Performance of Low-Level Motion Estimation Methods for Confocal Microscopy of Plant Cells *in vivo*

T. Roberts, S. McKenna
School of Computing, University of Dundee, Scotland
troberts@computing.dundee.ac.uk

N. Wuyts, T. Valentine, A. Bengough
Scottish Crop Research Institute,
Dundee, Scotland

Abstract

The performance of various low-level motion estimation methods applied to fluorescence labelled growing cellular structures imaged using confocal laser scanning microscopy is investigated. This is a challenging and unusual domain for motion estimation methods. A selection of methods are discussed that can be contrasted in terms of how much spatial or temporal contextual information is used. The Lucas Kanade feature tracker, a spatially and temporally localised method, was, as one would expect, accurate around resolvable structure. It was not able to track the smaller, repetitive cell structure in the root tip and was somewhat prone to identifying spurious features. This approach is improved by developing a full multi-frame, robust, Bayesian method, and it is demonstrated that by using extra frames with motion constraints reduces such errors. Next, spatially global methods are discussed, including robust variational smoothing and Markov Random Field (MRF) modelling. A key conclusion that is drawn from investigation of these methods is that generic low-level (robust) smoothing functions do not provide good results in this application and that this is probably due to the large regions with little stable structure. Furthermore, contrary to recently reported successes, graph cuts and loopy belief propagation for MAP estimation of the MRF labels provided often poor and inconsistent estimates. The results suggest the need for greater emphasis on temporal smoothing for generic low-level motion estimation tools and more task specific, spatial constraints, perhaps in the form of high level models in order to accurately recover motion from such data. Finally, the form of the estimated growth is briefly discussed and related to contemporary biological models. We hope that this paper will assist non-specialists in applying state-of-the-art methods to this form of data.

1 Introduction and Background

Confocal laser scanning microscopy (CLSM) combined with fluorescence labelling can provide rich, high spatiotemporal resolution data on the structure and dynamics of live biological specimens. The analysis of such data has often been qualitative in nature. In this workshop paper we discuss and evaluate various 2D low-level motion estimation methods with a view to developing a flexible, quantitative, automatic tool for confocal data. The initial focus here is on the estimation of *Arabidopsis thaliana* root growth, a model system in plant sciences. Success for this application would allow the large quantities of confocal data to be analysed thus improving our understanding of growth (cell production and expansion) and environmental and genetic factors on its regulation. This analysis is also interesting due to the unusual properties of the confocal data, as discussed below.

CLSM allows optical sections of living plant tissues to be acquired in a non-invasive manner [7, 9, 17]. It greatly reduces blurring from out-of-focus light and thereby provides higher resolution images than conventional 'wide-field' microscopy. 3D datasets can be produced by combining these sections and measurements can be repeated over time to investigate dynamics. When combined with targeted, fluorescent markers such as GFP, specific cell structures can be isolated, providing rich data for studying the morphology and physiology of living plant systems, e.g. [8, 9]. In this paper, *A. thaliana* plants expressing a single fluorescence marker targeted to the plasma membrane (LTI-eGFP) were grown *in vitro* under standard conditions [13]. Images of their roots were acquired on a Leica TCS SP1 confocal microscope using a $10\times$ (NA 0.3) or $20\times$ (NA 0.5) objective lens with the $488nm$ laser line of a $20mW$ Argon laser. Images were captured as either time series or as a series of z-stacks at high resolution (1024×1024 pixels). An example of an image produced under such conditions is shown in Figure 1. A typical quantity of interest to be derived from a sequence of such images might be the rate of growth of individual cells and its distribution throughout the root.

The data produced by these experiments have some unusual properties that make low-level motion estimation challenging. The most significant properties are: (a) the motion is a combination of non-rigid growth and global translation, with cells forming a complex articulated structure with small, but important, differences between cells, (b) using fluorescent markers results in sparse images with

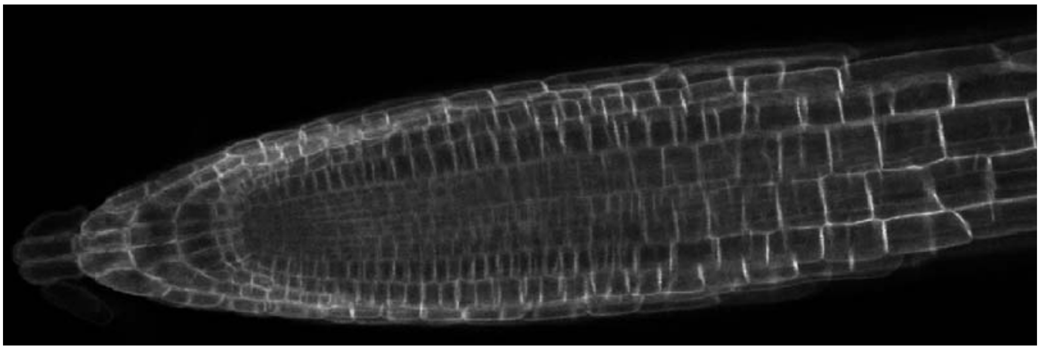

Figure 1: A confocal scan of live, LTI-eGFP labelled, *Arabidopsis thaliana root.*

