Selection of Local Optical Flow Models by Means of Residual Analysis

Björn Andres, Fred A. Hamprecht, and Christoph S. Garbe

Interdisciplinary Center for Scientific Computing
University of Heidelberg, 69120 Heidelberg, Germany

Abstract. This contribution presents a novel approach to the challenging problem of model selection in motion estimation from sequences of images. New light is cast on parametric models of local optical flow. These models give rise to parameter estimation problems with highly correlated errors in variables (EIV). Regression is hence performed by equilibrated total least squares. The authors suggest to adaptively select motion models by testing local empirical regression residuals to be in accordance with the probability distribution that is theoretically predicted by the EIV model. Motion estimation with residual-based model selection is examined on artificial sequences designed to test specifically for the properties of the model selection process. These simulations indicate a good performance in the exclusion of inappropriate models and yield promising results in model complexity control.

1 Introduction & Related Work

In their well-known contribution [1], Black and Jepson propose to estimate optical flow independently for segmented spatiotemporal regions. Parameters of optical flow models are hence allowed to depend on non-trivial subsets of the spatiotemporal volume. The exploitation of the full potential of this approach involves the three challenging problems of motion segmentation, noise estimation and motion model selection. These problems are connected by the fact that violations of suitable models that exceed the scale of noise indicate segment borders. Gheissari et al. [2] comprehensively discuss this interrelation and demonstrate how local optical flow estimation, motion segmentation and motion model selection can be incorporated into an unsupervised motion segmentation framework. This paper focuses on the selection of suitable parametric optical flow models. While a simple model fails to approximate data of higher intrinsic complexity under low noise conditions, a complex model is prone to over-fitting in the presence of noise. Various information criteria have been proposed that penalize model complexity in order to avoid over-fitting. Among the most popular are Akaike's Information Criterion [3] as well as the Bayesian Information Criterion [4]. In the context of motion estimation, the model selection problem has been discussed by Wechsler et al. [5] as well as by Gheissari et al. [2]. However, "[...] *none of the existing model selection criteria is capable of reliably identifying the true underlying model* [...]. *The main reason is that the available*

information theoretic model selection criteria are based on the assumptions that noise is very small and the data size is large enough" [2]. Hence, Gheissari et al. suggest to consider the constraint surfaces of parametric models as thin plates and to penalize the strain energy of these plates according to a physical model. They show a successful application of this surface selection criterion (SSC) in a motion segmentation framework. As the SSC incorporates only second order derivatives of the model surfaces, it cannot be used to distinguish different linear models. Moreover, if information on the distribution of noise is available from camera calibration measurements or noise estimation, probabilistic model selection criteria that incorporate this information should be employed. This paper is intended to fill the gap between information theoretic penalization and heuristic surface modeling. Following the general idea of Cootes et al. [6], we suggest to assess parametric optical flow models by measuring the discrepancy between the empirical distribution of regression residuals and the probability density function (PDF) predicted from theory. This paper is organized as follows. In the next section, we formalize the concept of local optical flow to cast new light on the interrelation of optical flow estimation, motion segmentation and motion model selection. In terms of local optical flow, we then outline in section 3 the specifics of parameter estimation with respect to motion model selection. This includes equilibrated total least squares (ETLS) estimation under a suitably defined Errors-in-Variables (EIV) model. In section 4, the probability distribution of regression residuals is derived from the EIV model. Section 5 deals with simulations conducted to test the proposed model selector for its specific properties. We applied this method to artificial sequences featuring grayvalue structure on multiple scales. Real world video data as well as standard benchmark sequences such as the Yosemite sequence are not suitable to test for the specifics of model selection as the model selector has no intrinsic capability of overcoming the aperture problem.

2 Local Optical Flow

We formalize the concept of local optical flow in order to strengthen the interrelation of optical flow estimation, motion segmentation and motion model selection. If $n_x, n_y, n_t, n_c \in \mathbb{N}$, $P = \{1, \ldots, n_x\} \times \{1, \ldots, n_y\} \times \{1, \ldots, n_t\}$ and $C = \{0, \ldots, n_c\}$ then, the mapping $g : P \to C$ shall be referred to as an *irradiance signal*. Moreover, any mapping $g : \mathbb{R}^3 \to \mathbb{R}$ shall be termed an *ideal irradiance signal*. Herein, for $x \in P$, $g(x, y, t)$ may represent the mean irradiance onto the pixel indicated by (x, y) over the time interval indicated by t as measured with the finite intensity range and resolution given by C [7].

