COMPARISON OF TWO ALGORITHMS FOR ROBUST M-ESTIMATION OF GLOBAL MOTION PARAMETERS

Rozenn Dahyot
Department of Statistics
Trinity College Dublin, Ireland
email: dahyot@mee.tcd.ie

Anil Kokaram
Electronic & Electrical Engineering Dpt
Trinity College Dublin, Ireland
email: akokaram@tcd.ie

Abstract

The estimation of Global or Camera motion from image sequences is important both for video retrieval and compression (MPEG4). This is frequently performed using robust M-estimators with the widely used Iterative Reweighted Least Squares algorithm. This article presents an investigation of the use of an alternative robust estimation algorithm and illustrates its improved computationnal efficiency. The paper also introduces two new confidence measures which can be used to validate camera motion measurements in the context of information retrieval.
Keywords: *Camera motion, M-estimators, Video analysis.*

1 Introduction

Content based information retrieval has been a highly active research area during the last decade [14]. The motivation has been that access via keywords allows only a primitive interaction with non-text media in general. To allow access on the basis of content (e.g. responses to questions like "find all aces in a game of tennis") is the key to efficient exploitation of these kinds of information. Typically the process begins by identifying and extracting feature primitives (colour, shape, motion,...) relevant to the content and these features might be related in some way to the human perception of the data. The features are then manipulated jointly in response to user queries or in order to identify events.

Motion is clearly an important feature in retrieval from video media, and Global Motion in particular captures the movement of the camera operator. This motion is well correlated to important events in video for example sport broadcasts [11] and can be used as a preliminary task before local motion analysis [9]. In this paper, global motion is considered to be that single motion representation which accounts for the largest moving area in the image.

A 6 parameter affine model is introduced in section 2 to represent this displacement. Those parameters are then estimated by minimising an energy function. It is local motion that complicates the estimation of global motion. In effect, that area of the image which undergoes local motion (or discontinuity) acts as an outlier in the global motion model. Standard estimation methods as presented in section 3, are sensitive to the presence of such outliers. By using instead, robust estimation processes [4] (M-estimators), it is possible to handle the presence of contaminated data in the observations i.e. the effect of global motion in this case. M-estimation has been applied to many problems in computer vision such as regularisation [6], motion optical flow estimation [1], object tracking [2], or object recognition and detection [8].

This paper considers two algorithms used for performing robust Global motion estimation with M-estimators. The first is the so-called *Iterative Reweighted Least Squares* (IRLS), and the second is the little known *Iterative Modified Residuals* (IMR) [10]. This article quantitatively analyses how well they perform with respect to the accuracy of the motion estimate itself and it compares their computation times. The paper shows that in fact, the two algorithms lead to the same performances for accuracy, the IMR is more computationally efficient.

Of special importance in any estimation problem on real data is to measure the confidence of the estimates. The paper also introduces two new confidence measures which can be use to validate camera motion measurements in the context of information retrieval.

2 Image sequence modelling

Motion estimation techniques presented in this article rely on the following image sequence model:

$$I_n(\mathbf{x}) = I_{n-1}(F(\mathbf{x}, \Theta)) + \epsilon(\mathbf{x}) \tag{1}$$

where $I_n(\mathbf{x})$ is the grey level of the pixel at the location given by position vector $\mathbf{x}$ in the frame n. The vector function $F(\mathbf{x}, \Theta)$ is the transformation of image coordinates induced by the motion between time $n-1$ and n. Θ is the vector formed by the motion parameters. In other words, it means that the current frame can be created by rearranging the position of the intensities from the previous frame.

Camera Motion Model. To represent motion such as zooming, rotation and translation between the current frame n and the previous frame $n-1$, a 6-parameter affine transformation $F(\mathbf{x}, \Theta) = \mathrm{A}\mathbf{x} + \mathbf{d}$ is used where A is a 2×2 matrix for affine transformation and $\mathbf{d}$ is the displacement vector, as below:

$$\begin{aligned} F(\mathbf{x}, \Theta) &= \mathrm{A}\mathbf{x} + \mathbf{d} \\ &= \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix} \\ &= \begin{pmatrix} a_1\, x + a_2\, y + d_x \\ a_3\, x + a_4\, y + d_y \end{pmatrix} = \mathrm{B}(\mathbf{x})\ \Theta \end{aligned} \tag{2}$$

where $\mathrm{B}(\mathbf{x}) = \begin{pmatrix} x & y & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 \end{pmatrix}$ and $\Theta = [a_1, a_2, d_x, a_3, a_4, d_y]^T$.

