Revisiting the Brightness Constraint: Probabilistic Formulation and Algorithms

Venu Madhav Govindu*

HIG-25, Simhapuri Layout,
Visakhapatnam, AP 530047, India
venu@narmada.org

Abstract. In this paper we introduce a principled approach to modeling the image brightness constraint for optical flow algorithms. Using a simple noise model, we derive a probabilistic representation for optical flow. This representation subsumes existing approaches to flow modeling, provides insights into the behaviour and limitations of existing methods and leads to modified algorithms that outperform other approaches that use the brightness constraint. Based on this representation we develop algorithms for flow estimation using different smoothness assumptions, namely constant and affine flow. Experiments on standard data sets demonstrate the superiority of our approach.

1 Introduction

Computing the optical flow field between images has been a central problem in computer vision. Thanks to numerous investigations over the past two decades, both our understanding of the problem and its algorithmic implementation have become increasing sophisticated (see [1, 2, 3, 4, 5, 6] and references therein). Most flow algorithms are based on the brightness constraint that is derived from an intensity conservation principle. Given two images taken at time-instants t and $t+1$ and denoting the flow at pixel (x, y) by (u, v), by conservation of intensity we have the relationship, $I(x, y, t) = I(x+u, y+v, t+1)$. By expanding this function as a Taylor series we have a first-order approximation $I(x+u, y+v, t+1) \approx I(x, y, t) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t}.1$ which simplifies to $I_x u + I_y v + I_t = 0$ where I_x, I_y and I_t are the derivatives in the x, y and t dimensions respectively. This relationship is known as the brightness constraint and can be interpreted as a line in the (u, v) flow space. Since the flow at a point consists of two values, a single brightness constraint is insufficient, i.e. flow estimation is *ill-posed*. Therefore, flow is estimated by imposing additional assumptions of smoothness on the flow field.

There are three significant issues with using the brightness constraint that need to be addressed *simultaneously* in any representation. Firstly, the brightness constraint is derived using a first-order Taylor approximation implying that the flow magnitude is assumed to be small. However many algorithms violate this underlying assumption and

* The work in this paper was partially supported by NSF Grant IIS-03-25715 during the author's visit to the University of Maryland, College Park, USA.

A. Leonardis, H. Bischof, and A. Pinz (Eds.): ECCV 2006, Part III, LNCS 3953, pp. 177–188, 2006.

treat the brightness constraint as an *algebraic* line with infinite extent[1]. Secondly, the interpretation of the brightness constraint as a single line in the $u-v$ space is based on the assumption that the image derivatives observed are 'true' values. Thus the existence of noise in the observed image data is not explicitly accounted for, leading to unprincipled algorithms. Thirdly and most importantly, the derivation of the brightness constraint itself is based on an incorrect model where the temporal dimension is treated differently from the spatial dimensions which introduces undesirable biases. Perhaps this derives from the early methods which assumed that only two images were available, i.e. with a time-step of 1. We shall demonstrate in this paper that the correct approach is to *model* the spatio-temporal volume in a uniform and continuous manner and introduce the specific discretisation of the spatio-temporal image data only as an *algorithmic* detail. This approach immediately allows us to explain the behaviour of well-known flow algorithms and also recast their assumptions into more accurate versions.

In this paper we simultaneously address all the three limitations mentioned above. We systematically account for the data noise and also naturally allow for incorporation of priors that agree with the small flow assumption. By treating the spatio-temporal dimensions in a uniform framework, a key insight that arises is that the correct *representation* for estimating image flow is not the two-dimensional vector field, but rather its homogeneous counterpart, i.e. *normalised volume-flow*[2]. We will also show that the popular least-squares (i.e. Lucas-Kanade) and Total Least Squares (henceforth referred to as TLS) methods for constant flow in a patch can both be seen as specific instances of our model. We also emphasise that an optic flow estimator consists of two components, namely the choice of data representation (brightness constraint in our case) and the computational model used to solve the estimation problem. Recent advances in flow estimation have been based on increasingly sophisticated computational approaches, eg. [4, 5, 6, 7]. In contrast this paper focuses on the choice of data representation and *not* on the computational model. The representation proposed here can be incorporated into any computational framework that uses the brightness constraint. We also point out that important issues like robustness to data outliers and motion segmentation are outside the scope of this paper.

2 Probabilistic Brightness Constraint

In this section we derive a probabilistic model for the image brightness constraint. We develop our solution assuming a continuous space-time image volume. We re-emphasise that a time-step of 1 is an artifact of image acquisition and should not influence our problem formulation. Thus, although the spatial and temporal resolutions are different, we make an essential distinction between the model and its algorithmic

[1] While multi-scale techniques exist they are designed to reduce the magnitude of the true flow in an image. This, in principle, does not impose any constraint on the magnitude of the *estimated* flow.