little stable texture in many regions, (c) structure at other depths and fluorescence in other cell regions can have different motion types, (d) biological variability and confocal imaging lead to a non-Gaussian, non-isotropic noise distribution as illustrated in Figure 2, (e) large inter-frame motion (around 8 pixels) can be present, especially when a depth stack is being acquired, (f) small motions in depth can be present (the roots were setup to reduce out of plane motion, by growing them directly on microscope slides in thin layers of growth medium) and (g) cells can divide. Cell division is not considered explicitly in this paper.

Previous work on motion estimation from confocal data has focussed mainly on local feature tracking, for example in the form of particle image velocimetry (PIV) algorithms. The authors are not aware of any detailed investigations into root growth using confocal data. However, there have been investigations into estimating growth using other (external) imaging modalities. van der Weele *et al.* [22] applied low-level motion estimation methods to non-fluorescence microscope images of the external surface of *A. thaliana* roots (without the application of graphite particles). Walter *et al.* [23] tracked maize root surfaces using infra-red illumination. Barron and Liptay [2] used differential flow from multi-view optical imaging and near infra-red imaging to recover the growth of maize seedlings. We decided that the different properties and advantages of confocal microscopy warranted this further investigation.

2 Motion Estimation Methods

In our discussion of motion estimation methods it will be useful to contrast methods in terms of how much spatial and temporal information is used. First, a local feature tracking method is discussed which, as one would expect, is found to be accurate around resolvable structure and provides the baseline for performance evaluation. This region tracking approach is improved by developing a full multi-frame, robust likelihood, Bayesian method and the improvements derived from such an approach are illustrated. Next, spatially global methods are discussed, including robust variational smoothing and Markov Random Field (MRF) modelling. The key problem with MRF models is solving the resulting inference problem, and a selection of techniques are discussed for this purpose, namely graph cuts, loopy belief propagation and multi-scale MCMC. The performance of all methods is illustrated on the same dataset, shown in Figure 3. This dataset contains a mixture of root growth and global translation. Results are presented after subjectively tuning the parameters to their reported optimal values (automatic parameter learning being non-trivial). Due to the lack of ground truth we compare the methods directly. *For these methods, the most important conclusions, such as the need for more sophisticated spatial constraints, can be drawn without a high accuracy comparison.* Tests in more controlled conditions, e.g. translating the data, have been performed for the local methods. In this application fast implementations are not important but the method must be feasible for deployment to standard PC workstations.

2.1 Lucas-Kanade-Tomasi Feature Tracker

One straightforward and popular approach to motion estimation is local region tracking as exemplified by the Lucas-Kanade-Tomasi tracker [1, 16]. As this method, in the form investigated here, uses only spatiotemporally localised data, the method is quite simple, popular and general and we use it as a baseline for comparing the methods. Here we only consider recovering locally rigid translations. A more realistic model would also account for expansion and deformation, and would be important for longer-term tracking. Results from Birchfield's publicly available implementation of this method (KLT $v1.3.2$) are shown in Figure 3. The default energy functions and parameters were used (linearised quadratic) with 200 feature tracks requested. Performance deteriorated dramatically when the region window size was below 11 pixels (for which results are shown). It is clear that the results are not accurate in the root tip region where confusion between the small, similar cells is considerable. *A*