Toward a linear residual. The parameter Θ can be estimated by minimising some function of the residual $\epsilon(\mathbf{x}) = I_n(\mathbf{x}) - I_{n-1}(F(\mathbf{x}, \Theta))$. This residual $\epsilon(\mathbf{x})$ is however not linear in Θ. A Taylor series expansion around the motion parameters Θ is used to linearise the parameter estimation problem as follows:

$$I_n(\mathbf{x}) - I_{n-1}(\mathrm{B}(\mathbf{x})\ \Theta) = \nabla I_{n-1}(\mathrm{B}(\mathbf{x})\ \Theta) \cdot \mathrm{B}(\mathbf{x}) \cdot \delta\Theta + \varepsilon(\mathbf{x}) \tag{3}$$

$\epsilon(\mathbf{x})$ and the higher order terms of the expansion are lumped together in the new residual $\varepsilon(\mathbf{x})$ linear with respect to the update $\delta\Theta$. The ∇ operator is the usual multidimensional gradient operator. The estimation proceeds by recursive estimation of $\delta\Theta$ and updating of $\Theta \leftarrow \Theta + \delta\Theta$. This simple idea unifies all previous approaches to Global Motion estimations.

The equation (3) is expressed at each location $\mathbf{x}$. Considering all the pixels, it can be rewritten using vectors and matrices such as:

$$\mathbf{z} = \mathrm{G} \cdot \delta\Theta + \boldsymbol{\varepsilon} \tag{4}$$

with the vectors (limiting the notation to the first two locations $\mathbf{x_1} = (x_1, y_1)$ and $\mathbf{x_2} = (x_2, y_2)$):

$$\mathbf{z} = \begin{bmatrix} I_n(\mathbf{x_1}) - I_{n-1}(\mathrm{B}(\mathbf{x_1})\ \Theta) \\ I_n(\mathbf{x_2}) - I_{n-1}(\mathrm{B}(\mathbf{x_2})\ \Theta) \\ \vdots \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon(\mathbf{x_1}) \\ \varepsilon(\mathbf{x_2}) \\ \vdots \end{bmatrix}$$

and the following matrix:

$$G = \begin{bmatrix} x_1 \cdot I^x_{n-1}(\mathrm{B}(\mathbf{x_1})) & y_1 \cdot I^x_{n-1}(\mathrm{B}(\mathbf{x_1})) & I^x_{n-1}(\mathrm{B}(\mathbf{x_1})) & x_1 \cdot I^y_{n-1}(\mathrm{B}(\mathbf{x_1})) & y_1 \cdot I^y_{n-1}(\mathrm{B}(\mathbf{x_1})) & I^y_{n-1}(\mathrm{B}(\mathbf{x_1})) \\ x_2 \cdot I^x_{n-1}(\mathrm{B}(\mathbf{x_2})) & y_2 \cdot I^x_{n-1}(\mathrm{B}(\mathbf{x_2})) & I^x_{n-1}(\mathrm{B}(\mathbf{x_2})) & x_2 \cdot I^y_{n-1}(\mathrm{B}(\mathbf{x_2})) & y_2 \cdot I^y_{n-1}(\mathrm{B}(\mathbf{x_2})) & I^y_{n-1}(\mathrm{B}(\mathbf{x_2})) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

3 Maximum likelihood estimation

A maximum likelihood approach to estimation of Θ would choose an estimate $\hat{\Theta}$ which maximises the likelihood of $\epsilon(\mathbf{x})$ as all the sites in the image simultaneously $\mathcal{P}(\varepsilon)$. This is equivalent to minimising the log liklelihood $-\log \mathcal{P}(\varepsilon)$. In practice, by exploiting the linearisation of the log-likelihood (as in the previous section) around a current estimate of Θ called Θ_i; it is possible to generate an estimate by successively updating $\hat{\Theta}$ through $\hat{\Theta}^{i+1} = \Theta^i + \delta\hat{\Theta}$. Where $\delta\Theta^i$ is the update to be estimated.