Definition 1 (Optical Flow). *Let $g : \mathbb{R}^3 \to \mathbb{R}$ be an ideal irradiance signal such that the first partial derivatives of g exist and let $(u, v)^T : \mathbb{R}^3 \to \mathbb{R}^2$. Then, $(u, v)^T$ shall be referred to as a field of* optical flow *precisely if*

$$\partial_t g + u\partial_x g + v\partial_y g = 0 \tag{1}$$

holds, which is is the well-known brightness change constraint equation *(BCCE).*

Definition 2 (Local Optical Flow). *Let $g : \mathbb{R}^3 \to \mathbb{R}$ be an ideal irradiance signal such that the first partial derivatives of g exist and let $(u, v)^T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^2$. Moreover, let $\omega : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}_0^+$ such that $\forall \boldsymbol{x} \in \mathbb{R}^3 : \omega(\boldsymbol{x}, \boldsymbol{x}) > 0$. Then, $(u, v)^T$ shall be referred to as a field of* local optical flow *with respect to ω precisely if*

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^3 : \ \omega(\boldsymbol{x}, \boldsymbol{x}')(\partial_t g(\boldsymbol{x}') + u(\boldsymbol{x}, \boldsymbol{x}')\partial_x g(\boldsymbol{x}') + v(\boldsymbol{x}, \boldsymbol{x}')\partial_y g(\boldsymbol{x}')) = 0 \ . \quad (2)$$

The mapping ω shall then be referred to as an aperture function *and, for all $\boldsymbol{x} \in \mathbb{R}^3$, the set $U_\omega(\boldsymbol{x}) := \{\boldsymbol{x}' \in \mathbb{R}^3 | \omega(\boldsymbol{x}, \boldsymbol{x}') > 0\}$ shall be termed the* motion neighborhood *of $\boldsymbol{x}$. In this paper, we refer to (2) as the* local brightness change constraint equation (LBCCE).

(Local) optical flow is often considered in conjunction with parametric models of u and v of which the parameters are estimated such that (1) and (2), respectively hold approximately. A way of looking at the definition of local optical flow is the following: Fix an $\boldsymbol{x} \in \mathbb{R}^3$. Now, the estimation of $(u, v)^T$ from the LBCCE for $\boldsymbol{x}$ is indeed the estimation of optical flow $(u', v')^T : \mathbb{R}^3 \to \mathbb{R}^2$ at this pixel on the data $U_\omega(\boldsymbol{x})$ namely, $\forall \boldsymbol{x}' \in \mathbb{R}^3 : u'(\boldsymbol{x}') = u(\boldsymbol{x}, \boldsymbol{x}') \wedge v'(\boldsymbol{x}') = v(\boldsymbol{x}, \boldsymbol{x}')$. Moreover, local optical flow $(u, v)^T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^2$ comprises optical flow $(u', v')^T : \mathbb{R}^3 \to \mathbb{R}^2$ as the special case in which it is assumed that $\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^3 : u(\boldsymbol{x}, \boldsymbol{x}') = u'(\boldsymbol{x}') \wedge v(\boldsymbol{x}, \boldsymbol{x}') = v'(\boldsymbol{x}')$ i.e., for all $\boldsymbol{x}' \in \mathbb{R}^3$, $(u(\boldsymbol{x}, \boldsymbol{x}'), v(\boldsymbol{x}, \boldsymbol{x}'))$ is independent of $\boldsymbol{x}$. The generality of the LBCCE affords that $U_\omega(\boldsymbol{x})$ need not be, for instance, topologically connected and that, for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^3$ such that $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, $U_\omega(\boldsymbol{x}_1)$ and $U_\omega(\boldsymbol{x}_2)$ need neither be disjoint nor otherwise related. The aim in motion segmentation is to find a suitable aperture function ω that partitions the preimage of the irradiance signal i.e., $\forall \boldsymbol{x_1}, \boldsymbol{x_2} \in \mathbb{R}^3 : U_\omega(\boldsymbol{x_1}) = U_\omega(\boldsymbol{x_2}) \vee U_\omega(\boldsymbol{x_1}) \cap U_\omega(\boldsymbol{x_2}) = \emptyset$. In terms of local optical flow, the classical approach by Lucas and Kanade [8], to estimate optical flow for small identical spatiotemporal neighborhoods of each pixel, is to consider, for a given extension $d_s, d_t \in \mathbb{R}_0^+$ of these neighborhoods, the aperture function ω such that $\forall (x, y, t)^T, (x', y', t')^T \in \mathbb{R}^3$: $\omega((x, y, t), (x', y', t')) = \Theta(d_s - |x' - x|)\Theta(d_s - |y' - y|)\Theta(d_t - |t' - t|)$ (with Θ denoting the Heaviside step function). Black and Jepson [1] investigate several parametric models of local optical flow, among these the local planarity assumption.

