French National Agency (France-Bioimaging ANR-10-INBS-04, DALLISH-ANR-16-CE23-0005), and the NIH grant R01-GM067230, HFSP grant LT000954/2015. It was carried out partly in the context of CytoDI project between: Serpico, Inria-Rennes, France and Lyda Hill Department of Bioinformatics, UT Southwestern, Texas, USA.

Table 4: ASAE for real sequences.

Sequence	HS3D	Amat *et al.* (2013)	Ours
Sequence 1	0.15	0.34	0.14
Sequence 2 (Cell)	0.6	0.46	0.42
Sequence 2 (Collagen)	0.42	0.39	0.27

References

Abramoff, M., Niessen, W., and A. Viergever, M. (2000). Objective quantification of the motion of soft tissues in the orbit. *IEEE trans. on Med. Imag.*, **19**, 986–95.

Ahmad, R. K. *et al.* (2015). 3D structure tensor analysis of light microscopy data for validating diffusion MRI. *NeuroImage*, **111**, 192–203.

Amat, F. *et al.* (2013). Fast and robust optical flow for time-lapse microscopy using super-voxels. *Bioinformatics*, **29**(3), 373–380.

Baker, S. *et al.* (2011). A database and evaluation methodology for optical flow. *Int. J. Computer Vision*, **92**(1), 1–31.

Barnes, C. *et al.* (2009). PatchMatch: A randomized correspondence algorithm for structural image editing. In *Proc. of Special Interest Group on Computer Graphics and Interactive Techniques*, pages 24:1–24:11, New Orleans.

Boquet-Pujadas, A. *et al.* (2017). Bioflow: a non-invasive, image-based method to measure speed, pressure and forces inside living cells. *Scientific Reports*, **7**(1), 9178.

Brox, T. *et al.* (2004). High accuracy optical flow estimation based on a theory for warping. In *Proc. of European Conference on Computer Vision*, pages 25–36, Heidelberg.

Brox, T. and Malik, J. (2011). Large displacement optical flow: descriptor matching in variational motion estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **33**(3), 500–513.

Chan, K. G. and Liebling, M. (2015). Estimation of divergence-free 3D cardiac blood flow in a zebrafish larva using multi-view microscopy. In *Proc. of Int. Symposium on Biomedical Imaging*, pages 385–388, New York.

Charras, G. (2008). A short history of blebbing. *J. of Microscopy*, **231**(3), 466–478.

David M. Young, J. (1971). In *Iterative Solution of Large Linear Systems*. Academic Press NY.

Davidenko, N. *et al.* (2016). Evaluation of cell binding to collagen and gelatin: a study of the effect of 2D and 3D architecture and surface chemistry. *J. of Materials Science*, **27**(10), 148.

Dean, K. M. *et al.* (2015). Deconvolution-free subcellular imaging with axially swept light sheet microscopy. *Biophysical J.*, **12**(108), 2807–15.

Dean, K. M. *et al.* (2016). Diagonally Scanned Light-Sheet Microscopy for Fast Volumetric Imaging of Adherent Cells. *Biophysical J.*, **6**(110), 1456–65.

Delpiano, J. *et al.* (2012). Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy. *Machine Vision and Applications*, **23**(4), 675–689.

Di Stefano, L. and Mattoccia, S. (2003). Fast template matching using bounded partial correlation. *Machine Vision and Applications*, **13**(4), 213–221.

Fortun, D. *et al.* (2013). Correlation and variational approaches for motion and diffusion estimation in fluorescence imaging. In *Proc. of European Signal Processing*, pages 1–5, Marrakech.

Fortun, D., Bouthemy, P., and Kervrann, C. (2015). Optical flow modeling and computation: A survey. *Computer Vision and Image Understanding*, **134**, 1–21.

Hafner, D. *et al.* (2013). Why is the census transform good for robust optic flow computation? In *Proc. of Scale Space and Variational Methods in Computer Vision*, pages 210–221, Graz.