Assuming the distribution of the residual $\boldsymbol{\varepsilon}$ is spherical (i.e. multidimensional gaussian with a covariance matrix proportional to the identity matrix), the algorithm can be written as :

$$\begin{array}{l} \text{Do} \\ \left| \begin{array}{l} \widehat{\delta\Theta} = \arg\min_{\delta\Theta}\{\mathcal{J}(\boldsymbol{\varepsilon}^{(i)}) = \sum_{\mathbf{x}}[\varepsilon^{(i)}(\mathbf{x})]^2\} \\ \widehat{\Theta}^{(i+1)} = \widehat{\Theta}^{(i)} + \widehat{\delta\Theta} \end{array} \right. \\ \text{Until convergence at final step } \boldsymbol{i}\ (\widehat{\Theta} = \widehat{\Theta}^{(\boldsymbol{i})}) \end{array} \tag{5}$$

At each step i, $\widehat{\delta\Theta}$ is estimated by Least Squares:

$$\widehat{\delta\Theta} = [\mathrm{G}^{(i)\ T}\mathrm{G}^{(i)}]^{-1}\mathrm{G}^{(i)\ T}\mathbf{z}^{(i)} \tag{6}$$

4 Robust M-estimation

Least Square estimation is sensitive to gross errors (or outliers) due to, for instance, local motion in the video different from the global motion or occlusion effects. M-estimators [10] are now widely used to perform robust estimation of global motion parameters [4, 15, 13]. The underlying assumption is that the probability density function of the residuals is no longer gaussian, and can be written as:

$$\mathcal{P}(\boldsymbol{\varepsilon}) \propto \exp\left[-\frac{1}{2}\sum_{\mathbf{x}} \rho\left(\frac{\varepsilon(\mathbf{x})}{\sigma_\rho}\right)\right] \tag{7}$$

Several functions ρ, convex or non-convex, have been proposed in the literature [4, 8, 13]. In our experiments (section 6), we have chosen the convex function $\rho(t) = 2\sqrt{1+t^2} - 2$, in order to avoid problems with local minima occurring with non-convex ones which could disturb the comparison of the algorithms. σ_ρ is the *scale parameter* that controls the limit where the influence of the outliers begins to decrease [10]. This parameter is fixed offline [4] but to simplify, it is set to 1 in the following equations.

The corresponding energy to minimize is non-quadratic and requires specific algorithms. Two have been proposed in the literature. The first is widely used and is called *The location step with modified weights* [10] in the robust statistic framework, or more commonly the *Iterative Reweighted Least Squares* [12], or as ARTUR in the Half Quadratic (HQ) formulation [5]. The IRLS algorithm is reviewed in paragraph 4.2.

The second algorithm, little known in computer vision literature, has first been called *The location step with modified residuals* in robust statistics [10] and as LEGEND in the HQ framework [5]. This algorithm is referred as Iterative Modified Residuals (IMR) and is explained in paragraph 4.3 for global motion parameter estimation. The section 4.1 briefly explains the origins of the two algorithms and they are compared in section 4.3.

4.1 Half-Quadratic Theory

Several explanations can account for both algorithms [10, 5, 8], but for simplicity we choose here the HQ framework. Maximising $\mathcal{P}(\boldsymbol{\varepsilon})$ in Θ is equivalent to minimize $\mathcal{J}(\boldsymbol{\varepsilon}) = \sum_{\mathbf{x}} \rho(\varepsilon(\mathbf{x}))$ iteratively in $\delta\Theta$. HQ theory defines an augmented energy $\mathcal{J}^*$ with the same global minimum:

$$\mathcal{J}(\boldsymbol{\varepsilon}) = \min_{\mathbf{b}} \left\{ \mathcal{J}^*(\boldsymbol{\varepsilon}, \mathbf{b}) = \sum_{\mathbf{x}} \rho^*(\varepsilon(\mathbf{x}), b(\mathbf{x})) \right\} \tag{8}$$