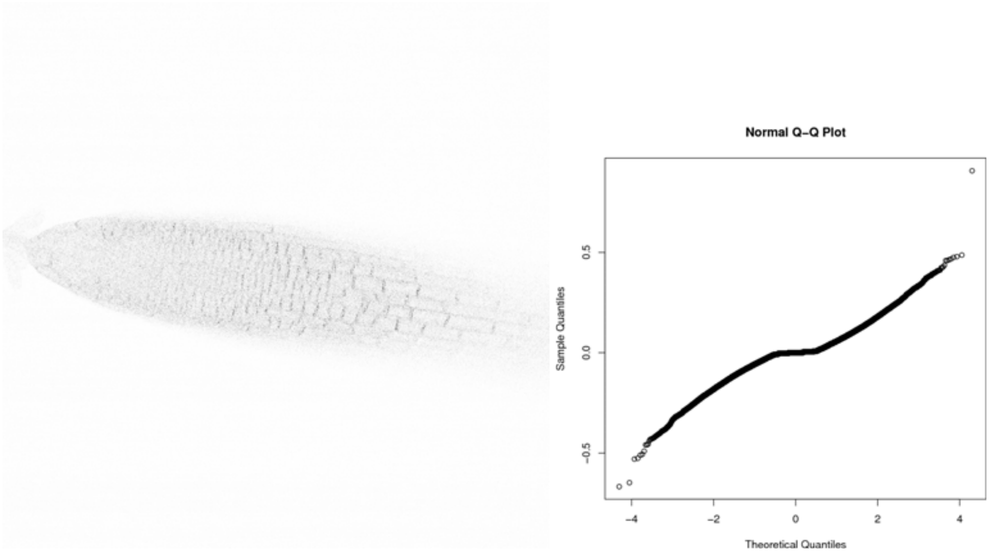

Figure 2: Left: The (scaled) absolute difference difference between two scans taken in rapid succession so that growth is negligible. Right: A QQ-normal plot showing the non-Gaussian nature of the noise (formed from the difference between two close scans). The form of the curve suggests three components: approx. Gaussian background noise (central linear section, corresponding to the majority of the pixels), approx. Gaussian 'biological noise' (larger linear section) and outliers.

more global method will be necessary to remove such ambiguities. Furthermore, around 10% of the feature points were assigned to intra-cellular regions where the appearance is more variable (as compared to cell junctions) and the associated tracks are more often incorrect. Using a method other than eigenvalue inspection, that is more application specific, could reduce such errors.

2.2 Multi-frame, Robust Bayesian Method

In this section we discuss three developments of the above region tracking approach, namely a more constrained temporal model, a robust likelihood and a better technique for estimating uncertainty. The method is similar to that proposed by Singh [19] but uses a better estimate of uncertainty and a robust likelihood. Since the dynamics of the root, over the small timescales investigated here, are constrained and simple (compared to the form of spatial constraints) we can easily make use of multiple frames to reduce gross errors and improve accuracy. In particular, a constant velocity model was fit through a series of frames (typically 3 to 5) rather than propagated forward as in the KLT implementation. A robust M-estimator of the form $r^2/(s^2 + r^2)$ where r denotes pixel residuals and $s = 0.02$ was employed. Due to the lack of computational constraints and the need for accurate estimates of uncertainty this method was developed in a different manner to the Lucas Kanade feature tracker. In particular, the posterior distribution over local motion was found by exhaustive evaluation at all foreground points (which are easily found in this application using an intensity threshold, resulting in a feasible runtime for 3-5 frames). Accurately estimating and reporting motion uncertainties is very important in applications requiring accurate, automatic large scale analysis. Note that only estimating motion at points of high intensity (partially) justifies the use of a single non-isotropic noise function, the effects of which are not investigated in more detail due to lack of space. The motion estimate is derived from an expectation over the dominant (MAP) posterior mode. The results for this method using 2 or 3 frames with an 11 pixel circular window are shown in Figure 4. It can be clearly seen that using a constrained motion model over 3 frames reduces the gross errors present in the Lucas Kanade results. Notice the errors due to aperture problems on cell walls. The certainty is correctly estimated to be low where such errors occur.

2.3 Black's Robust Smoothing Method

Using global spatial smoothing to remove local ambiguities and improve accuracy has a long history in optical flow estimation, e.g. Horn and Schunk [10]. In this section we describe and evaluate one such method due to Black and Anandan [3] based upon robust spatial smoothing that preserves motion discontinuities. In this method the equations resulting from the minimisation of both a robust function on the image residuals and a robust spatial smoothing function are solved using an iterative, multi-scale, graduated non-convexity algorithm, a key idea in which is approximation of the non-convex robust functions with convex functions (see [3] for more details). The results for this method are