[2] Volume-flow measures the flow field in the spatio-temporal volume and optical flow is its projection onto the image plane. The unit-norm vector, normalised volume-flow is projectively equivalent to optic flow and should not be confused with 'normal flow' which represents the projection of optical flow in a direction orthogonal to the brightness constraint line.

utilisation. We develop our method for continuous data and at the appropriate juncture replace the image derivatives involved by those calculated on discrete image data. This model is at the heart of the subsequent algorithms that we shall develop using different smoothness assumptions.

The estimated image derivatives are represented by $I_d = [I_x, I_y, I_t]^T$. We represent the error in the image derivatives using an additive Gaussian noise model, i.e. $I_d = I_{d0} + n$, where $I_{d0} = [I_{x0}, I_{y0}, I_{t0}]^T$ is the true value of the derivatives and $n = [n_x, n_y, n_t]^T$ is the noise term. For the sake of simplicity of presentation, we shall in the following assume that the noise is zero-mean, independent and identically distributed, i.e. $n \sim N(0, \sigma^2 I_3)$ where I_3 is the 3×3 identity matrix. However this does not preclude the use of more general forms of noise covariance matrices since the measurements can be whitened before applying our analysis. In general, it is realistic to assume that the spatial and temporal derivatives have different covariances due to the nature of sampling in space and time. We denote the three-dimensional volume-flow at a point as $F = [U, V, W]^T$ where U, V and W are the displacements in the x, y and t dimensions respectively. The two-dimensional optical flow is the projection of the volume-flow vector F onto the $x - y$ image plane and is denoted as (u, v) where $u = \frac{U}{W}$ and $v = \frac{V}{W}$. It will be noted that normalised volume-flow f is given by $f = \frac{F}{||F||}$ and is also projectively equivalent to the optical flow (u, v), i.e. $f \propto [u, v, 1]^T$. Using the principle of image brightness conservation, we have $I(x + U, y + V, t + W) = I(x, y, t)$. By a Taylor series expansion around the point (x, y, t) we have $I(x, y, t) + \frac{\partial I}{\partial x}U + \frac{\partial I}{\partial y}V + \frac{\partial I}{\partial t}W = I(x, y, t)$ leading to

$$\frac{\partial I}{\partial x}U + \frac{\partial I}{\partial y}V + \frac{\partial I}{\partial t}W = 0 \tag{1}$$

which is a brightness constraint equation in three-dimensions and can be simply expressed as ${I_d}^T F = 0$. It will be immediately observed here that we have an unknown scale factor for F, i.e. ${I_d}^T F = {I_d}^T(\alpha F) = 0$, implying that we can only derive F upto a scale factor. Hence we fix the scale by using the normalised volume-flow vector, $f = \frac{F}{||F||}$. However, Eqn. 1 applies to the true image derivatives, whereas we can only observe the estimated derivatives. Thus to define the conditional distribution of the flow given the observed image derivatives I_d using the relationship, ${I_{d0}}^T F = (I_d - n)^T F = 0$, we apply the chain rule for conditional probabilities resulting in

$$P(F|I_d) = \int P(F|I_{d0})P(I_{d0}|I_d)\ dI_{d0} \tag{2}$$

From Eqn. 1, for the true image derivatives we note that the linear constraint implies that only flow values that satisfy this equation are admissible. Thus the conditional probability $P(F|I_{d0})$ is described by our brightness constraint and is equal to $\delta({I_{d0}}^T F)$ where $\delta(.)$ is the Delta Function. Also since the true derivatives are perturbed by Gaussian noise to give the observed derivative estimates, we can represent the conditional probability $P(I_{d0}|I_d)$ by using the Gaussian noise prior. This is true since we can equivalently write $I_{d0} = I_d - n$. Thus $P(I_{d0}|I_d) = e^{-\frac{n^T n}{2\sigma^2}}$ where $n = [n_x, n_y, n_t]^T$ represent the noise in the image derivatives. Consequently

$$P(F|I_d) = \int \underbrace{\delta({I_{d0}}^T F)}_{P(F|I_{d0})} \underbrace{e^{-\frac{1}{2\sigma^2}({n_x}^2+{n_y}^2+{n_t}^2)}}_{P(I_{d0}|I_d)} dn_x dn_y dn_t$$

