

Significance tests and statistical inequalities for region matching

Guillaume Née, Stéphanie Jehan-Besson, Luc Brun, Marinette Revenu

▶ **To cite this version:**

Guillaume Née, Stéphanie Jehan-Besson, Luc Brun, Marinette Revenu. Significance tests and statistical inequalities for region matching. Structural, Syntactic, and Statistical Pattern Recognition; Joint IAPR International Workshops S SSPR 2008, Dec 2008, Orlando, Floride, United States. pp.350-360, 10.1007/978-3-540-89689-0_39 . hal-00812754

HAL Id: hal-00812754
https://hal.science/hal-00812754

Submitted on 12 Apr 2013

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.

Significance tests and statistical inequalities for region matching

G. Née[1,2]*, S. Jehan-Besson[1], L. Brun[1], and M. Revenu[1]

[1] GREYC Laboratory - Bd du Maréchal Juin. 6 - 14050 Caen - France
[2] General Electric Healthcare - Velizy - France
{gnee,jehan,brun,revenu}@greyc.ensicaen.fr

Abstract. Region matching - finding conjugate regions on a pair of images - plays a fundamental role in computer vision. Indeed, such methods have numerous applications such as indexation, motion estimation or tracking. In the vast literature on the subject, several dissimilarity measures have been proposed in order to determine the true match for each region. In this paper, under statistical hypothesis of similarity, we provide an improved decision rule for patch matching based on significance tests and the statistical inequality of McDiarmid. The proposed decision rule allows to validate or not the similarity hypothesis and so to automatically detect matching outliers. The approach is applied to motion estimation and object tracking on noisy video sequences. Note that the proposed framework is robust against noise, avoids the use of statistical tests and may be related to the *a contrario* approach.

1 Introduction

The notion of similarity between regions is a key feature for a wide range of applications in image analysis and pattern recognition such as clustering, indexation, image segmentation, shape matching or registration [1]. This notion is usually indirectly defined through a distance or a measure of dissimilarity. In many methods proposed for matching, the most similar pattern in a set of patterns is usually chosen. In this paper we are interested in the significance of the matching. One straightforward way to decide if two connected set of pixels (regions) are significantly similar consists in comparing the dissimilarity of both regions to a fixed threshold [18]. However, such a decision rule does not usually provide a clear interpretation of the threshold which is often difficult to set and which should be adapted to each data-set.

Deciding if two regions are similar may alternatively be stated as the decision between the two opposite hypothesis: "$\mathbf{H_0}$: both regions are similar" and "$\mathbf{H_1}$: both regions are dissimilar". Given a dissimilarity measure, one must measure the probability that an observation of this dissimilarity is greater than a given threshold under both hypothesis $\mathbf{H_0}$ and $\mathbf{H_1}$. The decision may then be stated using a likelihood ratio test. The main problem of such a decision scheme is that

* This work is funded by a grant co-financed by General Electric Healthcare and the Region Basse Normandie.

both hypothesis $\mathbf{H_0}$ and $\mathbf{H_1}$ are not symmetric. Indeed the set of regions similar to a given region is usually much simpler than the set of its dissimilar regions. Consequently, the design of a general model for $\mathbf{H_0}$ is usually tractable while the design of a model for $\mathbf{H_1}$ may be quite difficult or even impossible. This last point limits the use of likelihood ratio test.

The *a contrario* approach provides a way to solve this difficulty by setting the decision only on the null hypothesis $\mathbf{H_0}$. This approach has been used in several distinct fields such as detection of gelstat [5, 7, 6], motion detection [17], shape matching and recognition [13, 14], object matching [2], parameter estimation [8], blotch detection for digitized film restoration [16] and segmentation [15, 9]. The a contrario approach comes from the perception theory and particularly the grouping law of the Wertheimer's theory. This grouping law states that "objects having a quality in common get perceptually grouped". The Helmholtz principle [5] which states that "an event is meaningful if its number of occurrences is very small in a random model" is a quantitative version of the previous law. More formally, let us consider an event E which probability under $\mathbf{H_0}$ is bounded by a low threshold δ, the *a contrario* approach leads to reject the initial hypothesis if such an event occurs. Using such a decision scheme, δ may be interpreted as an upper bound of the probability of a false alarm (rejection of $\mathbf{H_0}$ while $\mathbf{H_0}$ is actually true). The probability δ may be fixed *a priori* in which case the test $P(E|H_0) < \delta$ is called a significance test [4]. Desolneux [6] proposed to set δ according to the expected number of false alarm. Such a method provides an elegant way to fix the threshold but reduces the adaptability of the method to user requirements.

