

# Robust selection of parametric motion models in image sequences

Patrick Bouthemy, Bertha Mayela Toledo Acosta, Bernard Delyon

► **To cite this version:**

Patrick Bouthemy, Bertha Mayela Toledo Acosta, Bernard Delyon. Robust selection of parametric motion models in image sequences. 2016 IEEE International Conference on Image Processing (ICIP), Oct 2016, Phoenix, United States. pp.3743 - 3747, 2016, <10.1109/ICIP.2016.7533059>. <hal-01400895>

**HAL Id: hal-01400895**

**<https://hal.inria.fr/hal-01400895>**

Submitted on 22 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ROBUST SELECTION OF PARAMETRIC MOTION MODELS IN IMAGE SEQUENCES

*Patrick Bouthemy, Bertha Mayela Toledo Acosta*

*Bernard Delyon*

Inria  
Centre Rennes - Bretagne Atlantique  
35042 Rennes cedex, France

University of Rennes 1  
IRMAR  
35042 Rennes cedex, France

## ABSTRACT

Parametric motion models are commonly used in image sequence analysis for different tasks. A robust estimation framework is usually required to reliably compute the motion model. The choice of the right model is also important. However, dealing simultaneously with both issues remains an open question. We propose a robust motion model selection method with two variants, which relies on the Fisher test. We also derive an interpretation of it as a robust Mallows'  $C_P$  criterion. The resulting criterion is straightforward to compute. We have conducted a comparative experimental evaluation on different image sequences demonstrating the interest and the efficiency of the proposed method.

*Index Terms*— robust model selection, parametric motion model, image sequence.

## 1. INTRODUCTION

Adopting 2D parametric motion models is a common practice in image sequence analysis, for video stabilization [16], image stitching [27], optical flow computation [4, 7, 29], motion segmentation [6], tracking [26, 30], or action recognition in videos [12], to name a few. Two key issues then arise: which motion model and which estimation framework?

Using a statistical information criterion such as Akaike information criterion (AIC) or Bayesian information criterion (BIC) [2, 5, 25], is the most classical way to deal with the motion model selection issue [8, 10, 28]. It amounts to add to the conditional likelihood (or model fit) a penalty on the model complexity or dimension given by the number of the model parameters. The likelihood term is quadratic accounting for a Gaussian distribution of the residuals involved in the regression issue. Let us add the Mallows'  $C_P$  criterion [15] and the Minimum Description Length criterion (MDL) [21] respectively equivalent to AIC and BIC under certain hypotheses.

On the other hand, adopting a robust motion estimation framework is unavoidable to cope with the possible presence of outliers within the model estimation support, due to occlusions, local motion or any local violation of the required assumptions associated with motion computation. However,

combining model selection and robust estimation for parametric motion computation has rarely been investigated [28]. In this paper, we propose a statistical criterion for robust motion model selection based on the Fisher test. An interpretation related to the Mallows'  $C_P$  criterion [15] is also provided.

The rest of the paper is organized as follows. Section 2 is devoted to related work. Section 3 briefly recalls the robust estimation of the motion model parameters and formulates the model fit evaluation. Section 4 describes the robust motion model selection method we have developed. Experimental results and an objective comparison are reported in Section 5. Concluding remarks are given in Section 6.

## 2. RELATED WORK

Robust model selection has been explored in the statistics literature along several directions [1, 18, 20, 22]. In [22], a robust extension of AIC called AICR is developed by substituting a general robust estimator  $\rho$  of the model parameters  $\theta$  for the maximum likelihood estimator. M-estimators are incorporated in BIC and the asymptotic performance is studied in [14]. A special case is the use of the Huber robust function [11], leading to the RBIC criterion. The Mallows'  $C_P$  criterion is revisited in [23] to yield a robust version. However, these approaches were not explicitly concerned with jointly maximizing the inlier set size.