For, simplicity of presentation we ignore the normalisation required here to ensure that the integral measure on the delta-function is equal to one. Expanding the constraint into its respective terms we have ${I_{d0}}^T F = {I_d}^T F - (n_x U + n_y V + n_t W)$. To solve for the integral, we integrate out one variable (n_t in this case) to derive the following

$$P(F|I_d) = \frac{1}{|W|} \int e^{-\frac{1}{2\sigma^2}({n_x}^2+{n_y}^2+\frac{(n_x U+n_y V-c)^2}{W^2})} dn_x dn_y \tag{3}$$

where $c = {I_d}^T F$. This is obtained by integrating out the constraint and substituting for n_t. After some simple algebra, we can rewrite the exponential term of Eqn. 3 as the form

$$(n-\mu)^T R(n-\mu) + \mu_0 \tag{4}$$

with $R = \frac{1}{W^2}\begin{bmatrix} U^2+W^2 & UV \\ UV & V^2+W^2 \end{bmatrix}$, $\mu_0 = \frac{(I_x U+I_y V+I_t W)^2}{U^2+V^2+W^2}$. Therefore ,

$$P(F|I_d) = \frac{e^{-\frac{\mu_0}{2\sigma^2}}}{|W|} \int e^{-\frac{1}{2\sigma^2}(\mathbf{n}-\mu)^T R(\mathbf{n}-\mu)} d\mathbf{n} \tag{5}$$

The integral can be seen to be that of a Gaussian with a covariance of R^{-1} implying that the integral is equal to $|R|^{-\frac{1}{2}}$ and is independent of the value of μ. Now $|R| = \frac{U^2+V^2+W^2}{W^2}$, implying that

$$P(F|I_d) \propto \frac{e^{-\frac{1}{2\sigma^2}\frac{(I_x U+I_y V+I_t W)^2}{(U^2+V^2+W^2)}}}{\sqrt{U^2+V^2+W^2}}$$

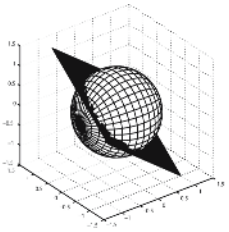
As observed earlier, the optical flow (u, v) is independent of the magnitude of the volume-flow vector F, hence we can set $||F|| = 1$. This implies that for the homogeneous image flow f (or normalised volume-flow) we have

$$P(f|I_d) \propto e^{-\frac{1}{2\sigma^2}\frac{f^T M f}{f^T f}} \tag{6}$$

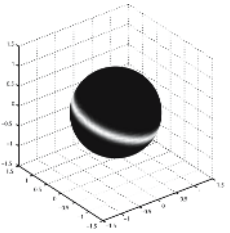
where the 3×3 matrix M is given by

$$M = \begin{bmatrix} I_x \\ I_y \\ I_t \end{bmatrix} \times \begin{bmatrix} I_x & I_y & I_t \end{bmatrix} = \begin{bmatrix} I_x I_x & I_x I_y & I_x I_t \\ I_x I_y & I_y I_y & I_y I_t \\ I_x I_t & I_y I_t & I_t I_t \end{bmatrix} \tag{7}$$

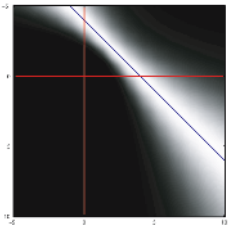
An analysis of this distribution is instructive. If we consider a single pixel, we will note that the probability value in Eqn. 6 is maximised when f is orthogonal to I_d, i.e. ${I_d}^T f_{max} = 0$. Thus f_{max} lies in the plane that is normal to I_d. However since $||f|| = 1$, we have f_{max} confined to the surface of a unit-sphere. Therefore, the locus of f_{max} is a *great circle* on the unit-sphere, see Fig. 1(a). As f deviates from the great circle f_{max},



(a) Single pixel constraint



(b) Probability form of constraint



(c) Projection of distribution on image plane

Fig. 1. Representations of a brightness constraint. (a) the brightness constraint plane intersects with the unit-sphere resulting in a great circle which is equivalent to the conventional brightness constraint line; (b) shows the probability distribution of normalised volume-flow for a single pixel. (c) shows the projection of the distribution in (b) on the $x - y$ image plane. Note the 'fuzzy bow-tie' form of the flow distribution. The maxima of this distribution is the conventional brightness constraint line.

the magnitude of the probability distribution decreases according to Eqn. 6. Thus the probability distribution of the normalised volume-flow vector for a single pixel is a Gaussian-like distribution on the unit-sphere centered on the great circle f_{max} as seen in Fig. 1(b). The great circle and distribution in Fig. 1(b) can be seen to be the unit-sphere equivalents of the brightness constraint line and a Gaussian distribution centered on the line respectively.

