

PHASE-BASED IMAGE MOTION ESTIMATION AND REGISTRATION

Magnus Hemmendorff
Mats T Andersson
Hans Knutsson

Department of Electrical Engineering
Linköping University
S-581 83 Linköping
SWEDEN
ma-hem@isy.liu.se, <http://www.isy.liu.se/cvl/>

ABSTRACT

Conventional gradient methods (optical flow), for motion estimation assume intensity conservation between frames. This assumption is often violated in real applications. The remedy is a novel method that computes constraints on the local motion. These constraint are given on the same form as in conventional methods. Thus, it can directly substitute the gradient method in most applications. Experiments indicate a superior accuracy, even on synthetic images where the intensity conservation assumption is valid. The conventional gradient methods seem obsolete.

1. INTRODUCTION

Gradient methods[8], often referred to as *optical flow*, are widely used in motion estimation. The method assumes that image intensity is conserved along a spatiotemporal path along the motion. This assumption gives constraints on the local motion $c_x v_x + c_y v_y + c_t = 0$, where (c_x, c_y, c_t) is the spatiotemporal gradient of image intensities and (v_x, v_y) is the local motion. Local constraints can be integrated over a region or the entire image to estimate any motion, e.g. translation or affine.

Computing the spatiotemporal gradient does not require more than two frames. Large motions can be accurately estimated using multiple scales and iterative refinement, where images are warped[10, 11].

The objective of this article is to introduce a novel method to estimate (c_x, c_y, c_t) . It does not assume strict intensity conservation, and accuracy is good even when there are intensity variations between frames.

2. PHASE BASED QUADRATURE FILTER METHOD

Using quadrature filters phase is a relatively common approach in stereo algorithms[12, 5]. The idea of using phase for motion estimation has previously been investigated by some researchers [3, 1, 4], but to our knowledge, nobody has tried this approach, which extends the accurate stereo algorithms to track motions. Our method is basically a gradient-based method with nonlinear preprocessing of the images. To improve accuracy, a confidence measure has been added.

Definition 2.1 *A filter is a quadrature filter[6] if its Fourier transform, $F(\mathbf{u})$, has zero amplitude on one side of a hyperplane through the origin, i.e. there is a direction $\hat{\mathbf{n}}$ such that*

$$F(\mathbf{u}) = 0 \quad \forall \hat{\mathbf{n}}^T \mathbf{u} \leq 0 \quad (1)$$

Quadrature filter responses are closely related to analytic signals. Note that quadrature filters must be complex in the spatial domain. We only use filters that are real in the Fourier domain.

2.1. Motion Constraint Estimation

A number of quadrature filters are applied in parallel on each of the image frames, producing the same number of filter responses. The quadrature filters are tuned in different directions and frequency bands to split dissimilar features into different filter responses, so that they do not interfere in the motion estimation. The quadrature filters also suppress undesired features like DC value and high frequencies. Unlike the conventional gradient method, our method is not sensitive to low pass variations in image intensity, that are frequent in medical X-ray images, or real world images where shadows and illumination vary.

This work was supported by SECTRA and The Swedish National Board for Industrial and Technical Development (NUTEK)

For each of the filter responses, we compute constraints on the local motion. The first step is to convolve both of the frames with quadrature filters

$$q_{A,j}(\mathbf{x}) = (f_j * I_A)(\mathbf{x}) \quad \text{and} \quad q_{B,j}(\mathbf{x}) = (f_j * I_B)(\mathbf{x}) \quad (2)$$

where $f_j(\mathbf{x})$ is a quadrature filter and $I_A(\mathbf{x})$ and $I_B(\mathbf{x})$ are image intensities of the two frames respectively. The phase is defined as

$$\theta_{A,j}(\mathbf{x}) = \arg q_{A,j}(\mathbf{x}) \quad \text{and} \quad \theta_{B,j}(\mathbf{x}) = \arg q_{B,j}(\mathbf{x}) \quad (3)$$

In all ensuing computations, we must remember that phase is always modulo 2π , but for readability we omit this in our formulas and notations. In most image points, the filter responses are strongly dominated by one frequency, which makes the phase nearly linear in a local neighbourhood. When the phase is linear, it can be represented by its value and gradient. Thus, a gradient method applied on the phase will be very accurate. Of course, the phase is not always linear in a local neighbourhood, but that can be detected, and reflected by a confidence measure.