Horn, B. and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, **17**, 185–203.

Hu, Y. *et al.* (2016). Efficient coarse-to-fine PatchMatch for large displacement optical flow. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 5704–5712, Las Vegas.

Hu, Y. *et al.* (2017). Robust interpolation of correspondences for large displacement optical flow. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 4791–4799, Honolulu.

Huang, L.-K. and Wang, M.-J. J. (1995). Image thresholding by minimizing the measures of fuzzines. *Pattern Recognition*, **28**, 41–51.

Hubený, J. *et al.* (2007). Estimating large local motion in live-cell imaging using variational optical flow towards motion tracking in live cell imaging using optical flow. In *Proc. of Computer Vision Theory and Applications*, pages 542–548, Barcelona.

Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, **35**, 73–101.

Hui, T.-W. *et al.* (2018). Liteflownet: A lightweight convolutional neural network for optical flow estimation. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 8981–8989, Salt Lake City.

Ilg, E. *et al.* (2017). Flownet 2.0: Evolution of optical flow estimation with deep networks. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1647–1655, Honolulu.

Kappe, C. P. *et al.* (2016). Reconstruction and visualization of coordinated 3D cell migration based on optical flow. *IEEE Trans. Visualization and Computer Graphics*, **22**(1), 995–1004.

Liu, K. *et al.* (2014). Optical flow guided cell segmentation and tracking in developing tissue. In *Proc. of Int. Symposium on Biomedical Imaging*, pages 298–301, Beijing.

Meister, S. *et al.* (2018). Unflow: Unsupervised learning of optical flow with a bidirectional census loss. In *Proc. of Advancement of Artificial Intelligence*, New Orleans.

Muja, M. and Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. In *Proc. of Computer Vision Theory and Applications*, pages 331–340, Lisbon.

Ourselin, S. *et al.* (2002). Robust registration of multi-modal images: Towards real-time clinical applications. In *Proc. of Medical Image Computing and Computer-Assisted Intervention*, pages 140–147, Tokyo.

Revaud, J. *et al.* (2015). EpicFlow: Edge-preserving interpolation of correspondences for optical flow. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Boston.

Sun, D. *et al.* (2010). Secrets of optical flow estimation and their principles. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2432–2439, San Francisco.

Tektonidis, M. and Rohr, K. (2017). Diffeomorphic multi-frame non-rigid registration of cell nuclei in 2D and 3D live cell images. *IEEE Trans. Image Processing*, **26**(3), 1405–1417.

Vogel, C. *et al.* (2013). An evaluation of data costs for optical flow. In *Proc. of Pattern Recognition*, pages 343–353, Heidelberg.

Xu, J. *et al.* (2017). Accurate optical flow via direct cost volume processing. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 5807–5815, Honolulu.

Zabih, R. and Woodfill, J. (1994). Non-parametric local transforms for computing visual correspondence. In *Proc. of European Conference on Computer Vision*, pages 151–158, Stockholm.

Zhu, X. and Milanfar, P. (2010). A no-reference image content metric and its application to denoising. In *Proc. ICIP*, pages 1145–1148, Hong Kong.