However instead of considering a representation of f on the unit-sphere, the conventional approach has been to use $F = [u, v, 1]$. If we substitute this form in Eqn. 6, we see that the exponential term is equal to $\frac{(I_x u + I_y v + I_t)^2}{(u^2+v^2+1)}$ which is identical to the TLS form used in [8, 9, 10]. In turn the equivalent probability distribution for (u, v) is shown in Fig. 1(c) and can be seen to have the so-called 'fuzzy bow-tie' form [11]. As is obvious from the above analysis and the distributions of f in Fig. 1, we note that the fuzzy bow-tie form of the flow distribution is nothing but an *artifact* of using a reduced representational space for the flow information. This arises from projecting a Gaussian-like form on the unit-sphere onto the image plane, i.e. the fuzzy bow-tie form in Fig. 1(c) is the projection of the distribution of Fig. 1(b) onto the image plane. Thus the fuzzy bow-tie distribution is not very illuminating and the probability form of Eqn. 6 is desirable as it leads to more accurate flow estimates. We also point out that our probability model is fundamentally different from that of [12] where a Gaussian noise model is applied to the flow (instead of the image derivatives) and a Gaussian distribution of flow on the image plane is derived. In our case the flow distribution in Eqn. 6 is the *natural representation* of the information in the image derivatives and as will be seen in the rest of the paper, this is a powerful, general representation that can be applied to various smoothness assumptions. It is also germane to point out that in this paper we are modeling the *optic flow field* based on image derivatives which should not be confused with modeling the *motion field* which would depend on a taxonomy of camera motions, zooming, rotating, translating etc. which results in specific types of motion fields.

3 Optic Flow Algorithms

In Sec. 2 we derived a probability distribution for optical flow at a pixel given its corresponding image derivatives. However, since the optical flow field consists of two values at each pixel, the probability distribution derived from a single pixel is insufficient to determine optical flow. In particular, matrix M in Eqn. 6 for a single pixel can be seen to be of rank one. In general, the ill-posedness of optical flow is addressed by making a variety of smoothness assumptions on the flow field which allows us to estimate the flow field using fewer parameters than the number of constraints available. The smoothness assumptions can be broadly characterised as being implicitly due to a parametric model or explicitly due to the use of a regularising smoothness term. Examples of the former are the constant flow assumption of Lucas and Kanade [3], affine flow [13, 14, 15], whereas [2, 6] are examples of an explicit smoothing strategy. In all of these methods, the estimation process is considerable affected by the assumption of a time-step of 1 in the corresponding formulations resulting in bias or a greater error. By explicitly applying our probabilistic formulation to these smoothness assumptions we derive modified algorithms that both clarify the behaviour of the conventional methods and significantly improve their performance. In the remainder of this section we consider constant and affine models and examine their implications for estimating optical flow.

3.1 Constant Flow

The simplest assumption for flow estimation is that of constant flow for an image patch which is the basis for the famous Lucas-Kanade algorithm [3]. Here the brightness constraint is represented by $I_x u + I_y v + I_t = 0$ and for a patch, the residual error is $\mathbf{E} = \sum_k ({I_x}^k u + {I_y}^k v + {I_t}^k)^2$ where k denotes the index of individual pixels in the patch. The minimiser of $\mathbf{E}$ is the Lucas-Kanade solution and is identical to the Ordinary Least Squares (henceforth OLS) solution:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \overline{I_x I_x} & \overline{I_x I_y} \\ \overline{I_x I_y} & \overline{I_y I_y} \end{bmatrix}^{-1} \begin{bmatrix} -\overline{I_x I_t} \\ -\overline{I_y I_t} \end{bmatrix} \tag{8}$$

where $\overline{I_x I_x} = \sum_k {I_x}^k {I_x}^k$ etc. As had been noted in [16] this yields a linear, biased estimate of the flow. The bias appears due to the implicit assumption that the temporal derivatives are noise-free and the use of the TLS method has been suggested to overcome this bias [8, 9]. This can also be explained using our probability distribution for optical flow. For constant flow over a patch, using the conditional probability distribution of Eqn. 6 we have $P(f|patch) = \prod_k P(f|{I_d}^k)$. The flow can be estimated by maximising the conditional probability distribution

$$\max_f \prod_k P(f|{I_d}^k) \Rightarrow \min_f e^{-\frac{1}{2\sigma^2} f^T (\sum_k M_k) f} \tag{9}$$