In contrast, only a few investigations of that type have been undertaken regarding motion estimation and segmentation. In [3], the authors designed a global energy function for both the robust estimation of mixture models and the validation of a MDL criterion. The overall goal is to get a layering representation of the moving content of an image sequence. The MDL encoding acts on the overall cost of the representation comprising the number of layers, residuals and motion parameters. However, the primary purpose was parsimonious motion segmentation, and not motion model selection per se. In [28], a robust extension of the Geometric Information Criterion (GIC) [13] is proposed in the vein of AICR. It was applied to the selection of the 3D geometric transformation attached to a rigid motion and estimated through the matching of image interest points. Geometrical and physical con-

straints are also explored in [9] for image motion segmentation, with the so-called surface selection criterion (SSC) primarily designed by the authors for range data segmentation. Better performance is reported than with several information criteria, but the use of SSC in this work is comparable to a regularization approach. We propose a different approach for robust motion model selection based on the Fisher statistic.

### 3. ROBUST MOTION MODEL ESTIMATION

In this paper, we are interested in selecting the most relevant parametric motion model accounting for the dominant image motion. The latter is usually due to the camera motion. Computing the dominant motion has numerous applications such as image stabilization, background subtraction for a free moving camera, image stitching or simply image registration. However, the proposed scheme can be applied to other motion configurations as well, for instance to select the right motion models on image regions.

#### 3.1. Computation of motion model parameters

We consider a set of 2D polynomial motion models. They will be precisely defined in Section 4.3. Let  $\theta_m$  denote the parameters of model  $m$ , that is, the polynomial coefficients for the two components of the velocity vector. The full model will be denoted by  $M$ , if we have  $M$  models to test.  $\mathbf{w}_{\theta_m}(p)$  is the velocity vector supplied by the motion model  $m$  at point  $p = (x, y)$  of the image domain  $\Omega$ .

To estimate the parameters of the motion model, we exploit the usual brightness constancy assumption [7], leading to the linear regression equation relating the velocity vector and the spatiotemporal derivatives of the image intensity  $I$ :

$$\nabla I(p) \cdot \mathbf{w}_{\theta_m}(p) + I_t(p) = 0. \quad (1)$$

Let us denote  $r_{\theta_m}(p)$  the left member of (1). Once we compute an estimate  $\hat{\theta}_m$  of the motion model parameters, we get the residuals  $r_{\hat{\theta}_m}(p)$  for  $p \in \Omega$  measuring the discrepancy between the input data and the estimated motion model. The estimation of the motion model parameters is formulated as:

$$\hat{\theta}_m = \arg \min_{\theta_m} \sum_{p \in \Omega} \rho(r_{\theta_m}(p)), \quad (2)$$

where  $\rho$  denotes a robust penalty function. As examples, the Lorentzian function is used in [4], in [26] the Hampel estimator is preferred, and in [19] the Tukey's function is adopted.

#### 3.2. Model fit evaluation

To evaluate how the estimated motion model fits the input data over the associated inlier set, we consider the residual sum of squares (RSS) obtained for the robustly estimated parameters  $\hat{\theta}_m$  of the motion model  $m$  given by:

$$RSS_m = \sum_{p \in \mathcal{I}_m} r_{\hat{\theta}_m}^2(p), \quad (3)$$

where  $\mathcal{I}_m$  represents the set of inliers associated with the estimated motion model  $m$ . The residual is in fact defined by  $r_{\hat{\theta}_m}(p) = I(p + \mathbf{w}_{\hat{\theta}_m}(p), t + 1) - I(p, t)$ . We compute  $RSS_m$  on the inlier set  $\mathcal{I}_m$  and not on the overall domain  $\Omega$ , to obtain the model fit evaluation precisely on the subset of points whose motion conforms with the estimated motion model.