Our approach is based on a conjoint use of the *a contrario* approach and a general formulation of the McDiarmid's theorem [11]. McDiarmid's theorem allows to design statistical tests for a wide variety of problems and provides a natural robustness to noise and to outliers. Within our framework this theorem is used to bound the probability that the dissimilarity between two regions is greater than a given threshold. Such a bound is incorporated into the *a contrario* framework in order to provide an automatic threshold on the dissimilarity between two regions. This general framework is illustrated by a dissimilarity measure between patches which corresponds to the Sum of Squared Difference (SSD) involved in Block Matching algorithms. This dissimilarity measure is used in conjunction with the McDiarmid Theorem within the *a contrario* approach in order to design a robust Variable Size Block Matching (VSBM) algorithm that can be used for motion estimation and tracking in noisy video sequences.

We first present our *a contrario* approach for region matching in section 2. We then illustrate this method using a first dissimilarity measure between patches in section 3. Several experiments presented in section 4 demonstrate the effectiveness of our approach for variable size block matching (VSBM) and tracking in noisy video sequences.

2 Problem statement and statistical tools

We present in this section our approach for region matching and the statistical tools which constitute the basis of our approach.

2.1 *A contrario* approach for region matching

Due to the random part in image acquisition systems, an image I is classically considered as an observation of a perfect statistical image I^*. Using such an image model, an ideal region is defined as a vector $\mathbf{X} = (X_1, \ldots, X_n)$ of n random variables (r.v.) representing the pixel intensities. A "real" region is then considered as an observation of this random vector which takes its values in $\prod_{k=1}^{n} A_k$. In natural images, the set of admissible values A_k usually corresponds to $[0; M]$ where $M = 255$. However, in medical images (e.g. : MRI, Echography), the set A_k may be larger.

Given two statistical regions $\mathbf{X}_1$ and $\mathbf{X}_2$ and a dissimilarity criterion $d(.,.)$, let us consider two observations R_1 and R_2 of respectively $\mathbf{X}_1$ and $\mathbf{X}_2$ and the event E: *the observed value* $d(R_1, R_2)$ *of the statistic* $d(\mathbf{X}_1, \mathbf{X}_2)$ *is greater than a threshold* T. As mentioned in section 1, the *a contrario* approach is based on the estimation of the probability of this event under the similarity hypothesis $\mathbf{H_0}$. Let us consider an upper bound δ of this probability:

$$\mathbf{P}\left\{d(\mathbf{X}_1, \mathbf{X}_2) \geq T | \mathbf{H_0}\right\} \leq \delta \tag{1}$$

We can remark that the probability δ and the threshold T are dependent. Indeed if the threshold T is set to a high value, this corresponds to a non probable event under $\mathbf{H_0}$ and δ should then be small. On the contrary if the threshold T is set to a small value, this corresponds to a probable event under $\mathbf{H_0}$ and so δ must be large. More generally, one may usually assume that the threshold T is a decreasing function of δ which may be denoted as $T(\delta)$.

Using the *a contrario* approach, if we take δ as a low probability value, the event E is considered as not probable under the similarity hypothesis $\mathbf{H_0}$ and this hypothesis is then rejected. Our decision rule for region matching is then defined as:

$$\text{if} \quad d(R_1, R_2) \geq T(\delta) \quad \text{then} \quad \mathbf{H_0} \,\text{is rejected} \tag{2}$$

2.2 Computation of the threshold

The main difficulty of the above approach lies in the computation of the threshold $T(\delta)$. In this work, we propose to use the extension of the McDiarmid theorem [11] which allows to bound the probability of a large class of events. Let us remind this theorem:

Theorem 1. *Let* $\mathbf{Y} = (Y_1, \ldots, Y_n)$ *be a family of random variables with* Y_k *taking values in a set* A_k*, and let* f *be a bounded real-valued function defined on* $\Omega = \prod_{k=1}^{n} A_k$*. If* μ *denotes the expectation of* $f(\mathbf{Y})$ *we have for any* $\alpha \geq \mu$*:*

$$\mathbf{P}\left\{f(\mathbf{Y}) \geq \alpha\right\} \leq \exp\left(\frac{-2(\alpha - \mu)^2}{r^2}\right) + \mathbf{P}\left\{\mathbf{Y} \in C\right\} \tag{3}$$

Where C *is a subset of* Ω *and* r^2 *is the maximal sum of squared range (section 6.1) defined on* $\overline{C} = \Omega - C$*.*

The set C in the above theorem corresponds to a set of outliers for $\mathbf{Y}$. Within our framework, we define $f(Y)$ as our dissimilarity measure $d(\mathbf{X}_1, \mathbf{X}_2)$ and Y as an appropriate combination of the two vectors $\mathbf{X}_1$ and $\mathbf{X}_2$. An example will be given in the next section. Theorem 1 then provides a bound for $P(d(\mathbf{X}_1, \mathbf{X}_2) \geq \alpha)$. The hypothesis $\mathbf{H_0}$ is introduced through the parameters r and μ. Some assumptions on the noise model can be used to estimate those parameters (see section 3.1).