The estimated flow is the smallest eigen-vector for matrix

$$\overline{M} = \frac{1}{N} \sum_{k=1}^{N} M_k = \frac{1}{N} \begin{bmatrix} \overline{I_x I_x} & \overline{I_x I_y} & \overline{I_x I_t} \\ \overline{I_x I_y} & \overline{I_y I_y} & \overline{I_y I_t} \\ \overline{I_x I_t} & \overline{I_y I_t} & \overline{I_t I_t} \end{bmatrix} \tag{10}$$

where N is the number of pixels in the patch. This is identical to the TLS solution. However, it must be pointed out that the above derivation of the Maximum Likelihood Estimate (MLE) of flow does not incorporate a prior distribution for the flow values. As has been noted in Sec. 1, the brightness constraint is valid only for a small deviation from the point around which the Taylor series expansion is made, i.e. flow cannot be large. This implicit assumption cannot be captured by treating the brightness constraint as an algebraic equation and is often ignored. In our case, since we represent the information at a pixel as a conditional probability distribution we can incorporate the small flow assumption as a prior on the flow field. For the flow values to be small, we note that since $(u, v) = (\frac{U}{W}, \frac{V}{W})$, we require the contribution of U and V to the magnitude of the volume-flow $||F|| = ||(U, V, W)||$ to be small. This notion can be captured by using a Gaussian distribution on the relative magnitudes of U and V, i.e. $\frac{U}{||F||}$ and $\frac{V}{||F||}$. This leads to a distribution of the form

$$P(.) = e^{-\frac{1}{2{\sigma_f}^2}\frac{U^2+V^2}{U^2+V^2+W^2}} = e^{-\frac{1}{2{\sigma_f}^2}\frac{f^T D f}{f^T f}}$$

where D is a diagonal matrix $D = diag([1, 1, 0])$ and the variable σ_f controls the influence of the prior on the estimator. If we reintroduce this prior into Eqn. 2 to weight the δ-function appropriately, our measurement matrix for flow estimation is modified into a Maximum A Posteriori (MAP) Estimate. Thus instead of averaged matrix $\overline{M}$ of Eqn. 10, the flow is seen to be the smallest eigen-vector of matrix $\overline{M}_{map}$ for

$$\overline{M}_{map} = \frac{1}{{\sigma_n}^2} N\overline{M} + \frac{1}{{\sigma_f}^2} D \tag{11}$$

where N is the number of pixels in the patch and σ_n and σ_f are the priors for the image derivative noise and flow magnitude respectively. It will be noted that the observation matrix in Eqn. 11 represents a regularised solution for the TLS problem [17]. For the sake of simplicity we reparametrise this matrix as $\overline{M} + \lambda D$ where λ represents the weight (influence) of the regularising term D. For a given value of λ, the estimated

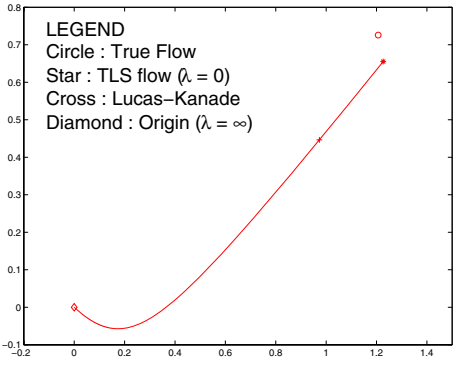


Fig. 2. Flow as a function of the regularisation term λ. The least-squares solution (Lucas-Kanade) lies on this curve. True flow and the TLS solution are also indicated. See Sec. 3.1 for details.

optical flow value is given by the smallest eigen-vector associated with the matrix $\overline{M} + \lambda D$. The behaviour of this parametrised form is particularly illuminating as illustrated in Fig. 2. In the case when $\lambda = 0$, the regulariser has no influence on the estimate and we get the TLS solution. When $\lambda = \infty$, the solution is determined solely by the null-space of the regularising matrix D, i.e. $[0, 0, 1]^T$ equivalent to a flow of $(0, 0)$. This is intuitively correct since here the flow is determined only by the prior which is a Gaussian centered at the origin. As λ varies from 0 to ∞ the estimated optic flow traces a curve from the TLS solution to the origin. Of particular significance is the fact that the Lucas-Kanade (or OLS) solution lies *exactly* on this parametrised curve, i.e. it is identical to a regularised TLS solution of optical flow for a particular value of λ! This relationship is formally described by the following lemma.