Definition 3 (Local Planarity (LPL)). *Let $(u, v)^T : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^2$ and $\omega : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}_0^+$. Then, $(u, v)^T$ shall be called* locally planar *with respect to ω precisely if $\exists p_1, \ldots, p_8 : \mathbb{R}^3 \to \mathbb{R}$ $\forall (x, y, t)^T = \boldsymbol{x} \in \mathbb{R}^3$ $\forall (x', y', t')^T = \boldsymbol{x}' \in U_\omega(\boldsymbol{x})$:*

$$\begin{pmatrix} u(\boldsymbol{x}, \boldsymbol{x}') \\ v(\boldsymbol{x}, \boldsymbol{x}') \end{pmatrix} = \begin{pmatrix} p_1(\boldsymbol{x}) \\ p_2(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} p_3(\boldsymbol{x}) & p_4(\boldsymbol{x}) \\ p_5(\boldsymbol{x}) & p_6(\boldsymbol{x}) \end{pmatrix} \begin{pmatrix} x' - x \\ y' - y \end{pmatrix} + \begin{pmatrix} (x' - x)^2 & (x' - x)(y' - y) \\ (x' - x)(y' - y) & (y' - y)^2 \end{pmatrix} \begin{pmatrix} p_7(\boldsymbol{x}) \\ p_8(\boldsymbol{x}) \end{pmatrix} \ . \quad (3)$$

More restrictive models are obtained from LPL by imposing constraints on the parameter functions such as those to be found in Table 1. Given the LPL model,

Table 1. Parametric local optical flow models obtained from restrictions imposed on LPL. k indicates the number of parameter functions.

Code	k	Description	Restriction on LPL
LPL	8	Planar	none
LAF	6	Affine	$p_1 = p_2 = 0$
LDR	4	Divergence and Rotation	$p_1 = p_2 = 0, p_3 = p_6, p_4 = -p_5$
LSS	4	Stretch and Shear	$p_1 = p_2 = 0, p_3 = -p_6, p_4 = p_5$
LC	2	Constant	$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = 0$

define $\boldsymbol{a}_g : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^8$, $b_g : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ and $\boldsymbol{p} : \mathbb{R}^3 \to \mathbb{R}^8$ such that $\boldsymbol{p} := (p_1, \dots, p_8)^T$ and $\forall (x, y, t)^T = \boldsymbol{x}, (x', y', t')^T = \boldsymbol{x}' \in \mathbb{R}^3$:

$$\boldsymbol{a}_g(\boldsymbol{x}, \boldsymbol{x}') := \omega(\boldsymbol{x}, \boldsymbol{x}') \begin{pmatrix} \partial_x g(\boldsymbol{x}') \\ \partial_y g(\boldsymbol{x}') \\ (x' - x)\partial_x g(\boldsymbol{x}') \\ (y' - y)\partial_x g(\boldsymbol{x}') \\ (x' - x)\partial_y g(\boldsymbol{x}') \\ (y' - y)\partial_y g(\boldsymbol{x}') \\ (x' - x)^2 \partial_x g(\boldsymbol{x}') + (x' - x)(y' - y)\partial_y g(\boldsymbol{x}') \\ (x' - x)(y' - y)\partial_x g(\boldsymbol{x}') + (y' - y)^2 \partial_y g(\boldsymbol{x}') \end{pmatrix} , \quad (4)$$

$$b_g(\boldsymbol{x}, \boldsymbol{x}') := -\omega(\boldsymbol{x}, \boldsymbol{x}')\partial_t g(\boldsymbol{x}') \ . \quad (5)$$

Then, the LBCCE shall be written as

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^3 : \quad \boldsymbol{a}_g^T(\boldsymbol{x}, \boldsymbol{x}')\boldsymbol{p}(\boldsymbol{x}) = b_g(\boldsymbol{x}, \boldsymbol{x}') \ . \quad (6)$$

Analogous definitions of $\boldsymbol{a}_g$, b_g and $\boldsymbol{p}$ exist for the parametric models LAF, LDR, LSS and LC. Optimized linear shift invariant (LSI) operators with finite impulse response (FIR) [9] are used to compute derivatives of (non-ideal) irradiance signals. The preimage of such a signal is finite and so is hence $U_\omega(\boldsymbol{x})$ for all $\boldsymbol{x} \in P$[1]. Finiteness allows to express (6) as a set of systems of equations.