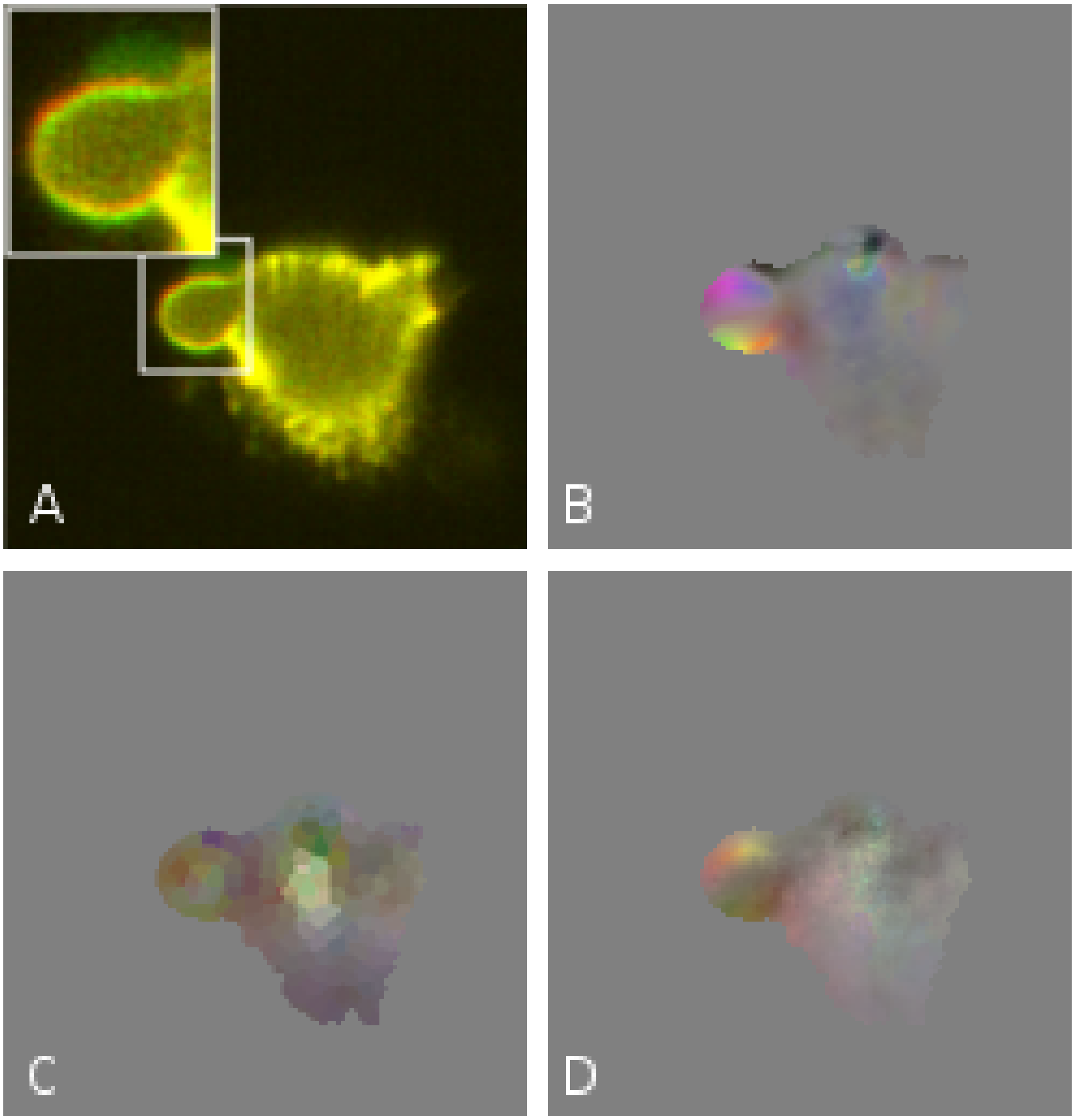

Figure 7: (A) Overlap of slices of sequence 2: source in red, target in green. 3DHSV motion map of (B) HS3D, (C) Amat's method, and (D) our variational method. Color code corresponds to Fig. 6 (left).