Lemma 1. *The Lucas-Kanade estimate of flow (i.e. OLS solution) is identical to the TLS solution for the regularised observation matrix* $\overline{M} + \lambda D$ *where* $\lambda = \frac{1}{N} \sum_{k=1}^{N} {I_t}^k ({I_x}^k u + {I_y}^k v + {I_t}^k)$*,* N *is the number of pixels in the patch and* (u, v) *is the Lucas-Kanade (or OLS) solution.*

Proof: We represent the three-dimension homogeneous co-ordinates of the flow vector as $[x, 1]^T = [u, v, 1]^T$. Further we partition the 3×3 observation matrix as $\overline{M} = \begin{bmatrix} A & b \\ b^T & c \end{bmatrix}$. Since the regularising matrix D=diag([1,1,0]) we have $\overline{M}+\lambda D = \begin{bmatrix} A+\lambda I & b \\ b^T & c \end{bmatrix}$ where I is the 2×2 identity matrix. For the TLS solution of the regularised observation matrix, we have

$$(\overline{M} + \lambda D) \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} A+\lambda I & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \alpha \begin{bmatrix} x \\ 1 \end{bmatrix}$$
$$\Rightarrow (A + \lambda I)x + b = \alpha x \tag{12}$$

Here α is the eigen-value associated with the TLS solution for a given λ. The lemma can now be proved by examination. Let us assume that the flow estimate x is the OLS solution x_{OLS}. By examining the observation matrix $\overline{M}$ of Eqn. 10 and the solution for x_{OLS} in Eqn. 8 we note that $Ax_{OLS} + b = 0$ which implies that for $x = x_{OLS}$ the relationship in Eqn. 12 is satisfied

$$(A + \lambda I)x_{OLS} + b = \lambda x_{OLS} + \underbrace{Ax_{OLS} + b}_{=0} = \alpha x_{OLS}$$

implying that $\lambda = \alpha$. Thus the eigen-relationship for $\overline{M} + \lambda D$ is satisfied for $x = x_{OLS}$ which proves that the OLS flow (i.e. Lucas-Kanade) is also a solution for the regularised TLS for $\lambda = \alpha$. The value of λ can now be easily derived by noting the lower relationship in Eqn. 12, i.e. $\alpha = \lambda = b^T x_{OLS} + c$. The terms b^T and c are the third row of the observation matrix $\overline{M}$ in Eqn. 10 implying that $\lambda = \frac{1}{N} \sum_{k=1}^{N} {I_t}^k ({I_x}^k u + {I_y}^k v + {I_t}^k)$. The form of λ is also intuitively satisfying. Informally speaking, it represents a measure of 'texturedness' in the temporal direction implying that as the temporal derivatives grow in magnitude, the Lucas-Kanade method introduces a greater amount of bias. It is well known that while the TLS solution is unbiased, compared to the OLS solution, the TLS has greater variance. In this context, the influence of the patch size

on λ is informative. When the patch size (N) is small, λ is large implying that our solution introduces a bias to reduce the variance of the solution. Conversely, when the patch size is large, λ is small implying that our solution is closer to the TLS estimate as desired. Thus our formulation can naturally capture the correct representation required for accurate flow estimation and also explains the behaviour of Lucas-Kanade and TLS algorithms.

3.2 Affine Flow

While the constant flow model is simple to implement, its accuracy is inherently limited as flow fields are seldom close to a piece-wise constant model. A more appropriate assumption is that of an affine model. An affine flow field is described by $\begin{bmatrix} u & v \end{bmatrix}^T = A\begin{bmatrix} x & y \end{bmatrix}^T + \begin{bmatrix} t_x & t_y \end{bmatrix}^T$ where (u, v) is the flow at position (x, y) and A is a 2×2 matrix. The affine model has been used to estimate optical flow in [13, 14, 15]. While the TLS estimator is unbiased it has a higher variance than the OLS solution. This implies that for small image patches, with few equations the Lucas-Kanade solution is preferable to the TLS solution. However as we noted in the previous subsection, when we have many equations the TLS solution is preferable to the biased OLS estimate. In general, the affine flow model is estimated for patches larger than those for constant flow since we need many more equations to reliably estimate the six parameters of the affine model. This implies that in our probabilistic model, the prior has little influence on affine estimation and can be neglected in our analysis here. By re-writing the optical flow in homogeneous co-ordinates we have $\mathbf{f} = \mathbf{P}\mathbf{a}$, where $\mathbf{P}$ represents terms relating to pixel position (x, y) and $\mathbf{a}$ is the vectorised representation for the affine parameters. Using this form in the probability model of Eqn. 6 we have