Definition & Proposition 4 (LPL Data). *Let $g : P \to C$ be an irradiance signal, $\omega : P \times P \to \mathbb{R}_0^+$ such that $\forall \boldsymbol{x} \in P : \omega(\boldsymbol{x}, \boldsymbol{x}) > 0$. Let $\forall \boldsymbol{x} \in P : U_\omega(\boldsymbol{x}) = \{\boldsymbol{x}' \in P | \omega(\boldsymbol{x}, \boldsymbol{x}') > 0\}$, $m : P \to \mathbb{N}$ and $\forall \boldsymbol{x} \in P : \boldsymbol{x}'_1, \dots, \boldsymbol{x}'_{m(\boldsymbol{x})}$ such that $\{\boldsymbol{x}'_1, \dots, \boldsymbol{x}'_{m(\boldsymbol{x})}\} = U_\omega(\boldsymbol{x})$. Moreover, consider $\boldsymbol{a}_g$ and b_g as defined in (4) and (5), respectively. Then,*

$$A_g(\boldsymbol{x}) := \begin{bmatrix} \boldsymbol{a}_g^T(\boldsymbol{x}, \boldsymbol{x}'_1) \\ \cdots \\ \boldsymbol{a}_g^T(\boldsymbol{x}, \boldsymbol{x}'_{m(\boldsymbol{x})}) \end{bmatrix} \quad \text{and} \quad \boldsymbol{b}_g(\boldsymbol{x}) := \begin{bmatrix} b_g(\boldsymbol{x}, \boldsymbol{x}'_1) \\ \cdots \\ b_g(\boldsymbol{x}, \boldsymbol{x}'_{m(\boldsymbol{x})}) \end{bmatrix} \ . \quad (7)$$

shall be termed the data matrix *and* data vector*, respectively of the LPL model. The LBCCE (6) is then equivalent to*

$$\forall \boldsymbol{x} \in P : \quad A_g(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x}) = \boldsymbol{b}_g(\boldsymbol{x}) \ . \quad (8)$$

[1] Only pixels $\boldsymbol{x} \in P$ at suitable distance to the border of P such that the derivatives can be computed for all $\boldsymbol{x}' \in U_\omega(\boldsymbol{x})$ are considered.

3 Parameter Estimation

We assume grayvalues to be corrupted by *additive* noise. The additive EIV model claims the existence of a true signal $\tau : P \to C$ and, for all $\boldsymbol{x} \in P$, a random variable $\epsilon(\boldsymbol{x})$ (noise) such that

$$\forall \boldsymbol{x} \in P : g(\boldsymbol{x}) = \tau(\boldsymbol{x}) + \epsilon(\boldsymbol{x}) \ . \tag{9}$$

Through the use of LSI operators, derivatives are approximated by linear combinations of grayvalues. The overlap of FIR masks in the computation of these derivatives at nearby pixels introduces correlation to the entries of A_g and $\boldsymbol{b}_g$. As these entries are linear in the derivatives, they can be decomposed with respect to (9) into

$$A_g(\boldsymbol{x}) = A_\tau(\boldsymbol{x}) + A_\epsilon(\boldsymbol{x}) \quad \text{and} \quad \boldsymbol{b}_g(\boldsymbol{x}) = \boldsymbol{b}_\tau(\boldsymbol{x}) + \boldsymbol{b}_\epsilon(\boldsymbol{x}) \ . \tag{10}$$

In the EIV model, it is assumed that $A_\tau(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x}) = \boldsymbol{b}_\tau(\boldsymbol{x})$ holds exactly as opposed to (8) which may be violated by the errors. As discussed comprehensively by Van Huffel [10], total least squares (TLS) would be the unique (with probability one) maximum likelihood estimator of the parameters $\boldsymbol{p}(\boldsymbol{x})$ if the entries of the matrix $[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]$ stemmed from a multivariate normal distribution with zero mean and covariance matrix $\sigma^2 \mathbb{1}$. If these entries were known to be uncorrelated with zero mean and equal variance, TLS would still be a strongly consistent estimator. But in the present case, mutual correlation is introduced by the overlapping FIR masks of derivative operators. The idea in equilibration is to derive from the covariance matrices of the vectors $\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])$ (column-wise vectorization of the matrix $[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]$) square equilibration matrices $W_L(\boldsymbol{x})$ and $W_R(\boldsymbol{x})$ to estimate $\hat{\boldsymbol{p}}(\boldsymbol{x})$ by TLS on the data $W_L(\boldsymbol{x})[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]W_R^T(\boldsymbol{x})$ instead of $[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]$. $W_R^T(\boldsymbol{x})\hat{\boldsymbol{p}}(\boldsymbol{x})$ is then taken as an estimate of the initial problem. Mühlich [11] derives properties of equilibration matrices from the perturbation theory of eigenvectors and presents an algorithm to compute these iteratively from the covariance matrices of the vectors $\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])$. If the aperture function $\omega : P \times P \to \mathbb{R}_0^+$ depends, for all $\boldsymbol{x}, \boldsymbol{x}' \in P$, only on the difference $\boldsymbol{x}' - \boldsymbol{x}$, there exist an $m \in \mathbb{N}$ such that $\forall \boldsymbol{x} \in P : |U_\omega(\boldsymbol{x})| = m$ as well as a common covariance matrix $C \in \mathbb{R}^{m \times (k+1)}$ (k being the number of model parameters) such that