A simplified version of theorem 1 has already been used [15, 9] within the segmentation framework. In this simplified version $\overline{C} = \Omega$ and $\mathbf{P}\{\mathbf{Y} \in C\} = 0$. To our knowledge, the general theorem 1 has never been used for region matching whereas it provides a natural tool for taking into account outliers which may occur in a statistical image. Indeed, the set $\overline{C}$ can be used to define the set of admissible values for a given type of images or regions.

Let us denote by $\Delta(\alpha)$, the probabilistic bound given by the McDiarmid's theorem:

$$\mathbf{P}\{f(\mathbf{Y}) \geq \alpha\} \leq \Delta(\alpha) \tag{4}$$

The threshold $T(\delta)$ introduced in (1) can then be computed by setting $\delta = \Delta(\alpha)$ and so $\alpha = \Delta^{-1}(\delta)$.

3 A first dissimilarity measure for patch matching

We consider two statistical regions defined in two images. In order to match heterogeneous regions we restrict our definition of region to rectangular patches. Such a restriction allows us to assume a one to one correspondence between the pixels of two matched regions.

Let us denote the considered regions by $\mathbf{X}_1 = \{X_i\}_{1 \leq i \leq n}$ and $\mathbf{X}_2 = \{X_i\}_{n+1 \leq i \leq 2n}$. The correspondence between the pixels of both regions is implicitly encoded by the index, since we suppose that a pixel X_i, $i \in \{1, \ldots, n\}$ of $\mathbf{X}_1$ corresponds to the pixel X_{i+n} of $\mathbf{X}_2$. We measure the similarity between corresponding pixels by the following dissimilarity measure which corresponds to the classical Sum of Square Differences (SSD):

$$d(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{n} \sum_{i=1}^{n} (X_i - X_{n+i})^2 \tag{5}$$

For the sake of simplicity, we introduce the random vector $\mathbf{Y} = (Y_1, \ldots, Y_n)$ where $Y_i = (X_i - X_{n+i})^2$, $i \in \{1, \ldots, n\}$. If $[0; M]$ is the interval of variation of the r.v. (X_i), then Y_i varies within $[0; M^2]$.

Our goal is to compute a decision rule that indicates if two observations R_1 and R_2 of $\mathbf{X}_1$ and $\mathbf{X}_2$ are similar or not. According to our approach (section 2), we have to upper bound the probability that the function $d(\mathbf{X}_1, \mathbf{X}_2)$ is greater than a given threshold α using the McDiarmid's theorem (theorem 1). Such an upper bound is provided by the following proposition:

Proposition 1. *Using the previously defined notations, we have for any couple* $(\mathbf{X}_1, \mathbf{X}_2)$ *of statistical regions and any threshold* $\alpha > 0$*:*

$$\mathbf{P}\{d(\mathbf{X}_1, \mathbf{X}_2) \geq \alpha\} \leq \exp\left(-\frac{2n(\alpha - \mu)^2}{N^4}\right) + K \tag{6}$$

with $K = \mathbf{P}\{\mathbf{Y} \in C\}$ *where* $C = [N^2; M^2]^n \subset \Omega$*, is defined by the parameter* N *and* $\mu = \mathbf{E}(f(\mathbf{Y}))$ *with* $f(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i = d(\mathbf{X}_1, \mathbf{X}_2)$*.*

Proof. see Appendix, section 6.2.

The exact values of K and μ should be determined according to a specific model of similarity between regions (section 3.1). The parameter N determines the size of our set of outliers. Intuitively, if each pixel X_i, $i \in \{1, \ldots, n\}$ of $\mathbf{X}_1$ corresponds to the pixel X_{i+n} of $\mathbf{X}_2$, observations of the random variables Y_i should be around 0 with a large probability. We thus define *a priori* the set $[0; N^2]$ on which we expect a large number of observations and consider the set $C =]N^2; M^2]$ as our set of outliers. The parameter N should be set for each application and each dissimilarity measure. However, experiments, not reported here shown us that the optimal value of N remains stable on a wide class of images. The value of N as been fixed to 65 in all experiments reported below.