$$P(flow|patch) = \max_{model} \prod_k e^{-\frac{1}{2\sigma^2}\frac{\mathbf{f}^T M_k \mathbf{f}}{\mathbf{f}^T \mathbf{f}}}$$

$$\Rightarrow P(\mathbf{a}|patch) = \max_{\mathbf{a}} \prod_k e^{-\frac{1}{2\sigma^2}\frac{(\mathbf{P}_k\mathbf{a})^T M_k \mathbf{P}_k \mathbf{a}}{(\mathbf{P}_k\mathbf{a})^T \mathbf{P}_k \mathbf{a}}}$$

$$\Rightarrow \mathbf{a} = \arg\min_{\mathbf{a}} \sum_k \frac{\mathbf{a}^T {\mathbf{P}_k}^T M_k \mathbf{P}_k \mathbf{a}}{\mathbf{a}^T {\mathbf{P}_k}^T \mathbf{P}_k \mathbf{a}} \tag{10}$$

Thus the problem of estimating the affine parameters is reduced to the minimisation of a sum of Rayleigh quotients[3]. This particular quotient form occurs frequently in computer vision problems like ellipse fitting etc. and a significant body of work has been devoted to its minimisation. In our solution for the affine parameters we use the First-Order Renormalisation of [19].

3.3 Performance of Affine Flow Estimation

In this subsection we evaluate our affine flow estimation scheme using the standard image sequences of Barron *et al* [1]. All experiments are performed with a fixed set of

[3] In [18], the authors use algebraic arguments to approximate the above objective function as a single ratio of quadratic forms where the numerator is an average over the patch for the terms $\mathbf{P}^T M \mathbf{P}$ and the denominator is held to be $\mathbf{P}^T \mathbf{P}$ for a given pixel co-ordinates (x, y).

parameters. Each image sequence is smoothed using a separable Gaussian kernel with uniform spatial and temporal standard deviation of 1.4 pixels. The derivative filter is the series-design filter used in [18]. Apart from the image derivative filter the patch size is an important parameter that influences performance by determining the trade-off between estimation accuracy (requiring large patches) and resolution (requiring small patches). Throughout our experiments we use a constant patch size of 31×31 pixels and estimate the affine flow for such patches with a shift of 5 pixels in each direction. Thus each pixel is present in multiple patches and the flow estimate is the average over all patch estimates. We tabulate our results in Tables 1- 4. The error measure is identical to that of [1] and can be seen to measure the angle between the normalised volume-flow representations of the ground truth and the estimate. The error values for the first four methods are taken from [1]. As can be easily observed, our algorithm performs very well with respect to the other procedures. In particular, we point out that our accuracy is achieved without the use of any adaptive schemes. Also the standard deviation of our error values are *significantly* smaller compared to other methods. For the Yosemite sequence, we note that in comparison with the adaptive scheme of [4], our estimator has almost the same performance (error of 1.16° compared to 1.14°) whereas our standard deviation is significantly smaller (1.17° compared to 2.14°). While, the results of [5] on the Yosemite sequence are superior to ours, we reiterate that our performance

Table 1. Sinusoid Sequence Results

Method	Error (in °)		Density
	μ	σ	
Lucas-Kanade	2.47	0.16	100 %
Horn-Schunck	2.55	0.59	100 %
Fleet-Jepson	0.03	0.01	100 %
Uras *et al.*	2.59	0.71	100 %
Farneback [20]	0.74	0.03	100 %
Liu *et al.* [18]	0.31	0.05	100 %
Our method	**0.09**	**0.03**	**100 %**

Table 2. Translating Tree Results

Method	Error (in °)		Density
	μ	σ	
Lucas-Kanade	0.66	0.67	39.8 %
Horn-Schunck	2.02	2.27	100 %
Fleet-Jepson	0.32	0.38	74.5 %
Uras *et al.*	0.62	0.52	100 %
Farneback [20]	0.62	1.99	100 %
Liu *et al.* [18]	0.20	0.62	100 %
Our method	**0.15**	**0.10**	**100 %**

Table 3. Diverging Tree Results

Method	Error (in °)		Density
	μ	σ	
Lucas-Kanade	1.94	2.06	48.2 %
Horn-Schunck	2.55	3.67	100 %
Fleet-Jepson	0.99	0.78	61.0 %
Uras *et al.*	4.64	3.48	100 %
Farneback [20]	0.75	0.69	100 %
Liu *et al.* [18]	0.65	1.73	100 %
Our method	**0.51**	**0.21**	**100 %**

Table 4. Yosemite Results (without clouds)