$\mathcal{J}^*$ is minimized iteratively in $\delta\Theta$ and $\mathbf{b} = \{b(\mathbf{x})\}_{\mathbf{x}}$:

Do
- Do
  - $\delta\Theta^{(j)} = \arg\min_{\delta\Theta} \left\{ \mathcal{J}^*(\boldsymbol{\varepsilon}^{(j)}, \mathbf{b}^{(j)}) \right\}$
  - $\mathbf{b}^{(j+1)} = \arg\min_{\mathbf{b}} \left\{ \mathcal{J}^*(\boldsymbol{\varepsilon}^{(j+1)}, \mathbf{b}^{(j)}) \right\}$
- Until convergence at final step $\boldsymbol{j}$ ($\widehat{\delta\Theta} = \delta\Theta^{(\boldsymbol{j})}$)
- $\widehat{\Theta}^{(i+1)} = \widehat{\Theta}^{(i)} + \widehat{\delta\Theta}$

Until convergence at final step $\boldsymbol{i}$ ($\widehat{\Theta} = \widehat{\Theta}^{(\boldsymbol{i})}$)

The different interactions of the auxiliary variable $\mathbf{b}$ with the residual $\boldsymbol{\varepsilon}$ defines the different robust algorithms of the M-estimation. $\mathbf{b}$ corresponds to weights on the residuals in the IRLS algorithm and is denoted $\mathbf{w}$ in section 4.2.

4.2 Iterative Reweighted Least Squares (IRLS)

The first proposed augmented energy can be written as:

$$\mathcal{J}^*\,(\boldsymbol{\varepsilon}, \mathbf{w}) = \sum_{\mathbf{x}} w(\mathbf{x})\ [\varepsilon(\mathbf{x})]^2 + \Psi(w(\mathbf{x})) \tag{9}$$

When the auxiliary variable $\mathbf{w} = \{w(\mathbf{x})\}_{\mathbf{x}}$ is fixed, the update is estimated by weighted Least Squares:

$$\delta\Theta^{(j)} = [\mathrm{G}^{(i)\ T}\mathrm{W}^{(j)}\mathrm{G}^{(i)}]^{-1}\ (\mathrm{G}^{(i)})^T\mathrm{W}^{(j)}\mathbf{z}^{(i)} \tag{10}$$

The diagonal matrix $\mathrm{W} = \mathrm{diag}(\mathbf{w})$ is then updated by $w^{(j+1)}(\mathbf{x}) = \frac{\rho'(\varepsilon(\mathbf{x}))}{2\cdot\varepsilon(\mathbf{x})}$. The weights act to reduce the effect of large residuals in the estimation process. A parametric expression of the function Ψ is proposed in [8]:

$$\left| \begin{array}{l} w = \frac{\rho'(\varepsilon)}{2\varepsilon} \\ \Psi = \rho(\varepsilon) - \frac{\rho'(\varepsilon)}{2}\varepsilon \end{array} \right.$$

Under some hypothesis on ρ [7, 3], this can be expressed as $\Psi(w) = \phi((\phi')^{-1}(w)) - w\ (\phi')^{-1}(w)$ with $\phi(x^2) = \rho(x)$.

4.3 Iterative Modified Residuals (IMR)

The second augmented energy can be written as:

$$\mathcal{J}^*\,(\boldsymbol{\varepsilon}, \mathbf{b}) = \sum_{\mathbf{x}} [\varepsilon(\mathbf{x}) - b(\mathbf{x})]^2 + \xi(b(\mathbf{x})) \tag{11}$$

When $\mathbf{b} = \{b(\mathbf{x})\}_{\mathbf{x}}$ is fixed, the update is computed by :

$$\delta\Theta^{(j)} = [\mathrm{G}^{(i)\ T}\ \mathrm{G}^{(i)}]^{-1}\ \mathrm{G}^{(i)\ T}\ (\mathbf{z}^{(i)} - \mathbf{b}^{(j)}) \tag{12}$$