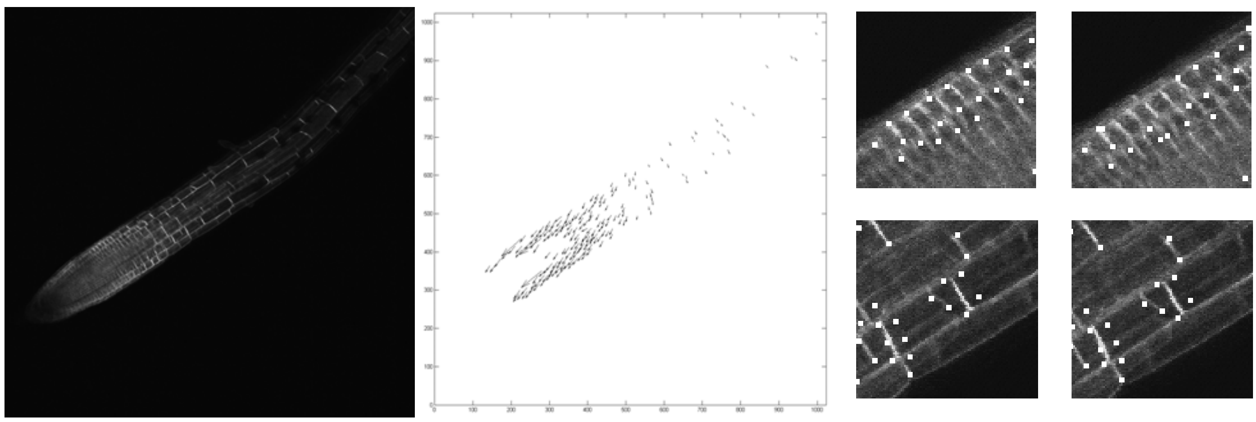

Figure 3: Left: The first frame in the sequence (used for all results). Middle: A (scaled) quiver plot of the sparse motion estimate from the Lucas Kanade Tomasi feature tracker. Right: Two zoomed views illustrating some typical errors. The top row shows gross errors in the root tip area where cells are small and similar. The bottom row illustrates the detection of features on less stable structures.

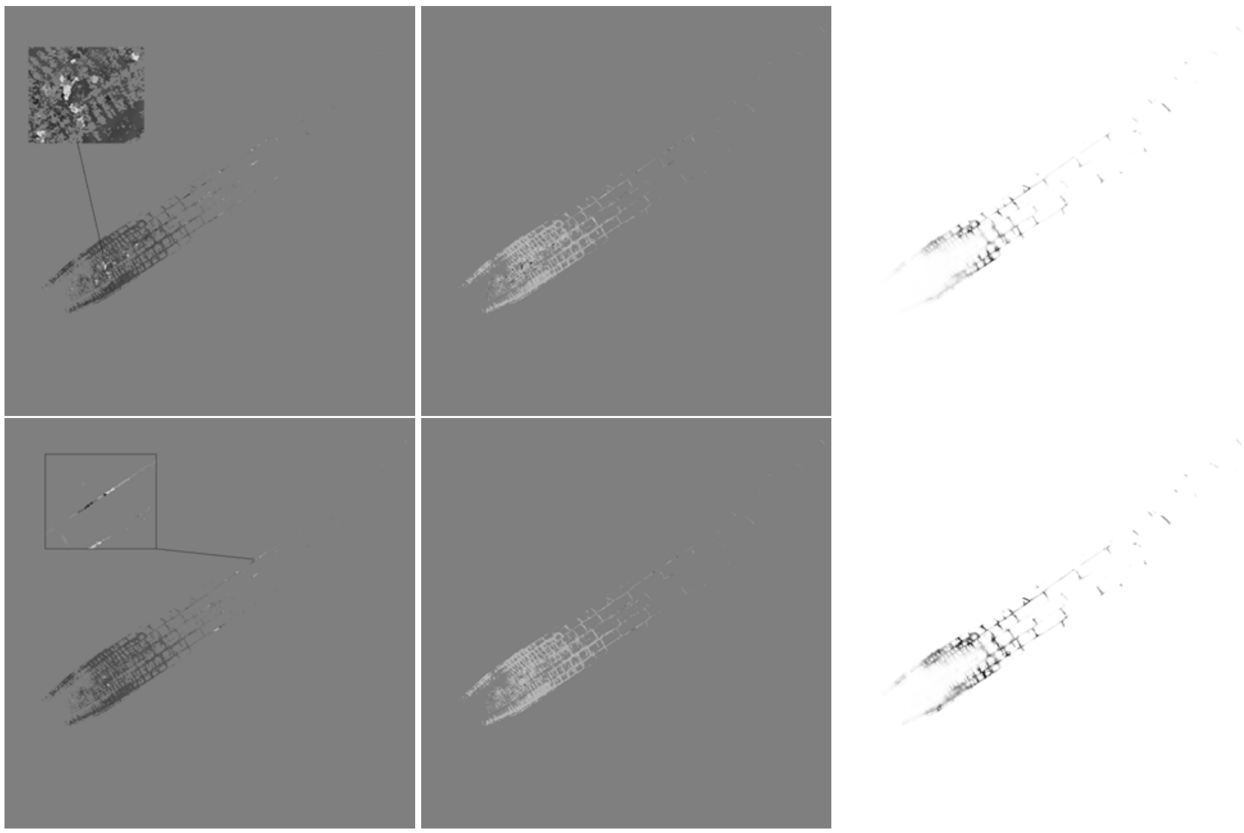

Figure 4: Results from the multi-frame method. The top row shows results using two frames and the bottom using three. The left and middle columns show the estimated horizontal and vertical flow respectively, at foreground points (i.e. those with high intensity). Mid-grey values corresponds to zero motion component, white to right/down and black to left/up with a range of 8 in all directions. Points where velocity is not estimated are shown in grey. The right column shows the uncertainty, with black corresponding to certainty. It can be seen that using even a single extra frame with a constrained motion model can significantly increase tracking certainties. Gross tracking errors are significantly reduced.