For each point in the image, and for each quadrature filter response, a constraint on the local motion is computed. To simplify notations, we drop the index, j , of the quadrature filter.

$$\mathbf{c} = \begin{pmatrix} c_x \\ c_y \\ c_t \end{pmatrix} = C \begin{pmatrix} \frac{1}{2} \frac{\partial}{\partial x} (\theta_B + \theta_A) \\ \frac{1}{2} \frac{\partial}{\partial y} (\theta_B + \theta_A) \\ \theta_B - \theta_A \end{pmatrix} \quad (4)$$

The motion constraint vector is the spatiotemporal gradient of the phase, weighted by the confidence measure, C , which will be introduced in next section.

2.2. Confidence Measure

Using a confidence measure is necessary to give strong features precedence over weaker features and noise. In addition, it is necessary to avoid phase singularities[12, 9] which occur when two frequencies interfere in the filter response. These singularities must be discovered and treated as outliers. All this is done by assigning a confidence value to each constraint. Our confidence measure is inspired by the stereo disparity algorithm by Westelius [12], which in turn is inspired by Wilson-Knutsson[2]. It is a product of several factors, where the most important feature is the magnitude.

Our confidence measure for magnitude may seem complicated at first glance. Except for suppressing weak features, it is also sensitive to difference between the two frames. This reduces the influence of structure that only exist in one of the images, such as moving shadows, appearing objects

and other features not moving according to the motion we estimate.

$$C_{mag} = \frac{\|q_A\|^2 \|q_B\|^2}{(\|q_A\|^2 + \|q_B\|^2)^{3/2}} \quad (5)$$

Other factors have been added to reflect whether the gradient, is sound for the specific quadrature filter in use. Negative frequencies are illegal and indicate phase singularities[9, 12].

$$C_{freq>0} = \begin{cases} 1 & \text{if } \hat{\mathbf{n}}^T \nabla \theta > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

We have also used some more features in our confidence measure, that are omitted here. Those are sensitive to consistency between the two frames and probability of phase wrap around (2π).

3. EXPERIMENTAL RESULTS

We have used the phase-based method on various image data, and it has always turned out advantageous to the conventional gradient method. One important application is motion compensation in sequences of medical X-ray images, digital subtraction angiography. Conventional gradient methods fail to estimate motions accurately, due to different DC level in the frames and motions of the injected contrast agent. Suppressing low frequencies helps a lot, but still our phase-based method is superior.

Figures 3 - 6 show a comparison for a medical X-ray angiography sequence. Image subtraction is used to filter out the vasculature and take away the bones and tissue. We get much less motion artifacts when using phase-based motion estimation. Constraints over the image are integrated, to fit a local-global deformable motion model[7] in least square sense. We have used four quadrature filters in different directions in conjunction with multiple scales and iterative refinement.

We have also compared accuracy on images where motions come from synthetic shifts. A real world test image has been shifted different amounts in different directions. To avoid influence from subpixel warps, the image has been subsampled after the warp. One might expect conventional gradient methods to work pretty good on these images that have perfect intensity conservation between frames. But still, our phase-based method is more accurate, as shown in figures 1 - 2.

4. FUTURE DEVELOPMENT

The confidence measure in this article is designed without much theory and experiments. It might be possible to get

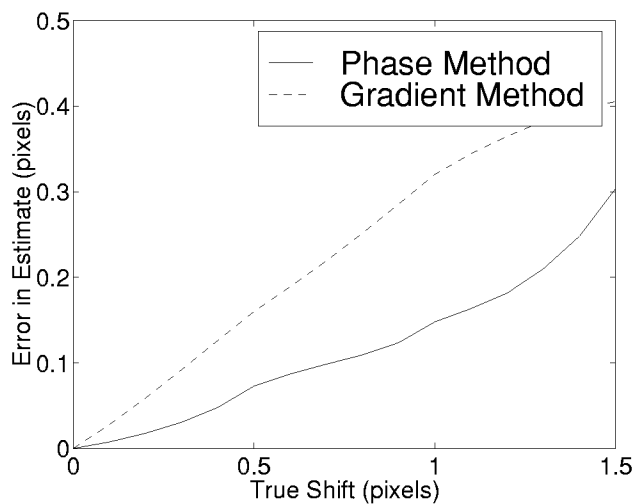


Figure 1: The phase-based method is more accurate than the conventional gradient method. This figure shows a comparison on images (Lena 256x256) that are shifted synthetically. (One pass estimation, i.e. no iterative refinement)