$$\forall \boldsymbol{x} \in P : \quad \mathrm{cov}(\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])) = \mathrm{cov}(\mathrm{vec}([A_\epsilon(\boldsymbol{x}), \boldsymbol{b}_\epsilon(\boldsymbol{x})])) = C \ . \tag{11}$$

The equilibration matrices $W_L \in \mathbb{R}^{m \times m}$ and $W_R \in \mathbb{R}^{(k+1)\times(k+1)}$ are in this case independent of $\boldsymbol{x}$. Equilibration in the context of motion model selection is discussed in detail in [12]. The effect of equilibration is illustrated in Figure 1. It can be seen that the unequilibrated data is highly correlated as well as that some correlation remains after equilibration.

4 Residual Analysis

If the distribution of noise in the grayvalues is known, we propose to test regression residuals to be in accordance with the theoretically expected distribution.

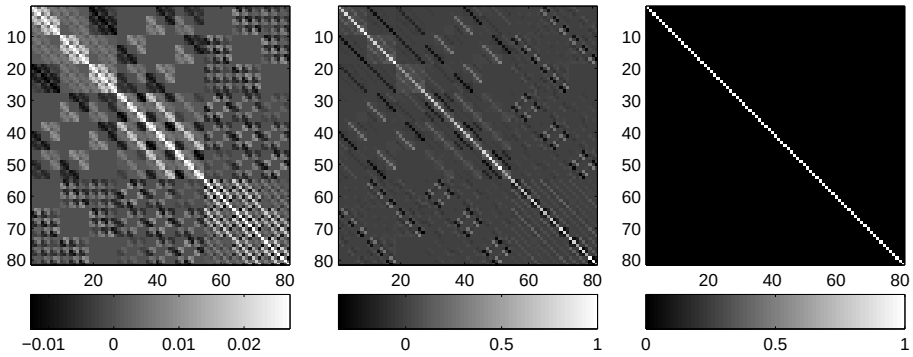

Fig. 1. LC ($k = 2$) local optical flow estimation from identical cuboidal $3 \times 3 \times 3$ motion neighborhoods ($m = 27$) is considered. Left: Identical covariance matrix $C \in \mathbb{R}^{81 \times 81}$ of the vectors $\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])$, with structure owing to FIR masks. Middle: Covariance matrix of $\mathrm{vec}(W_L[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]W_R^T)$. Right: Unity matrix.

Given ETLS estimates $\hat{\boldsymbol{p}} : P \to \mathbb{R}^k$, the residuals are given by the mapping $\hat{\boldsymbol{r}} : P \to \mathbb{R}^m$ such that

$$\forall \boldsymbol{x} \in P : \quad \hat{\boldsymbol{r}}(\boldsymbol{x}) := W_L[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]W_R^T \begin{pmatrix} \hat{\boldsymbol{p}}(\boldsymbol{x}) \\ -1 \end{pmatrix} \quad . \tag{12}$$

In principle, the theoretical PDF of these residuals is determined by the joint PDF of the entries of $A_g(\boldsymbol{x})$ and $\boldsymbol{b}_g(\boldsymbol{x})$. The latter is obtained from the EIV model, the motion models, and the derivative operators. However, there is a direct influence to the residual PDF by the factor $[A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]$ as well as an indirect influence by the PDF of the estimates $\hat{\boldsymbol{p}}(\boldsymbol{x})$. In the following, we assume $\hat{\boldsymbol{p}}$ to be deterministic. Then, the residuals (12), expressed as