show in Figure 5. The third smoothing level is shown, with parameters close to their default and recommended values (l1=10,l2=1,S1=10.0,S2=1.0,s1=5.0,s2=0.05). The method does not provide estimates of spatial uncertainty.

The performance of this method with these parameters agrees, as one might expect, around highly informative data (e.g. the region 1/3 up from the root tip) with the two previous spatially localised methods. *However, in regions of ambiguity, such as the root tip, the low-level spatial constraints fail to provide any significant advantages. Indeed the global smoothing reduces accuracy in some regions due to localised over-smoothing.* This can be clearly seen by the incorrect under-estimated motion on the root tip. Notice that there is no visual evidence for a discontinuity at the root agar boundary so the method, without further prior knowledge, could not be expected to identify this accurately.

2.4 MRF Models and Inference Schemes

Markov Random Fields (MRF) are another approach that can be used to elegantly model many low-level vision tasks such as motion estimation, stereo disparity and segmentation that require global spatiotemporal context. They share similarities with the variational minimisation approach discussed above, but emphasise a discrete and probabilistic approach. MRFs are discussed in detail in Li's book [15]. Our plan when investigating MRFs was to develop the global smoothing approach by doing joint motion estimation and segmentation, an approach that has had some success in other applications, e.g. [24]. This would be difficult with the variational approach. However, here we report our initial findings using MRFs for spatial smoothing and the performance of associated inference schemes. For this purpose we tested a range of data and smoothing energy functions. In particular, we tested the following data energy functions: quadratic, r^2, McClure Geman, $r^2/(s^2+r^2)$, and truncated quadratic, r^2 or c^2 if $r > c$ (with c constant). We tested the following smoothing functions: linear, $|r|$, truncated linear and quadratic, and McClure Geman with a range of different smoothing strengths. Here we are interested in inference in MRF models with these parameters fixed (MAP optimisation and computing expectations). As with variational methods the key problem in employing MRFs, which model the overall probability of a configuration using a cyclic network of local functions, is solving the resulting global, high dimensional, inference problem. However, there have been many recent papers reporting robust, high performance inference for MRF models, in particular graph cuts and loopy belief propagation [4, 6, 12, 21]. We evaluate these methods based upon the publicly available Middlebury MRF energy minimisation code ($v1.6$) [20], which contains algorithms contributed by various authors. We used discrete (i.e. not sub-pixel) velocity labels, and assume this will not affect the results appreciably. The results do not employ spatial weighting of energy terms as the correct form of weighting is not clear for this application.

Reliably fitting MRFs to our data using these algorithms proved difficult. The algorithms were quite sensitive to the parameters used, often providing clearly inaccurate results that were often inconsistent between algorithms. Partially successful results (using the regularised quadratic image noise and linear smoothing function) for expansion and swap graph cuts algorithms are shown in Figure 7. Loopy belief propagation (BP), in the form of max-product (Middlebury and in-house implementation) and min-sum (in-house implementation) also produced poor results and had prohibitively high memory requirements (a recognised problem with BP that has drawn some interest recently, e.g. [6, 11]). We note that all methods were able to produce accurate results when applied to synthetic transformations (i.e. root images with simple translation) with a wide range of energies/parameters. The final algorithm in the Middlebury suite was iterated conditional modes (ICM) and, as would be expected, this gave very poor results due to getting trapped in local minima early on (with energy values much higher than graph cuts). One disadvantage of these MAP methods is that, although they use probabilistic models, they do not produce spatial estimates of uncertainty (although marginals could be estimated using min-sum BP). It is important when discussing MRFs to mention the problem of automatic parameter learning, for example ML estimation of θ in $p(x;\theta) = Z^{-1}(\theta)e^{-E(x;\theta)}$ given a collection of IID datasets. This is complicated due to the global nature of the partition function, Z, and the large amount of data. Some recent progress, e.g. Hinton's contrastive divergence method [5], may make this more feasible, but sensitivity to the smoothing parameters and poor or inconsistent inference suggest this will be difficult on these data.