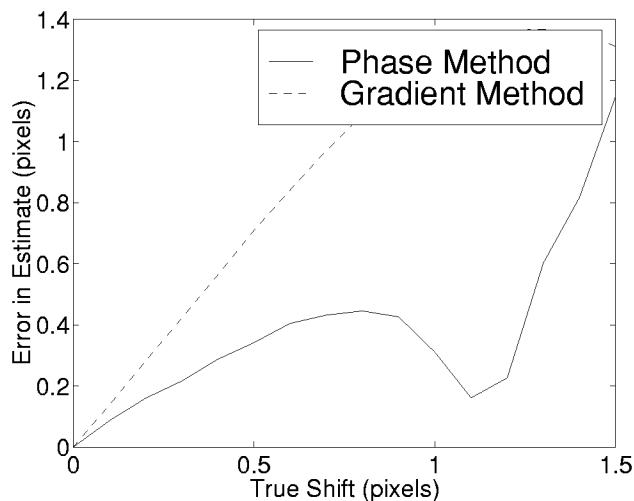


Figure 2: Another accuracy comparison between our phase-based method and the conventional gradient method. This figure shows a comparison on images (Debbie 128x128) that are shifted synthetically. (One pass estimation, i.e. no iterative refinement)

better accuracy with application specific confidence measures. For instance, in some applications it may be more or less important to check consistency between frames. In general, it can be that the confidence measure factor on magnitude, eq. (5) should depend on the noise level. Instead of being linear to magnitude, it should be a sigmoid function that give almost equal confidence to all features that are well above the noise level.

5. REFERENCES

- [1] A. D. Calway, H. Knutsson, and R. Wilson. Multiresolution estimation of 2-d disparity using a frequency domain approach. In *Proc. British Machine Vision Conf.*, Leeds, UK, September 1992.
- [2] A. D. Calway, H. Knutsson, and R. Wilson. Multiresolution estimation of 2-d disparity using a frequency domain approach. In *Proc. British Machine Vision Conf.*, Leeds, UK, September 1992.
- [3] A. D. Calway, H. Knutsson, and R. Wilson. Multiresolution frequency domain algorithm for fast image registration. In *Proc. 3rd Int. Conf. on Visual Search*, Nottingham, UK, August 1992.
- [4] D. J. Fleet and A. D. Jepson. Computation of Component Image Velocity from Local Phase Information. *Int. Journal of Computer Vision*, 5(1):77–104, 1990.
- [5] D. J. Fleet, A. D. Jepson, and M. R. M. Jenkin. Phase-based disparity measurement. *CVGIP Image Understanding*, 53(2):198–210, March 1991.
- [6] G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision*. Kluwer Academic Publishers, 1995. ISBN 0-7923-9530-1.
- [7] Magnus Hemmendorff. Motion compensated digital subtraction angiography. Master's thesis, Linköpings universitet, 1997. LiTH-ISY-EX-1750.
- [8] B. K. P. Horn and B. G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.
- [9] A. D. Jepson and D. J. Fleet. Scale-space singularities. In O. Faugeras, editor, *Computer Vision-ECCV90*, pages 50–55. Springer-Verlag, 1990.
- [10] B. Lucas and T. Kanade. An Iterative Image Registration Technique with Applications to Stereo Vision. In *Proc. Darpa IU Workshop*, pages 121–130, 1981.
- [11] J. Shi and C. Tomasi. Good features to track. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
- [12] C-J. Westelius. *Focus of Attention and Gaze Control for Robot Vision*. PhD thesis, Linköping University, Sweden, S-581 83 Linköping, Sweden, 1995. Dissertation No 379, ISBN 91-7871-530-X.