$$\forall \boldsymbol{x} \in P : \quad \hat{\boldsymbol{r}}(\boldsymbol{x}) = \underbrace{\left(\begin{pmatrix} \hat{\boldsymbol{p}}(\boldsymbol{x}) \\ -1 \end{pmatrix}^T W_R \otimes W_L \right)}_{=:\, R(\boldsymbol{x})} \mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]) \quad , \tag{13}$$

are obtained from the deterministic linear mapping defined by the matrix $R(\boldsymbol{x})$, applied to the vector $\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])$ of which the covariance matrix (11) is known. The covariance matrices of the residual vectors are therefore given by $C_r : P \to \mathbb{R}^{m \times m}$ such that $\forall \boldsymbol{x} \in P : C_r(\boldsymbol{x}) := \mathrm{cov}(\hat{\boldsymbol{r}}(\boldsymbol{x})) = R(\boldsymbol{x})CR^T(\boldsymbol{x})$. From the Cholesky factorizations $L : P \to \mathbb{R}^{m \times m}$ such that $LL^T = C_r$ follows that $\hat{\boldsymbol{s}} := L^{-1}\hat{\boldsymbol{r}}$ is decorrelated i.e.,

$$\forall \boldsymbol{x} \in P : \quad \mathrm{cov}(\hat{\boldsymbol{s}}(\boldsymbol{x})) = \mathbb{1}_m \quad , \tag{14}$$

while $\forall \boldsymbol{x} \in P : \mathbb{E}(\hat{\boldsymbol{s}}(\boldsymbol{x})) = L^{-1}(\boldsymbol{x})R(\boldsymbol{x})\mathbb{E}(\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})]))$. From (10) follows $\mathbb{E}(\mathrm{vec}([A_g(\boldsymbol{x}), \boldsymbol{b}_g(\boldsymbol{x})])) = \mathrm{vec}([A_\tau(\boldsymbol{x}), \boldsymbol{b}_\tau(\boldsymbol{x})]) + \mathbb{E}(\mathrm{vec}([A_\epsilon(\boldsymbol{x}), \boldsymbol{b}_\epsilon(\boldsymbol{x})]))$. Under the

assumption that the entries of $[A_\epsilon(\boldsymbol{x}), \boldsymbol{b}_\epsilon(\boldsymbol{x})]$ have zero mean, it follows

$$\forall \boldsymbol{x} \in P: \quad \mathbb{E}(\hat{\boldsymbol{s}}(\boldsymbol{x})) = L^{-1}(\boldsymbol{x}) W_L [A_\tau(\boldsymbol{x}), \boldsymbol{b}_\tau(\boldsymbol{x})] W_R^T \begin{pmatrix} \hat{\boldsymbol{p}}(\boldsymbol{x}) \\ -1 \end{pmatrix} . \tag{15}$$

In practice, it depends on the appropriateness of the parametric model as well as on the empirical distribution of noise whether or not

$$[A_\tau(\boldsymbol{x}), \boldsymbol{b}_\tau(\boldsymbol{x})] W_R^T \begin{pmatrix} \hat{\boldsymbol{p}}(\boldsymbol{x}) \\ -1 \end{pmatrix} = 0 \tag{16}$$

holds, in which case it follows from (15) that

$$\mathbb{E}(\hat{\boldsymbol{s}}(\boldsymbol{x})) = 0. \tag{17}$$

If, in addition, the noise in the grayvalues is i.i.d. according to a normal distribution with known variance then, it follows from (14) and (17) that the entries of the decorrelated residual vector $\hat{\boldsymbol{s}}(\boldsymbol{x})$ from ETLS estimation form a set of independent standard normally distributed random variables. We therefore suggest to adaptively test this set of residuals, for each pixel $\boldsymbol{x} \in P$, to be standard normally distributed. Deviations from the standard normal distribution are then taken as indications of inappropriateness of the motion model. We have therefor employed the Kolmogorov-Smirnov test, Pearson's χ^2 test, the Anderson-Darling test as well as the absolute difference of the vectors of the first 2,3,4 and 5 non-centered moments of the empirical and theoretical distribution.