We conclude the discussion of MRFs by briefly mentioning a novel multi-scale MCMC technique we have begun to develop to overcome these problems and provide accurate estimates of uncertainty. In this method the proposal functions are blocks that gradually reduce in size. Results, with identical model parameters to the graph cuts, are shown in the bottom row of Figure 7 and illustrate that such an approach is feasible and reconfirm the effect of low-level spatial smoothing. Early results suggest such a multi-scale approach can provide quite reliable inference, although computational cost is high. Questions on the effect of transition adaptivity need to be answered in future work.

3 Growth Estimation

The ultimate aim of the work described here is to deliver an open-source workbench application capable of accurate analysis of large amounts of confocal data. An example automated result found from the projection of the multi-frame

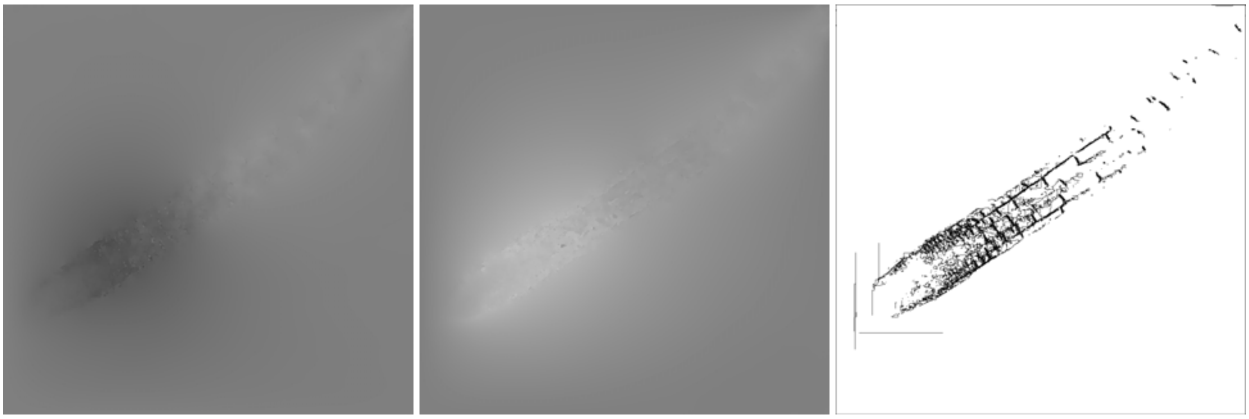

Figure 5: Results from Black and Anandan's robust smoothing method. The left and middle images show horizontal and vertical components of flow scaled to be comparable with the other results. The method is accurate around informative structures but low-level spatial smoothing provides little advantage, sometimes even reducing accuracy due to localised over-smoothing, compared with local methods. The right image shows violations of the spatial coherence model.

results to a spline fit to the central axis of the root is shown in Figure 6. The form of the result agrees with previous investigations e.g. [22]. Particular cell regions can also be identified and tracked allowing the study of small changes in cell expansion rate with a temporal resolution of minutes and so represents an exciting tool for plant physiology studies.

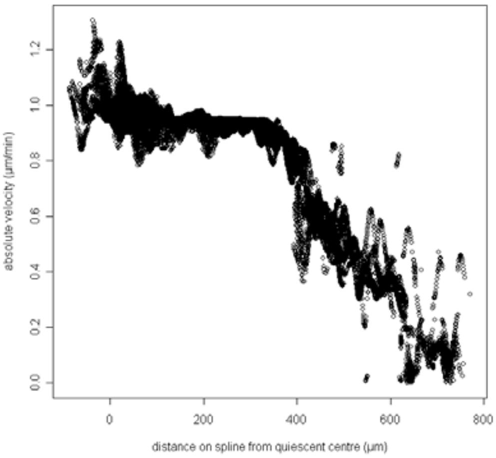

Figure 6: Velocity as a function of distance along the root from the tip.

4 Conclusions and Future Work

Local region tracking is a popular, flexible motion estimation approach, with relatively low computational requirements, that has been applied to many biological problems. However, we emphasis the need for good estimates of uncertainty since these methods are susceptible to gross errors due to local ambiguity and large motions. The focus here was investigation of more sophisticated, generic low-level methods to reduce such errors. We demonstrated that by employing simple temporal constraints a significant degree of local ambiguity was removed. *Unfortunately, using global low-level generic spatial constraints did not conclusively increase the accuracy for this type of data.* The large motions, lack of evidence in many regions, secondary motions and significant non-Gaussian noise clearly motivate the need for more sophisticated spatial constraints in order to produce more accurate motion estimates. There have been criticisms elsewhere of simple MRF models, e.g. those involving pairwise smoothing functions, for other applications, leading to proposals such as the product-of-experts method for modelling, and learning, higher order clique potentials, e.g. [14]. *Furthermore, contrary to recent positive investigations, we found inference in MRF spatial smoothing models on these data, using graph cuts and loopy belief propagation to be unreliable.*