Furthermore, we introduce the expression  $RSS_m^+$ , which represents the residual sum of squares computed over the inlier set  $\mathcal{I}_m$  attached to model  $m$ , but for the full model  $M$ :

$$RSS_m^+ = \sum_{p \in \mathcal{I}_m} r_{\hat{\theta}_M}^2(p). \quad (4)$$

Solving for (2) with the publicly available Motion2D<sup>1</sup> software [19], amounts to apply the Iteratively Reweighted Least Squares algorithm within a coarse-to-fine framework. At convergence, the final weights  $\alpha_m(p)$ ,  $p \in \Omega$ , are given by  $\alpha_m(p) = \frac{\psi(r_{\hat{\theta}_m}(p))}{r_{\hat{\theta}_m}(p)}$ , where the influence function  $\psi(\cdot)$  is the derivative of the robust function  $\rho(\cdot)$ .  $p$  belongs to  $\mathcal{I}_m$  if its final weight  $\alpha_m(p)$  is greater than  $\tau$ , with  $\alpha_m(p)$  normalized within  $[0, 1]$  and  $\tau$  set to 0.6 in practice.

### 4. ROBUST MOTION MODEL SELECTION

As in [8], we first adopt a two-class hypothesis test approach. This is motivated by the fact that we are dealing with an unnested set of parametric motion models. For instance, both the rotation and the scaling models involve three parameters as described in Section 4.3, but respectively account for quite different motions. In addition, we aim to select the model  $m$  which explains the motion of the maximum number of points in the image domain  $\Omega$ , i.e., with the largest possible inlier set. Let us point out that retaining the least complex motion model while maximizing the size of the inlier set may be contradictory, which is a key issue specific to robust model selection.

#### 4.1. Fisher statistic

First, we want to compare any model  $m$  of the set of tested models to the full model  $M$ . To this end, we resort to the Fisher statistic [24] formulated in our case as:

$$\mathcal{F}(m) = \frac{(RSS_m - RSS_m^+) / (q_M - q_m)}{RSS_m^+ / (|\mathcal{I}_m| - q_M)}, \quad (5)$$

where  $|\cdot|$  designates the set cardinality,  $q_m$  represents the number of parameters of model  $m$ . Both  $RSS_m$  and  $RSS_m^+$  are evaluated on the inlier set  $\mathcal{I}_m$  attached to the estimated model  $m$ . The denominator can be interpreted as a non-biased empirical estimate of the full model variance computed on  $\mathcal{I}_m$ , which will be denoted by:

$$\hat{\sigma}_M^2(\mathcal{I}_m) = \frac{RSS_m^+}{|\mathcal{I}_m| - q_M}. \quad (6)$$

This statistic allows us to decide whether model  $m$  is a more significant representation of the unknown true motion

<sup>1</sup><http://www.irisa.fr/vista/Motion2D/>

than the full model  $M$  over  $\mathcal{I}_m$  which is the validity domain of model  $m$  in  $\Omega$ . However, it will supply all the models  $m$  of that type. We need to take into account the dimension  $q_m$  of model  $m$  to further select the right one.

## 4.2. Model selection criterion

Starting from (5), and penalizing the complexity of the model expressed by the number  $q_m$  of model parameters, we define:

$$\mathcal{S}_1(m) = \mathcal{F}(m)(q_M - q_m) + 2q_m. \quad (7)$$

Under the assumption of validity of model  $m$ ,  $\mathcal{F}(m)$  follows a Fisher distribution  $F(q_M - q_m, |\mathcal{I}_m| - q_M)$ . Then, the first term of the right member of (7) (approximately) follows a  $\chi^2$  distribution with  $q_M - q_m$  degrees of freedom. We can now express the test for selecting the best motion model  $\hat{m}$ :

$$\hat{m} = \arg \min_m \mathcal{S}_1(m). \quad (8)$$

The theoretical behaviour of this test can be qualitatively described as follows.  $\mathcal{S}_1(m)$  is supposed to decrease when evaluating in turn the first successive models in decreasing (or equivalently increasing) complexity order up to the optimal model  $m^*$ , and then to increase for the subsequent models.