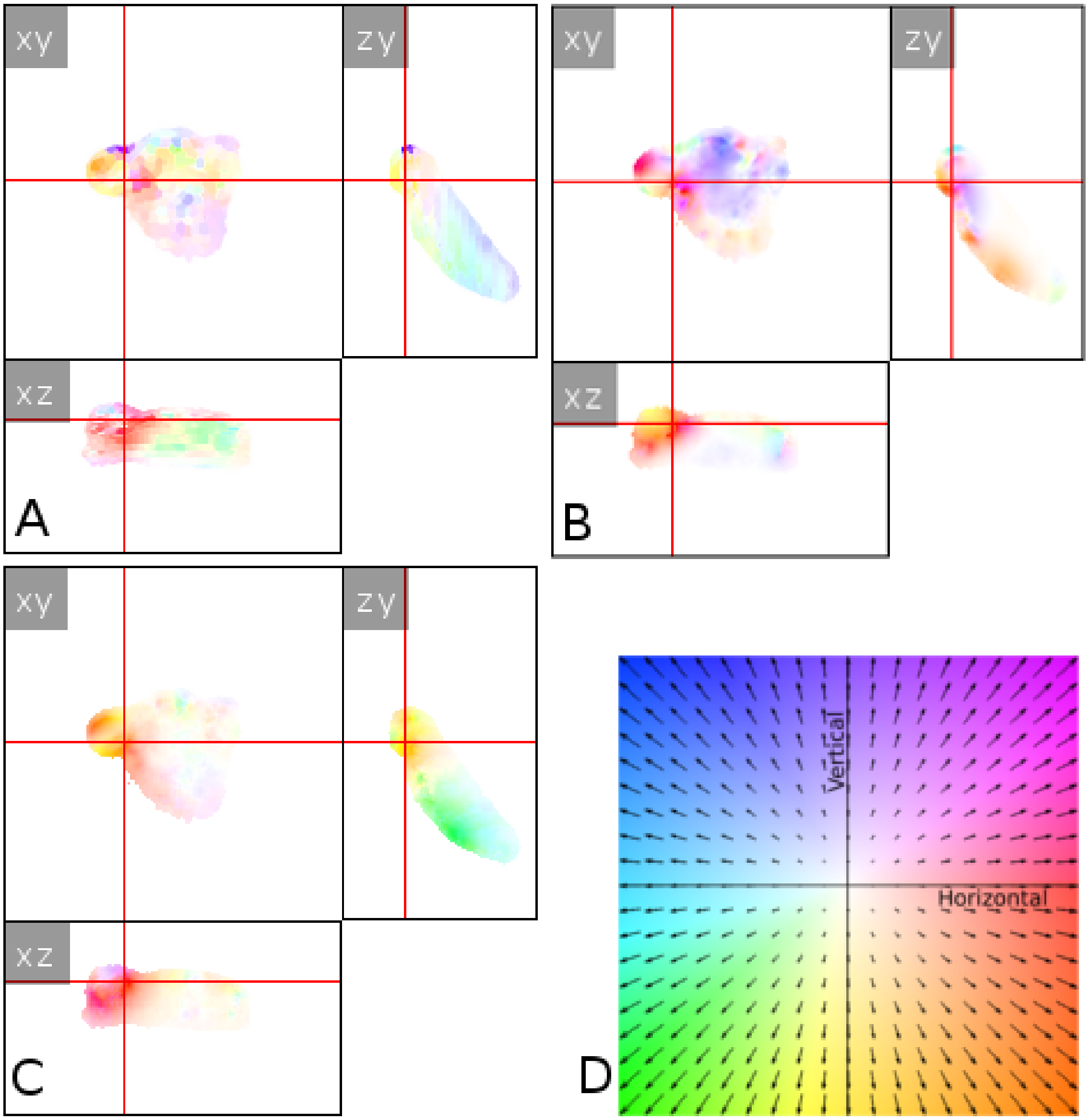

Figure 8: 3PHS map of (A) Amat's method, (B) HS3D, (C) our variational method (Eq. 2), and (D) 3PHS color code.