There are many possible directions for future work. Originally, we had planned to develop an MRF model for joint segmentation and motion estimation, allowing for the incorporation of non-isotropic and more accurate smoothing and observation functions. Another direction would be to learn higher order MRF potentials. However, the issue of reliable inference and high resource requirements would be compounded. Given the current performance of these low-level methods it is likely that we will be placing more emphasis on developing task specific, higher level, spatial constraints, such as those discussed in [18], that are able to incorporate stronger prior knowledge on cell structure and kinematics. One novel challenge of building and inferring such high level models is varying topology, both between datasets and over time.

Acknowledgements We thank BBSRC and SEERAD for funding this work, M. Black, S. Birchfield and those who contributed to the Middlebury MRF suite, for their source code, J. Haseloff for providing LTI lines and F. Bransby for helpful discussions.

[1] S. Baker and I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *IJCV*, 56(3):221–255, 2004. ISSN 0920-5691.

[2] J.L. Barron and L. Liptay. Measuring 3D Plant Growth Using Optical Flow. *BioImaging*, 5:82–86, 1997.

[3] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow-fields. *CVIU*, 63(1):75–104, January 1996.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. In *ICCV*, pages 377–384, 1999.

[5] M. A. Carreira-Perpiñán and Geoffrey E. Hinton. On contrastive divergence learning. In *Workshop on Artificial Intelligence and Statistics*, pages 33–40, 2005.

[6] P.F. Felzenszwalb and D.P. Huttenlocher. Efficient belief propagation for early vision. *IJCV*, 70(1):261–268, October 2006.

[7] S. Gilroy. Fluorescence microscopy of living plant cells. *Annual Review Plant Physiol Plant Mol Biol*, 48:165–190, Jun 1997.

[8] M. R. Hanson and R. H. Köhler. GFP imaging: methodology and application to investigate cellular compartmentation in plants. *J Exp Bot*, 52(356):529–539, Apr 2001.

[9] P Hepler and Gunning B. Confocal fluorescence microscopy of plant cells. *Protoplasma*, 201:121–157, 1998.

[10] B.K.P. Horn and B.G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[11] P. Kohli and P.H.S. Torr. Efficiently solving dynamic markov random fields using graph cuts. In *ICCV*, volume II, pages 922–929, 2005.

[12] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? In *ECCV*, pages 65–81, 2002.

[13] S. Kurup, J. Runions, U. Köhler, L. Laplaze, S. Hodge, and J. Haseloff. Marking cell lineages in living tissues. *Plant*, 42(3):444–453, May 2005.

[14] X. Lan, S. Roth, D. Huttenlocher, and M.J. Black. Efficient belief propagation with learned higher-order markov random fields. In *ECCV*, 2006.

[15] S. Li. *Markov Random Field Modeling in Computer Vision*. Springer Verlag, 1995. ISBN 0-387-70145-1.

[16] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI*, pages 674–679, 1981.

[17] J.B. Pawley, editor. *Handbook of Biological Confocal Microscopy*. Springer, 4th edition, 2006.

[18] T. J. Roberts, S.J. McKenna, J. Hans, T. A. Valentine, and A. G. Bengough. Part-based multi-frame registration for estimation of the growth of cellular networks in plant roots. In *ICPR*, August 2006.

[19] A. Singh. *Optic Flow Computation: A Unified Perspective*. IEEE Computer Society Press, California, 1991.

[20] R. Szeliski, R Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. In *ECCV*, volume 2, pages 19–26, Graz, Austria, May 2006.

[21] M. Tappen and W. Freeman. Comparison of graph cuts with belief propagation for stereo using identical MRF parameters. In *ICCV*, volume 2, pages 900–908, October 2003.

[22] C.M. van der Weele, H.S. Jiang, K.K. Palaniappan, V.B Ivanov, K. Palaniappan, and T.I Baskin. A new algorithm for computational image analysis of deformable motion at high spatial and temporal resolution applied to root growth. *Plant Physiol*, 132(3): 1138–1148, Jul 2003.

[23] A. Walter, H. Spies, S. Terjung, R. Küsters, N. Kirchgessner, and U. Schurr. Spatio-temporal dynamics of expansion growth in roots: automatic quantification of diurnal course and temperature response by digital image sequence processing. *J Exp Bot*, 53 (369):689–698, Apr 2002.

[24] Paul P Wyatt and J. Alison Noble. MAP MRF joint segmentation and registration of medical images. *Medical Image Analysis*, 7(4):539–552, Dec 2003.