The vector $\mathbf{b}$ is updated by $b^{(j+1)}(\mathbf{x}) = \varepsilon(\mathbf{x}) \left(1 - \frac{\rho'(\varepsilon(\mathbf{x}))}{2\ \varepsilon(\mathbf{x})}\right)$. Here the auxiliary variable acts again to reduce the effect of large residuals but here by subtraction of each outlier. ξ has been defined as [7]: $\xi(b) = \sup_u\{-(u-b)^2 + \rho(u)\}$. As this expression is not analytically exploitable, a parametric expression of ξ has also been proposed [8]:

$$\left| \begin{array}{l} b = \varepsilon \left(1 - \frac{\rho'(\varepsilon)}{2\varepsilon}\right) \\ \xi = \rho(\varepsilon) - \left(\frac{\rho'(\varepsilon)}{2}\right)^2 \end{array} \right. \tag{13}$$

Remarks It has been shown that the IRLS algorithm converges in less steps $\boldsymbol{j}$ to the estimate $\widehat{\delta\Theta}$ than the IMR [10, 8] (cf. figure 1). But in comparing equations (12) and (10), we see that the IMR algorithm involves less computation (product of matrixes $[\mathrm{G}^{(i)\ T}\ \mathrm{G}^{(i)}]$) at each j step than the IRLS ($[\mathrm{G}^{(i)\ T}\mathrm{W}^{(j)}\mathrm{G}^{(i)}]$).

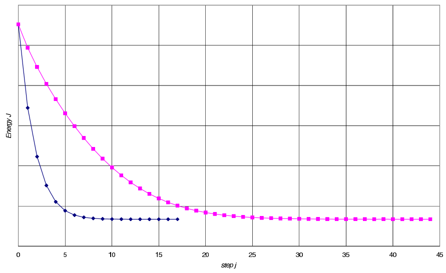

Figure 1: Energy $\mathcal{J}$ with respect to the step j. Convergence to the minimum faster for IRLS (diamond-blue) than for IMR (square-pink).

5 Confidence measures

The local motion behaviour or discontinuity contaminates the observations of the global motion. The level of contamination can be evaluated using the following measures.

Using the residuals. The probability density function of the residuals $\max_\Theta \mathcal{P}(\boldsymbol{\varepsilon})$ (or its corresponding energy defined as $\min_\Theta\{-\log \mathcal{P}(\boldsymbol{\varepsilon})\}$) is directly connected to estimator behaviour. If the estimate exactly accounts for the motion of each pixel within the image that is undergoing global motion, then the residual energy is zero (consequently, the likelihood probability is high). Conversely, high residual energy implies poor estimation and low likelihood probability.

Using the weights. The auxiliary variable used in the IRLS algorithm collects the weights defined on each residual. This weight is close to one when the residual is an inlier for the estimation of the global motion parameter, and close to zero otherwise. The image of the weights can be seen as a confidence map on the data. We propose a measure using those weights: $E[w^2] = \frac{\sum_{\mathbf{x}}[w(\mathbf{x})]^2}{\sum_{\mathbf{x}} 1}$. The measure is the Mean Square weight (MSW) across the entire image, and if most of the image can be accounted for by global motion, one would expect the IRLS MSW to be 1.0 and the IMR MSW to be 0.0. This measure is slightly different from that proposed by Bouthemy et al.[4] in that it does not require any prior thresholding.

6 Experimental results

Artificial sequences Three artificial video sequences (50 images of 360 × 288 pixels) were generated by applying a motion with known parameters on an original frame. The sequences show accelerating global motion in order to simulate rapid camera action.

Accuracy. As the table 1 shows, the accuracy of both algorithms on each parameter are the same. The error on the translation parameters is bigger than the one of matrix A, but is still very small (<< 1 pel).

	a_1	a_2	a_3	a_4	d_x	d_y
IRLS	4.10^{-4}	3.10^{-5}	3.10^{-5}	4.10^{-4}	1.10^{-2}	1.10^{-2}
IMR	4.10^{-4}	3.10^{-5}	4.10^{-4}	3.10^{-5}	1.10^{-2}	1.10^{-2}

Table 1: Average error on the parameters.