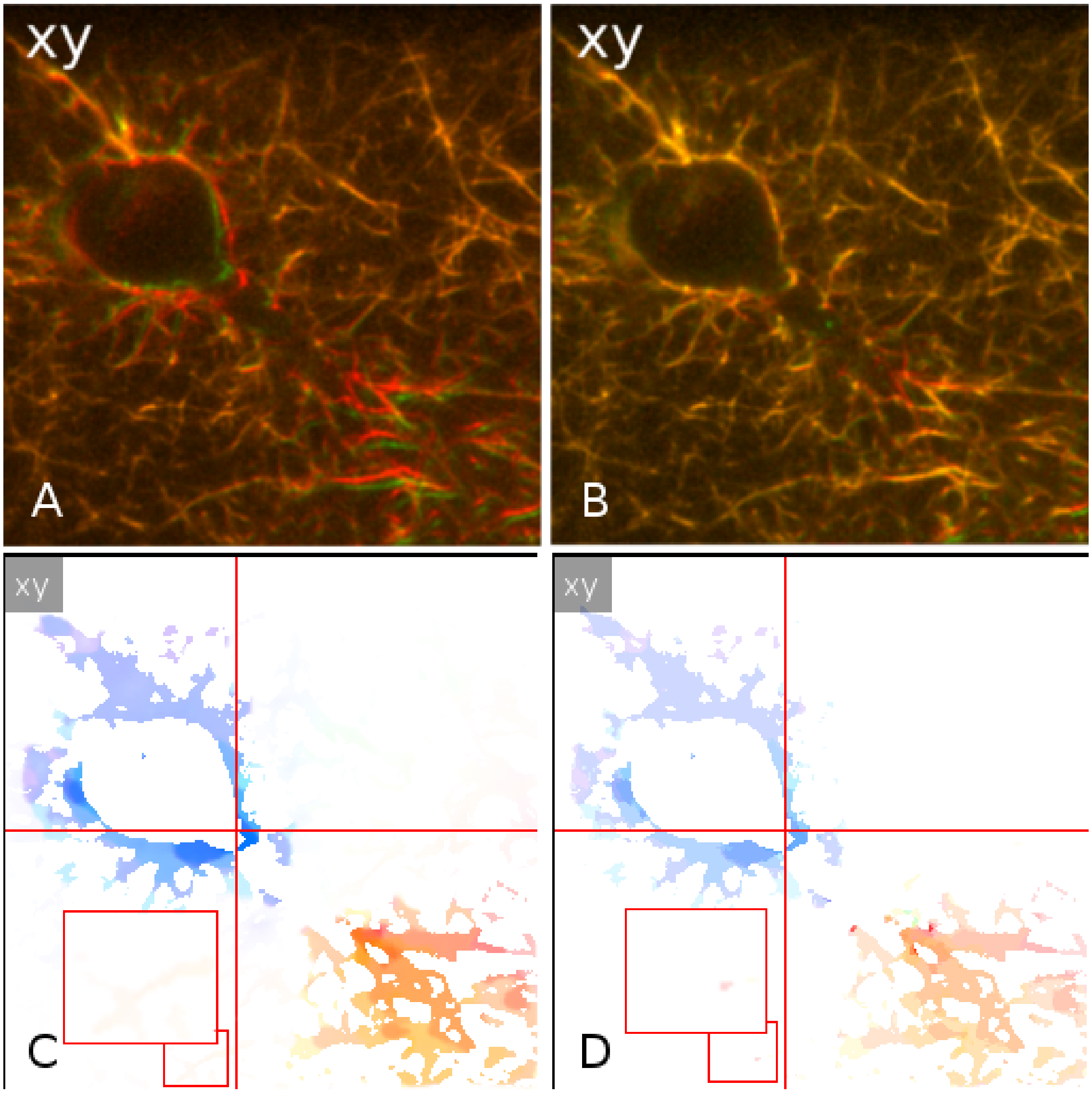

Figure 9: (A) Overlap of slices of sequence 2 (collagen): source in red, target in green (showing only xy-plane). (B) Source in red, reconstructed source after 3D PatchMatch and proposed variational refinement in green; the overlapped voxels appear in yellow. (C) 3PHS motion map after 3D PatchMatch and proposed variational refinement (best viewed in electronic form). The red rectangle highlights the small motion captured by the combination. (D) 3PHS motion map using 3D PatchMatch alone. The small motion is not captured.