Method	Error (in °)		Density
	μ	σ	
Lucas-Kanade ($\lambda_2 \geq 1.0$)	3.21	5.34	39.5 %
Horn-Schunck	3.68	4.90	100 %
Uras *et al.*	6.47	9.48	84.6 %
Memin-Perez [7]	1.58	1.21	100 %
Weickert *et al.* [6]	1.46	*	100 %
Liu *et al.* [18]	1.39	2.83	100 %
Farneback [20]	1.40	2.57	100 %
Farneback [4]	1.14	2.14	100 %
Papenberg *et al.* [5]	0.99	1.17	100 %
Our method	**1.16**	**1.17**	**100 %**

is achieved by focusing on the *representation* of the brightness constraint and not on sophisticated numerical minimisers. In summary, we note that our probability representation is powerful and even a straight-forward application of this model outperforms almost all other flow estimators. Other refinements like robustness, adaptive patches, and more accurate minimisers can be expected to further improve our results.

4 Conclusions

In this paper we have introduced a principled approach to modeling the brightness constraint. The resultant probabilistic model is shown to be powerful and can both explain the behaviour of existing flow algorithms and significantly improve their performance. Future work will address more sophisticated minimisation approaches and also the utilisation of our probabilistic model to solve for volume-flow in the spatio-temporal volume of images and for direct motion estimation and segmentation.

Acknowledgments

Thanks are due to Rama Chellappa for his continuous support and encouragement. Yonathan Wexler, David Jacobs and Sameer Agarwal provided useful comments on various drafts of this paper.

References

1. Barron, J., Fleet, D., Beauchemin, S.: Performance of optical flow techniques. International Journal of Computer Vision **12** (1994) 43–77
2. Horn, B., Schunck, B.: Determining optical flow. Artificial Intelligence **17** (1981) 185–203
3. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Proc. of DARPA Workshop. (1981) 121–130
4. Farneback, G.: Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. In: Proc. International Conf. on Computer Vision. Volume 1. (2001) 77–80
5. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. International Journal of Computer Vision (to appear)
6. Bruhn, A., Weickert, J., Schnorr, C.: Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. International Journal of Computer Vision **61** (2005) 211–231
7. Memin, E., Perez, P.: Hierarchical estimation and segmentation of dense motion fields. International Journal of Computer Vision **46** (2002) 129–155
8. Weber, J., Malik, J.: Robust computation of optical-flow in a multiscale differential framework. International Journal of Computer Vision **14** (1995) 67–81
9. Nestares, O., Fleet, D., Heeger, D.: Likelihood functions and confidence bounds for total-least-squares problems. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. (2000) I: 523–530
10. Wang S., Markandey V., R.A.: Total least squares fitting spatiotemporal derivatives to smooth optical flow field. In: Proc. of the SPIE: Signal and Data processing of Small Targets. Volume 1698. (1992) 42–55

11. Weiss, Y., Fleet, D.J.: Velocity likelihoods in biological and machine vision. In: Probabilistic Models of the Brain: Perception and Neural Function. MIT Press (2002) 77–96
12. Simoncelli, E.P., Adelson, E.H., Heeger, D.J.: Probability distributions of optical flow. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition. (1991) 310–315
13. Bergen, J., Anandan, P., Hanna, K., Hingorani, R.: Hierarchical model-based motion estimation. In: Proc. of European Conference on Computer Vision. (1992) 237–252
14. Wang, J., Adelson, E.: Representing moving images with layers. IEEE Transactions on Image Processing **3** (1994) 625–638
15. Ju, S., Black, M., Jepson, A.: Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency. In: IEEE Conf. on Computer Vision and Pattern Recognition. (1996) 307–314
16. Van Huffel, S., Vandewalle, J.: The Total Least Squares Problem : Computational Aspects and Analysis. SIAM (1991)
17. Golub, G.H., Hansen, P.C., O'Leary, D.P.: Tikhonov regularization and total least squares. SIAM Journal on Matrix Analysis and Applications **21** (1999) 185–194
18. Liu, H., Chellappa, R., Rosenfeld, A.: Accurate dense optical flow estimation using adaptive structure tensors and a parametric model. IEEE Transactions on Image Processing **12** (2003) 1170–1180
19. Chojnacki, W., Brooks, M., van den Hengel, A.: Rationalising the renormalisation method of kanatani. Journal of Mathematical Imaging and Vision **14** (2001) 21–38
20. Farneback, G.: Fast and accurate motion estimation using orientation tensors and parametric motion models. In: Proc. of 15th International Conference on Pattern Recognition. Volume 1. (2000) 135–139