We design a second version of the robust model selection criterion, by incorporating the number of inliers in the model complexity penalization as in the BIC criterion, that is:

$$\mathcal{S}_2(m) = \mathcal{F}(m)(q_M - q_m) + 2 \log(|\mathcal{I}_m|)q_m. \quad (9)$$

We now provide another interpretation of the statistical criterion (7). Let us first make  $\hat{\sigma}_M^2(\mathcal{I}_m)$  appear in the expression of  $\mathcal{S}_1(m)$  as follows:

$$\mathcal{S}_1(m) = \frac{(RSS_m - RSS_m^+)}{\hat{\sigma}_M^2(\mathcal{I}_m)} + 2q_m, \quad (10)$$

By exploiting (3) and (6), it can be further developed into:

$$\mathcal{S}_1(m) = \frac{1}{\hat{\sigma}_M^2(\mathcal{I}_m)} \sum_{p \in \mathcal{I}_m} r_{\hat{\theta}_m}^2(p) - |\mathcal{I}_m| + q_M + 2q_m. \quad (11)$$

If we neglect  $q_M$  which is a constant term for the test (8), expression (11) can be viewed as the Mallows'  $C_P$  criterion computed over the inlier set attached to model  $m$  with  $|\mathcal{I}_m|$  as the number of observations. Then, our test (8) could also be interpreted as a robust version of the Mallows'  $C_P$  criterion.

Let us point out that (11) explicitly involves the aforementioned trade-off between maximizing the size  $|\mathcal{I}_m|$  of the inlier set and minimizing the complexity (i.e., the number  $q_m$  of parameters) of the selected motion model. In contrast, in existing robust model selection criteria as AICR or RBIC:

$$\mathcal{R}(m) = \sum_{p \in \Omega} \rho(r_{\theta_m}(p)) + \beta q_m, \quad (12)$$

the model selection is only implicitly influenced by the size of the inlier set attached to model  $m$  through the values of the robust function  $\rho(\cdot)$  at the outlier points. Hence, the impact depends on the asymptotic behaviour of the robust function. The same holds for [23] where in addition the penalty term requires additional expensive computation to be evaluated.

**Table 1.** Set of 2D polynomial motion models

Motion model	Dim.	Expression
Translation (T)	2	$\mathbf{w}_\theta = (a_1, a_4)$
Tr.+Rotation (TR)	3	$\mathbf{w}_\theta(p) = (a_1 + a_3y, a_4 - a_3x)$
Tr.+Scal. (TS)	3	$\mathbf{w}_\theta(p) = (a_1 + a_2x, a_4 + a_2y)$
Tr.+Rot.+Scal. (TRS)	4	$\mathbf{w}_\theta(p) = (a_1 + a_2x + a_3y, a_4 - a_3x + a_2y)$
Full affine (FA)	6	$\mathbf{w}_\theta(p) = (a_1 + a_2x + a_3y, a_4 + a_5x + a_6y)$
Pan-tilt (PT)	2	$\mathbf{w}_\theta(p) = (a_1 + a_1x^2 + a_4xy, a_4 + a_1xy + a_4y^2)$
Pan-tilt-zoom (PTZ)	3	$\mathbf{w}_\theta(p) = (a_1 + a_2x + a_1x^2 + a_4xy, a_4 + a_2y + a_1xy + a_4y^2)$
Planar surface rigid motion (PSRM)	8	$\mathbf{w}_\theta(p) = (a_1 + a_2x + a_3y + a_7x^2 + a_8xy, a_4 + a_5x + a_6y + a_7xy + a_8y^2)$
Full quadratic (FQ)	12	$\mathbf{w}_\theta(p) = (a_1 + a_2x + a_3y + a_7x^2 + a_8xy + a_9y^2, a_4 + a_5x + a_6y + a_{10}x^2 + a_{11}xy + a_{12}y^2)$