Computation time. The estimation has been computed using a three level pyramid as in [4] both to speed up the computation and to have accurate results. Figure 2 shows the advantage of this approach in terms of computation time: the notations IRLS and IMR (respectively PYR-IRLS and PYR-IMR) mean that the estimation has been performed without (resp. with) the pyramid of resolution. The curves show the computation time (in ms) for each frame of the sequence, and have been computed for the pan-only sequence which presents a decelerating translation for each frame n such that: $d_x(n) = 7 \exp\left[-\frac{n}{25}\right]$. The initial guess of the algorithms for each frame n is the identity transformation (i.e. in particular $d_x^{(i=0)} = 0$). At the beginning of the sequence, the initial guess is far from the solution, and therefore both algorithms (without using the pyramid decomposition) require more time to converge than at the end. This is not the case using the pyramid decomposition where a coarse to fine refinement is used. The

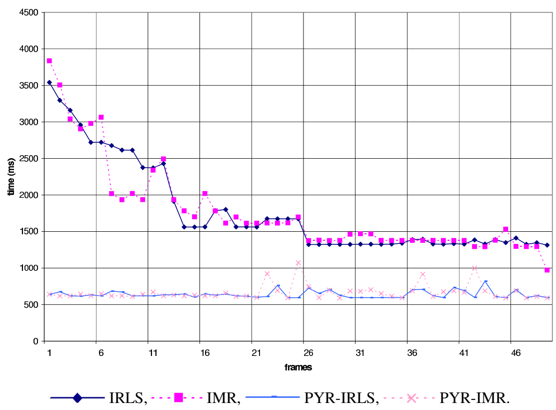

Figure 2: Computation time (pan-only).

table 2 presents the average time of computation on all our sequences using the pyramid decomposition. No obvious difference appears between the performance of IMR and IRLS.

	zoompan	pan only	zoom only
IRLS	641	639	605
IMR	633	666	607

Table 2: Average time (ms) on artificial sequences.

Real video sequences A video sequence of cricket has been processed using both algorithms (1500 images of 720×576 pixels).

Computation time. On the overall sequence, the IMR is reaching the estimate 10% faster than the IRLS (in comparing their average times $5612ms$ for IMR and $6476ms$ for IRLS over the sequence). As noticed in section 4.3, when the pixels in the image are numerous (that increases the size of the matrixes involved in the computation), the computation costs at each j step involving the products of matrixes in the IRLS algorithm can become time consuming.

Confidence measures. The figure 3 shows the confidence measures computed for the global motion parameter over the cricket sequence: there is an obvious correlation between the shot changes and weak values of the confidence measures (high values of the energy using the residuals, and low values of the confidence measure using weights). Abrupt transitions like cuts are well detected by both confidence measures, but not gradual transitions. As noticed in [4], confidence measures on the estimated parameters can also be used to detect shot changes in the video sequences. Figure 6 presents the weights estimated