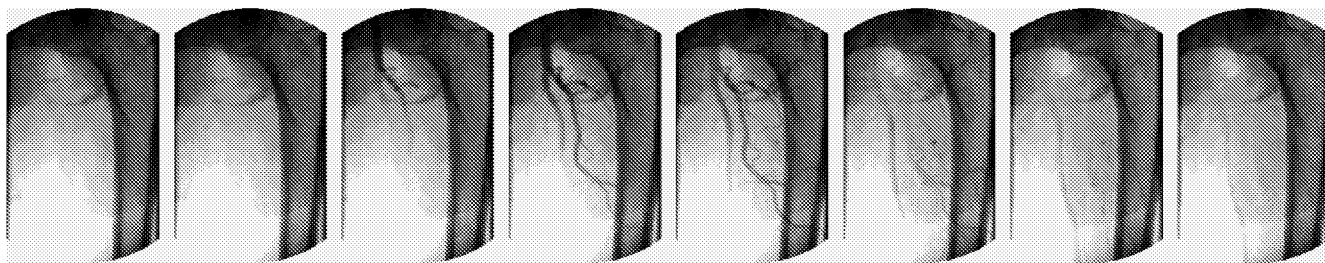


Figure 3: *Original X-ray images*

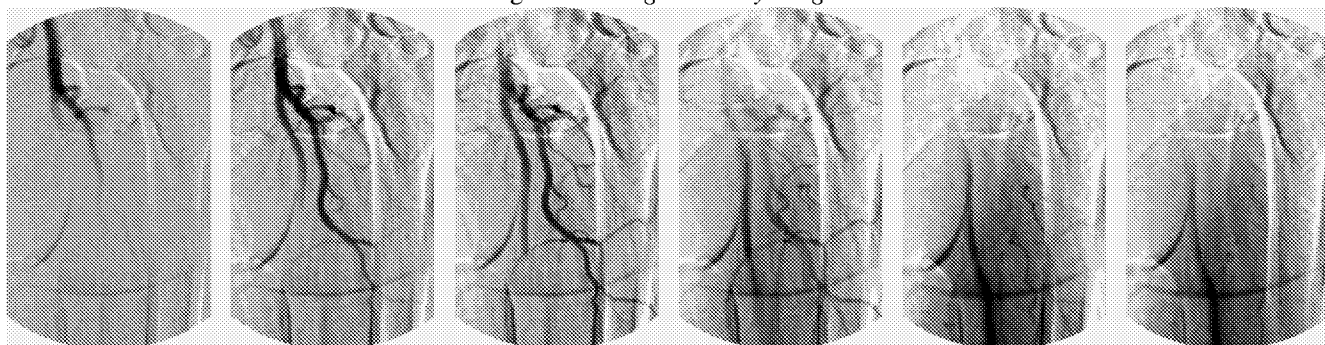


Figure 4: *Subtraction Images, no motion compensation*

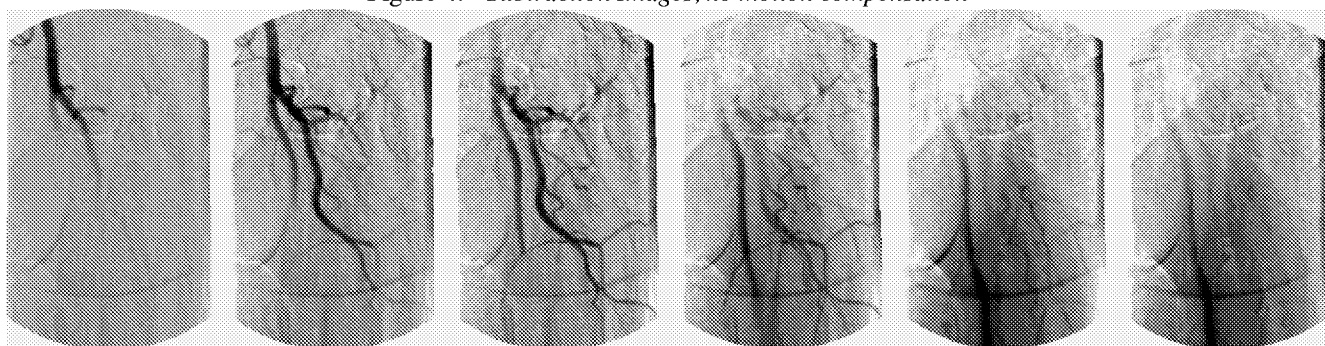


Figure 5: *Subtraction Images, motion compensation based on conventional gradient method, after filtering out low frequencies.*

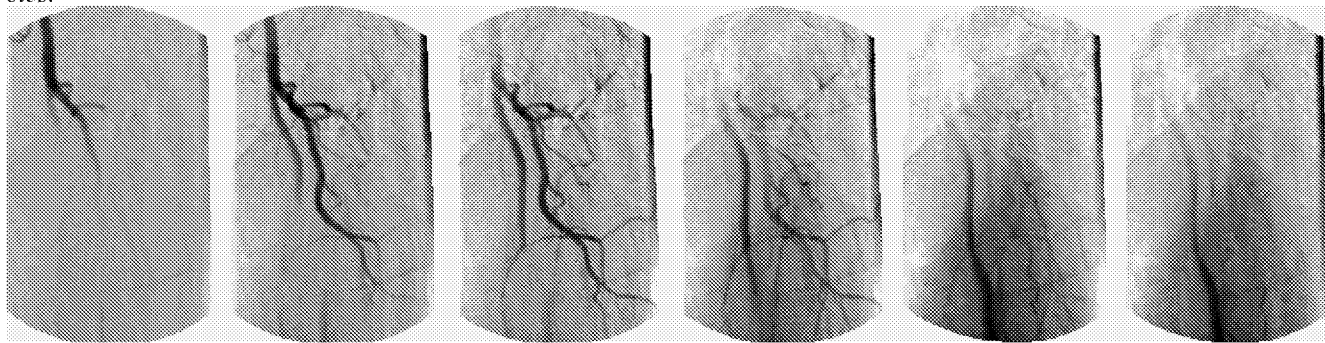


Figure 6: *Subtraction Images, motion compensation based on our phase-based method. Note there are less artifacts compared to figure 5*