## 4.3. Set of motion models

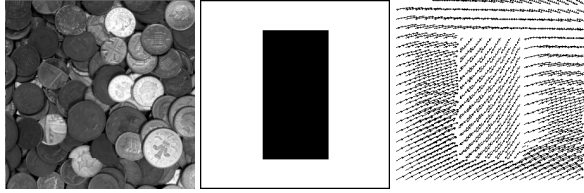
We are dealing with 2D polynomial motion models ranging from translation (polynomial of degree 0) to quadratic models (polynomials of degree 2), including different affine models (polynomials of degree 1). They are forming a set of models which is only partly nested. The model complexity ranges from dimension 2 to dimension 12. The full set of motion models is given in Table 1 with their main features. The 8-parameter quadratic motion model corresponds to a rigid motion between a planar scene and the camera. The 2-parameter quadratic model accounts for a pan-tilt camera motion.

## 5. EXPERIMENTAL RESULTS

We first conducted a comparative objective evaluation of the proposed selection criteria on a set of synthetic examples to quantitatively assess performances. As illustrated in Fig.1, we applied to a real image a velocity field to create an image pair with known motion, i.e., with available ground truth. The velocity field consists of two parametric subfields chosen in the list given in Table 1. The first parametric motion subfield is the dominant motion, and the outliers, forming a rectangular region in the middle of the image, undergo the second one.

We generated 1200 image pairs. The first 400 ones involve a translation (T) motion model as dominant motion and a full affine (FA) as secondary motion, the next 400 ones a FA motion model and a PSRM model respectively, and the last 400 ones a PSRM model and a T model respectively. We randomly select values in  $[-0.11, 0.11]$  for the parameters  $a_1$  and  $a_4$  of the T model, in  $[-0.05, 0.05]$ ,  $[-0.02, 0.02]$ ,  $[-0.04, 0.04]$ , and  $[-0.03, 0.03]$  respectively for the first-order parameters  $a_2, a_3, a_5$  and  $a_6$  of the FA model and the PSRM model, while parameters  $a_7$  and  $a_8$  are fixed to 0.0004 and 0.0002 respectively. In the FA model,  $a_1$  and  $a_4$  are fixed to 0.1 and -0.1 resp., while they resp. vary within  $[-0.5, 0.5]$  and  $[-1, 1]$  for the PSRM model.

The true dominant motion T is selected by criterion  $\mathcal{S}_1$  in 19.5% of the first 400 image pairs, in 55.75% of them by criterion  $\mathcal{S}_2$ , and in 7.25% of them by RBIC, respectively. For the next 400 image pairs, the true FA model is correctly selected respectively in 30.75% of them for  $\mathcal{S}_1$ , 47% for  $\mathcal{S}_2$ , and 54.75% for RBIC. Finally, the true PSRM model is correctly chosen by  $\mathcal{S}_1$  for 70.75% of the last 400 image pairs, for 82.75% of them by  $\mathcal{S}_2$ , and for 23.25% of them by RBIC.



**Fig. 1.** From left to right: the input image, the outlier subset (in black) in the middle of the image, typical velocity field applied to the input image formed by a dominant motion and a secondary one.

Clearly, criterion  $S_2$  outperforms the two others. Table 2 details the scores obtained with  $S_2$  for all the tested models and for the three subsets of experiments. As expected, selection errors are spread and concern only more complex models than the true one.

Tested models	True dominant motion models		
	T	FA	PSRM
T	<b>55.75</b>	0	0
TR	10	0	0
TS	4.75	0	0
TRS	1.5	0	0
FA	0.25	<b>47</b>	0
PT	16.25	1.75	0
PTZ	4.25	0	0
PSRM	1.5	31	<b>82.75</b>
FQ	5.75	20.25	17.25

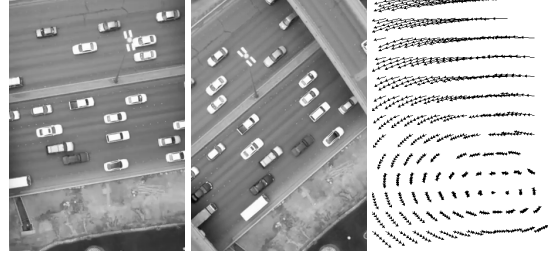
**Table 2.** Scores obtained with criterion  $S_2$  for all the tested models and for the three subsets of experiments.