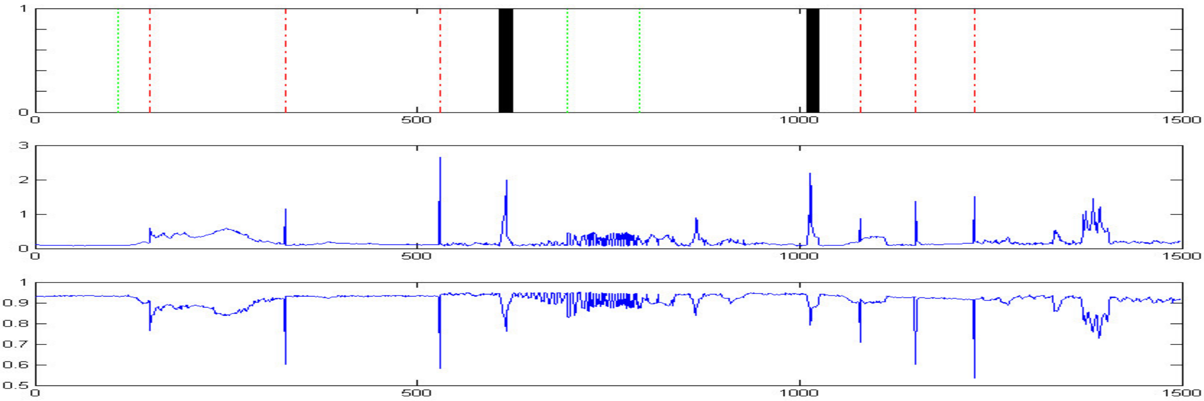

Figure 3: From top to bottom : ground truth of shot transitions (red dashdot lines for cuts, green dot lines for dissolves, black solid lines for wipes), confidence measures using residuals (middle) and weights (bottom).

for some images of the sequence. The weights are presented scaled by 255. High brightness represents pixels with high weights and dark pixels represent those with low weights. Low weights indicate pixels which are *not* part of the area undergoing global motion. This is in effect a representation of objects that are not following the camera motion such as the wipe in image $n = 616$ or people in frame 1400. Strong camera activity is detected in image $n = 251$ (travelling $d_x = 16$) through the presence of outliers on the right and left borders of the weight map. These are caused by off-scene locations that cannot be matched in the successive images because of the large inter-frame motion.

7 Conclusion

We have presented two algorithms performing the robust M-estimation. Depending on the size of the images, we have shown that the IMR algorithm can be faster, for the same accuracy, than the IRLS algorithm usually used to solve M-estimation for the global motion estimation problem. Global motion parameters are used for instance to index sport events [11] since the movement of the camera is highly correlated to the game. The two measures characterising the amount of contaminated data can help to detect shot changes [4]. Finally, weight maps provide a interesting start for local motion analysis and object segmentation in videos.

$n = 616$ $n = 551$ $n = 251$ $n = 1400$

Figure 4: Images from the cricket sequence with their weighting map $\{w(\mathbf{x})\}_{\mathbf{x}}$. Black pixels correspond to weights close to 0 (outliers), and white ones are weights close to 1 (inliers).

References

[1] M. J. Black and P. Anandan. A framework for the robust estimation of optical flow. In *Proceedings of the 4th International Conference on Computer Vision*, pages 231–236, Berlin, Allemagne, May 1993.

[2] M. J. Black and A. D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal on Computer Vision*, 26(1):63–84, January 1998.

[3] M. J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–92, July 1996.

[4] P. Bouthémy, M. Gelgon, and F. Ganansia. A unified approach to shot change detection and camera motion characterization. *IEEE Transactions on Circuits and Systems for Video Technology*, 9:1030–1044, 1999.

[5] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proceedings of the International Conference on Image Processing ICIP-94*, volume 2, pages 168–172, 1994.

[6] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2):298 – 311, February 1997.

[7] Pierre Charbonnier. *Reconstruction d'image: régularisation avec prise en compte des discontinuités*. PhD thesis, Université de Nice-Sophia Antipolis, France, 1994.

[8] R. Dahyot. *Analyse d'images séquentielles de scènes routières par modèles d'apparence pour la gestion du réseau routier*. Editions LCPC, ISBN: 2-7208-2028-1, France, September 2003.

[9] R. Fablet, P. Bouthemy, and P. Perez. Nonparametric motion characterization using causal probabilistic models for video indexing and retrieval. *IEEE Transactions on Image Processing*, 11(4):393–407, April 2002.

[10] P.J. Huber. *Robust Statistics*. John Wiley and Sons, 1981.

[11] A. Kokaram and P. Delacourt. A new global estimation algorithm and its application to retrieval in sport events. In *IEEE International Workshop on Multimedia Signal Processing, MMSP'01*, Cannes, France, October 2001.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, 1995.

[13] Wei Qi and Yuzhuo Zhong. New robust global motion estimation approach used in mpeg-4. *Journal of Tsinghua University Science and Technology*, 2001.

[14] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, December 2000.

[15] A. Smolic and J.-R. Ohm. Robust global motion estimation using a simplified m-estimator approach. In *IEEE International Conference on Image Processing*, Vancouver, Canada, September 2000.