5 Application & Results

In order to allow for motion estimation on real world video data or standard benchmark sequences such as the Yosemite sequence, the model selector has to be incorporated into a motion estimation framework that is capable of handling the aperture problem. If the model selector was examined separately on real data, the aperture problem as well as a possible incoherence of the true displacement and optical flow due to changes in illumination would distract from properties of the model selection process. Benchmark results from motion estimation frameworks on the other hand include effects from all components, be it confidence measures, motion segmentation or noise estimation techniques. Hence, in order to specifically test for properties of the model selector, we generated a variety of sequences from given two-dimensional displacement fields by warping of an initial frame. Grayvalue structure on multiple scales was introduced to this frame in order to avoid the aperture problem. Zero mean Gaussian noise was added to the sequences. In this special case, no framework is needed. A systematic study of the discrepancy between the true displacement field and optical flow estimates on this data can be trusted to indicate precisely the properties of the model selection process. Results from model selection are shown in Figure 2 for a sequence featuring motion patterns of different parametric form (top) as well as for a simulated continuous current (bottom). From the different shading in

Figure 2b, it can be seen that model selection is in accordance with the true displacement field. Motion patterns are identified correctly. The incidental choice of overly complex models is explained by the fact that a higher order model with the additional parameters correctly estimated as zero cannot be distinguished from the simpler model by means of residual analysis. The most complex model is correctly selected at motion discontinuities. Apart from the identification of motion patterns, an important application of model selection is to limit model complexity with respect to noise. While a complex model is appropriate at low noise levels model complexity has to be controlled with increasing noise in order to avoid over-fitting. Figure 3 shows the effect of model selection on the mean deviation of optical flow estimates from the true displacement field at 0.5%, 3% and 10% noise added to the continuous current sequence. The means were taken over the entire sequence. Amplitude and direction of the deviation were calculated together with the well-known angular error. Global choices of a single model (left part of each bar graph) are compared to adaptive model selection per pixel. In the latter case, errors from different (the selected) models are cumulated in the mean. Regardless which of the models LC (C), LSS (S), LRD (R), LAF (A) and LPL (P) is chosen globally, a situation exists in which the error is intolerably high compared to another model. This effect from global model assumptions causes a problem in applications with complex motion patterns and changing noise where a complex model, although needed to yield good estimates at low noise, performs weak at increasing noise levels. Considering model selection by Pearson's χ^2 test (x) or the Anderson-Darling test (a) which are the best performing model selectors, it can be seen from the top row of figure 3 that these adaptive estimators yield errors comparable to those obtained from the best global choices. The reason is that the residual-based model selector precisely excludes inappropriate models. At 3% and 10% noise, global model choices exist which are favorable to adaptive model selection. Nevertheless, the discrepancy is tolerable if the aim is to exclude the most complex models which perform poorly in this case. Results in model complexity control hence prove to be useful for applications where noise as well as the complexity of the displacement field vary. The slight limitation in model complexity control at high noise is due to the idealization of parameter estimates to be deterministic. However, this is not a principle drawback of residual analysis in motion model selection and effects are tolerable.

6 Conclusion & Perspectives

We have demonstrated that statistical testing of regression residuals is a viable approach to the model selection problem in motion estimation. The residual-based model selector is capable of precisely excluding inappropriate models. Its performance in model complexity control makes this model selector a particularly useful tool for applications where noise as well as the complexity of the displacement field vary. Slight limitations of the proposed method with tolerable effects are due to the idealization of equilibrated total least squares estimates

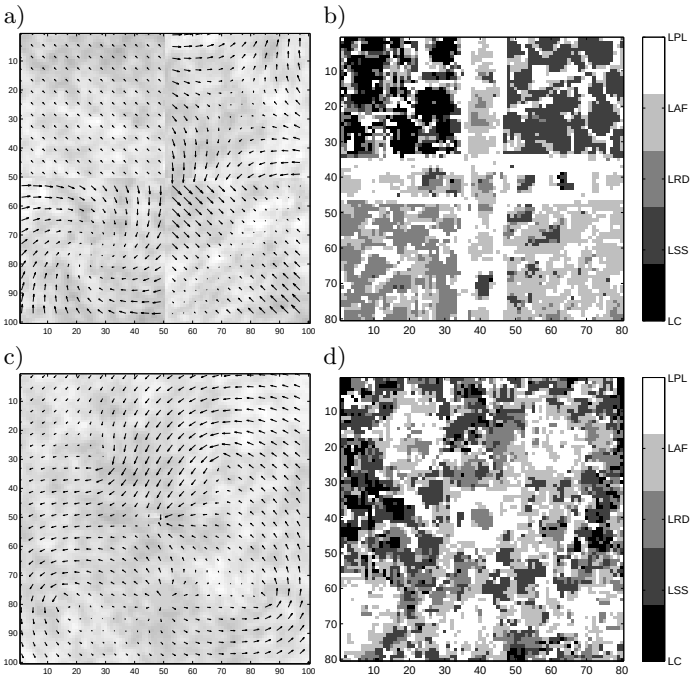

Fig. 2. Model selection from 11x11x3 motion neighborhoods of simulated sequences at 0.5% noise-to-signal amplitude ratio by comparison of 5 moments of the residual distribution. a) displacement field of the types sequence, b) according model selection, c) displacement field of the current sequence, d) according model selection.

to be deterministic. The incorporation of approximations to the distribution of TLS estimates is a promising starting point for future research.