We now report results on two different real image sequences. The first one depicts a traffic scene acquired from an airborne camera. It can be observed that the scene is almost planar and the main camera motion component is a rotation around the axis of view, which advocates the PRSM model. The sequence comprises 49 images, that is, 48 successive image pairs. Table 3 contains the model selection results provided by the M-likelihood (L) alone (i.e., the first term of (11)), our criteria  $S_1$  and  $S_2$ , and RBIC. In contrast to the Gaussian case, the M-likelihood does not necessarily decrease with the model complexity, since it also depends on the inlier set which may shrink for simple models. Clearly, a penalized likelihood is required.  $S_1$  and  $S_2$  exhibit close satisfactory behaviour, while RBIC always selects the full model. Let us stress that all the motion models selected by  $S_1$  and  $S_2$  over the sequence involve a rotation component. By the way, TRS and FA models can be considered as reasonable choices.

	T	TS	TR	TRS	FA	PT	PTZ	PSRM	FQ
RBIC	0	0	0	0	0	0	0	0	48
L	11	8	2	5	5	1	0	16	0
$S_1$	0	0	1	10	14	0	0	20	3
$S_2$	0	0	1	11	15	0	0	18	3

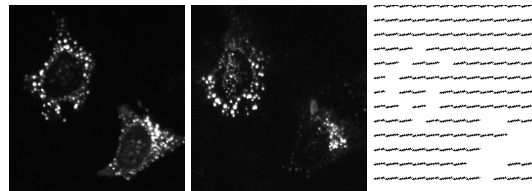
**Table 3.** Motion models selected by the tested criteria over the sequence of Fig.2 (48 image pairs).

In the second real experiment, we deal with a light fluorescence microscopy sequence containing 32 frames (Fig.3).



**Fig. 2.** First and last frames of the traffic scene sequence, and the computed dominant flow between frames 1 and 2.

The global motion is due to a shift of the acquisition setup, and it was assessed as a global translation by the biologist. Two cells are moving in the sequence, yielding local motions. Here, we deal with the overall model selection within the sequence and we report in Table 4 the values of the criteria for the 9 tested motion models. The right motion model T is correctly selected by  $S_1$  and  $S_2$ , but not by RBIC which remains stuck in the full model.



**Fig. 3.** Two images out of the fluorescence microscopy sequence and the computed dominant motion.

	RBIC	$S_1$	$S_2$
T	260439.74	<b>38127519.72</b>	<b>38128065.67</b>
TS	257574.24	38263744.05	38264562.99
TR	258755.35	38359977.80	38360796.72
TRS	255532.74	38505565.00	38506656.96
FA	257111.36	39094278.00	39095915.92
PT	262053.84	39642634.06	39643725.97
PTZ	257330.38	38462489.79	38463854.67
PSR	256541.99	39211062.78	39213246.69
FQ	<b>254030.65</b>	39789352.81	39792628.89

**Table 4.** Values of RBIC,  $S_1$  and  $S_2$  for 9 tested motion models, averaged over the sequence of Fig.3.

## 6. CONCLUSION

We have proposed a new approach for robust model selection in the framework of motion model estimation. It relies on the Fisher statistic and explicitly tackles the tradeoff between the size of the inlier set (to be maximized) and the complexity of the motion model (to be minimized). In addition, this new criterion can be viewed as a proposition for a robust Mallows'  $C_P$  criterion. We propose two variants  $S_1$  and  $S_2$ , both easy to compute. Experiments on synthetic and real sequences along with comparison with RBIC, demonstrate that our criterion  $S_2$  provides superior performance.