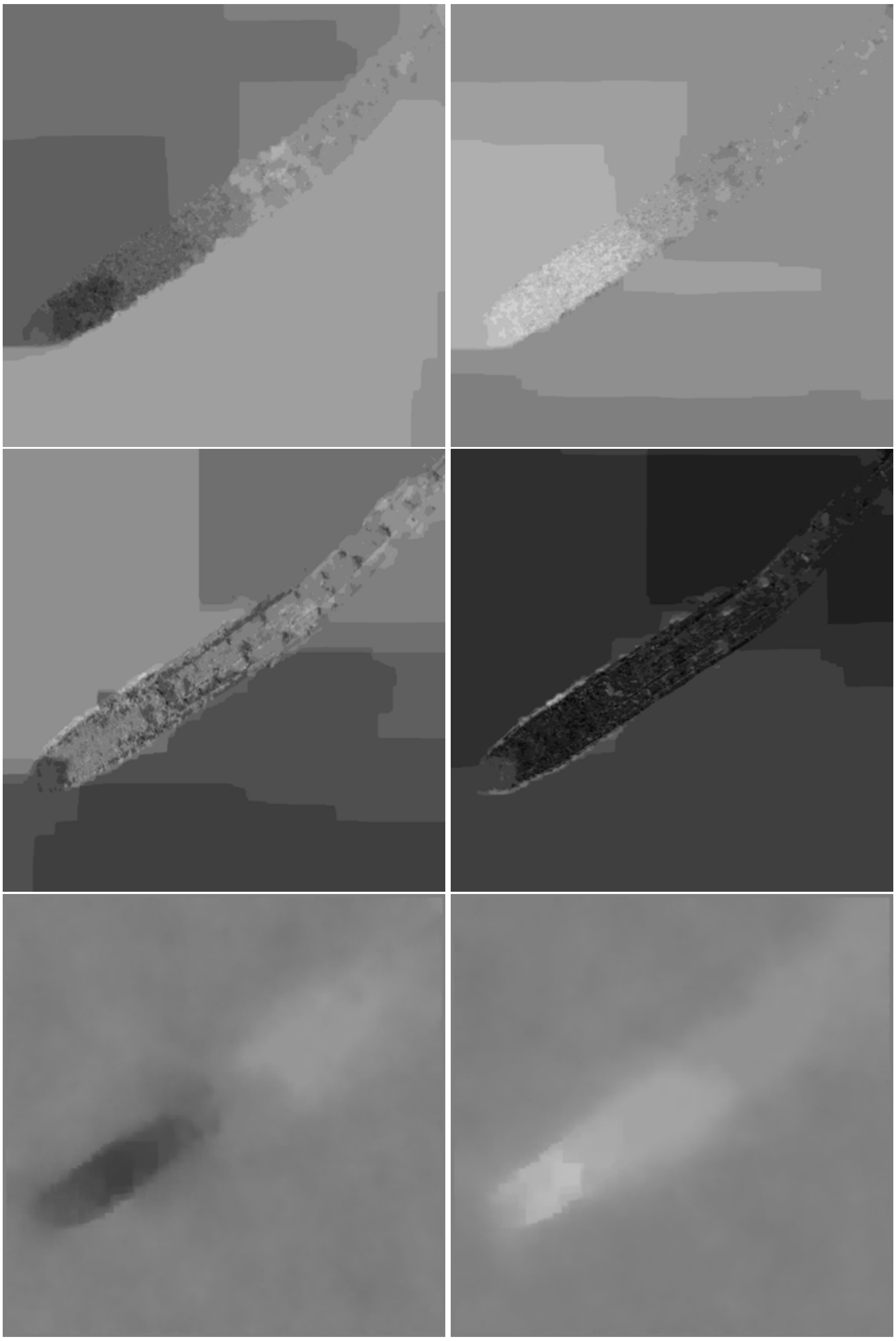

Figure 7: Results from the MRF model with various inference schemes. The first row shows the expansion graph cuts algorithm, the second row the swap graph cuts algorithm. These algorithms proved sensitive to the choice of energy functions and parameters and often, as in this case, provided inconsistent results. From the top row, which shows the best result achieved after extensive manual parameter tuning, it can be seen that lower velocity estimates in the root tip region are incorrect however the result are better that Blacks method in this region which suggests that the linear smoothing function is more appropriate for this data. However, as with the other method, smoothing elsewhere provides inaccurate estimates. Note that the sharp changes are due to the discrete encoding of velocities and are not an inherent problem with the method. The bottom row shows expectations (rather than the MAP labelling) from an MCMC algorithm. Notice that even after long simulation times some artifacts from the blockwise transition function remain.