References

1. Black, M.J., Jepson, A.: Estimating multiple independent motions in segmented images using parametric models with local deformations. In: IEEE Workshop on Motion of Non-Rigid and Articulated Objects. (1994) 220–227
2. Gheissari, N., Bab-Hadiashar, A., Suter, D.: Parametric model-based motion segmentation using surface selection criterion. Computer Vision and Image Understanding **102** (2006) 214–226
3. Akaike, H.: A new look at statistical model identification. IEEE Transactions on Automaic Control **19** (1974) 716–723
4. Schwarz, G.: Estimating the dimension of a model. Annals of Statistics **6**(461) (1978) 464
5. Wechsler, H., Duric, Z., Li, F.Y., Cherkassky, V.: Motion estimation using statistical learning theory. PAMI **26**(4) (2004) 466–478
6. Cootes, T.F., Thacker, N., Taylor, C.J.: Automatic model selection by modelling the distribution of residuals. In: ECCV 2002, LNCS 2353, Springer (2002) 621–635

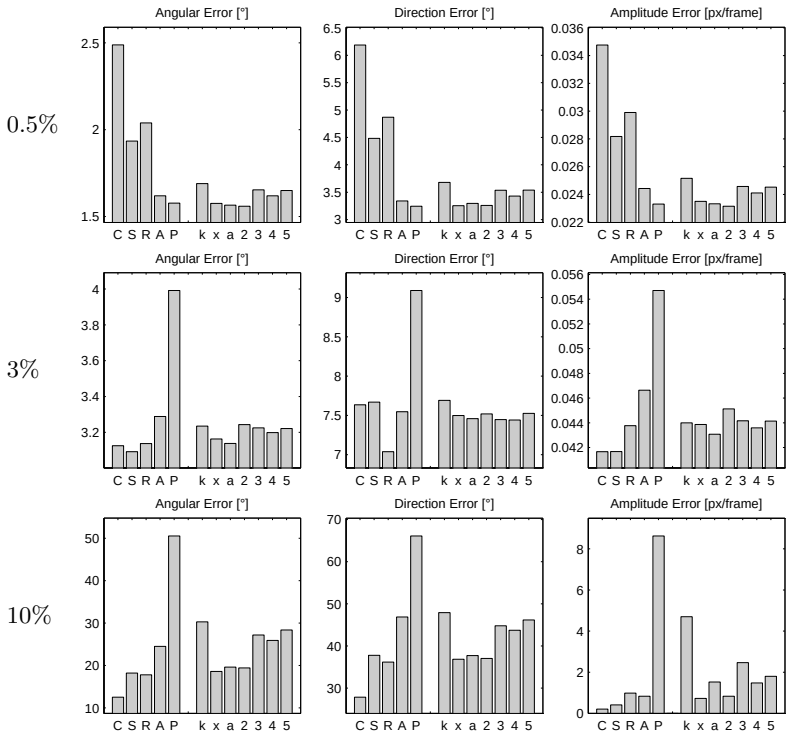

Fig. 3. Mean errors of optical flow estimation from 9x9x3 motion neighborhoods of a simulated continuous 2D current at different noise levels (top to bottom). Results for the models LC (C), LSS (S), LRD (R), LAF (A) and LPL (P), as well as for the adaptively selected models chosen by the KS test (k), the χ^2 test (x), the AD test (a) and the absolute difference of the vectors of the first 2,3,4 and 5 moments of the empirical and the theoretical distribution.

7. Jähne, B.: Digital Image Processing. Concepts, Algorithms and Scientific Applications. 6th edn. Springer, Berlin (2005)
8. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: DARPA Image Understanding Workshop. (1981) 121–130
9. Scharr, H.: Optimale Operatoren in der Digitalen Bildverarbeitung. PhD thesis, University of Heidelberg, Heidelberg, Germany (2000)
10. Van Huffel, S., Vandewalle, J.: The Total Least Squares Problem: Computational Aspects and Analysis. SIAM, Philadelphia (1991)
11. Mühlich, M.: Estimation in Projective Spaces and Application in Computer Vision. PhD thesis, Johann Wolfgang Goethe Universität Frankfurt am Main (2005)
12. Andres, B.: Model selection in optical flow-based motion estimation by means of residual analysis. Diploma thesis, University of Heidelberg (2007)