## 7. REFERENCES

- [1] C. Agostinelli. Robust model selection in regression via weighted likelihood methodology. *Statistics & Probability Letters*, 56:289–300, 2002.
- [2] H. Akaike. A new look at the statistical model identification. *IEEE Trans. on Automatic Control*, 19(6):716–723, Dec. 1974.
- [3] S. Ayer and H. S. Sawhney. Layered representation of motion video using robust maximum likelihood estimation of mixture models and MDL encoding. In *ICCV*, Cambridge, Jun. 1995.
- [4] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, 1996.
- [5] K. P. Burnham and D.R. Anderson. *Model selection and multimodel inference: A practical information-theoretic approach (2nd ed.)*. Springer-Verlag, 2002.
- [6] D. Cremers and S. Soatto. Motion competition: A variational approach to piecewise parametric motion segmentation. *Int. J. of Computer Vision*, 62(3):249–265, May 2005.
- [7] D. Fortun, P. Bouthemy, and C. Kervrann. Optic flow modeling and computation: a survey. *Computer Vision and Image Understanding*, 134:1-21, May 2015.
- [8] E. François and P. Bouthemy. Derivation of qualitative information in motion analysis. *Image and Vision Computing Journal*, 8(4):279–287, Nov. 1990.
- [9] N. Gheissari and A. Bab-Hadiashar. Motion analysis: Model selection and motion segmentation. In *ICIAP*, Mantova, Sep. 2003.
- [10] N. Gheissari and A. Bab-Hadiashar. A comparative study of model selection criteria for computer vision applications. *Image and Vision Computing*, 26(12):1636–1649, Dec. 2008.
- [11] P.J. Huber. *Robust Statistics*. Wiley Series in Probability and Statistics, Wiley, 1981.
- [12] M. Jain, H. Jégou, and P. Bouthemy. Improved motion description for action classification. *Frontiers in ICT: Computer Image Analysis*, doi: 10.3389/fict.2015.00028, Jan. 2016.
- [13] K. Kanatani. Geometric information criterion for model selection. *Int. J. of Computer Vision*, 26(3):171–189, 1998.
- [14] J.A.F. Machado. Robust model selection and M-estimation. *Econometric Theory*, 9:478–493, 1993.
- [15] C. Mallows. Some comments on Cp. *Technometrics*, 15:661-675, 1973.
- [16] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H.-Y. Shum. Full-frame video stabilization with motion inpainting. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(7):1150–1163, Jul. 2006.
- [17] P. Meer. Robust techniques for computer vision. In *Emerging Topics in Computer Vision*, G. Medioni and S. B. Kang (eds.), Prentice Hall, 107-190, 2004.
- [18] S. Müller and A. H. Welsh. Outlier robust model selection in linear regression. *Journal of the American Statistical Association*, 100(472):1297–1310, Dec. 2005.
- [19] J.-M. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models. *J. Visual Communication and Image Representation*, 6(4):348–369, Dec. 1995.
- [20] G. Qian and H.R. Künsch. On model selection via stochastic complexity in robust linear regression. *Journal of Statistical Planning and Inference*, 75:91–116, 1998.
- [21] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.
- [22] E. Ronchetti. Robust model selection in regression. *Statistics & Probability Letters*, 3:21–23, 1985.
- [23] E. Ronchetti and R.G. Staudte. A robust version of Mallows’s Cp. *Journal of The American Statistical Association*, 89(426):550–559, Jun. 1994.
- [24] G.G. Roussas. *A Course in Mathematical Statistics*. Academic Press, 1997.
- [25] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [26] T. Senst, V. Eiselen and T. Sikora. Robust local optical flow for feature tracking. *IEEE Trans. on Circuits and Systems for Video Technology*, 22(9):1377–1387, Sep. 2012.
- [27] R. Szeliski. Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Computer Vision*, 2(1):1-104, Dec. 2006.
- [28] P. Torr. An assessment of information criteria for motion model selection. In *CVPR*, San Juan, Jun. 1997.
- [29] J. Yang and H. Li. Dense, accurate optical flow estimation with piecewise parametric model. In *CVPR*, Boston, June 2015.
- [30] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Computing Surveys*, 38(4), Dec. 2006.