As described in section 2, given the parameter N, the probability K remains fixed and the right side of equation 6 defines a function Δ of α such that:

$$\mathbf{P}\{d(\mathbf{X}_1, \mathbf{X}_2) \geq \alpha\} \leq \Delta(\alpha).$$

The function Δ being bijective this last equation is equivalent to:

$$\mathbf{P}\left\{d(\mathbf{X}_1, \mathbf{X}_2) \geq \Delta^{-1}(\delta)\right\} \leq \delta.$$

For a threshold δ.

According to our decision rule (section 2.1), we thus reject the hypothesis $\mathbf{H_0}$ that both regions are similar if $d(\mathbf{X}_1, \mathbf{X}_2)$ is greater than:

$$T(\delta) = \Delta^{-1}(\delta) = \mu + N^2\sqrt{\frac{1}{2n}\ln\left(\frac{1}{\delta - K}\right)} \tag{7}$$

3.1 Statistical prior for parameter estimation

In order to explicitly compute the parameters μ and K given in equation 7, we need to formalize our similarity hypothesis $\mathbf{H_0}$. This task is achieved by the notion of σ-similar regions which relies on a Gaussian noise model:

Definition 1. *(σ-similarity)*
Two statistical regions $\mathbf{X}_1 = (X_i)_{i\in\{1,\ldots,n\}}$ *and* $\mathbf{X}_2 = (X_i)_{i\in\{n+1,\ldots,2n\}}$ *are called* σ*-similar if the r.v.* $(X_i - X_{i+n})$ *are independent and identically distributed (iid) and follow a normal distribution* $\mathcal{N}(0, \sigma)$ *for any* $i \in \{1, \ldots, n\}$*.*

Let us consider two σ-similar regions $\mathbf{X}_1$ and $\mathbf{X}_2$. As all r.v. $X_i - X_{i+n}$ are iid, the r.v. Y_i are iid and follow $\mathbf{P}\{Y_i = y\} = \frac{1}{\sigma\sqrt{2y\pi}}\exp\left(-\frac{y}{2\sigma^2}\right)$. We have thus:

$$\mu = \mathbf{E}(f(\mathbf{Y})) = \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(Y_i) = \sigma^2$$

$$K = \mathbf{P}\left\{\mathbf{Y} \in C = [N^2; M^2]^n\right\} = \prod_{i=1}^{n}\mathbf{P}\left\{Y_i \in [N^2; M^2]\right\}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\int_{N^2}^{M^2}\frac{1}{\sqrt{y}}\exp\left(-\frac{y}{2\sigma^2}\right)dy\right)^n = \left(\operatorname{erf}\left(\frac{M}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{N}{\sigma\sqrt{2}}\right)\right)^n$$

Note that our threshold (equation (7)) takes into account the amount of noise through the two terms $\mu = \sigma^2$ and K.

3.2 Patch matching

In order to illustrate our approach, let us give some basic results for patch matching between two frames of a video sequence.

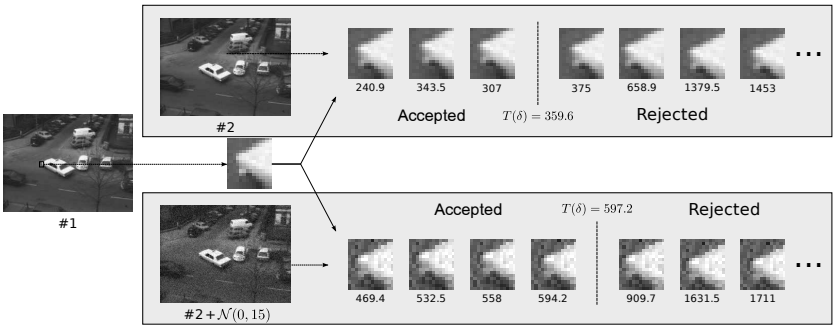


Fig. 1. Examples of acceptation/rejection of 13x15-patch matching. The candidate patches from frame #2 are selected around the position of the reference patch (from frame #1).

Given a reference patch from frame #1, we present the set of accepted/rejected patches of frame #2 in Fig. 1. The frame #2 is considered with or without the addition of a Gaussian noise $\mathcal{N}(0, 15)$. The candidate patches are presented with their associated value of dissimilarity. To compute the threshold $T(\delta)$, the ratio δ of authorised false alarms is set to 0.1 and N is set to 65 for both experiments (with or without noise on the frame #2). The value of σ is estimated by extracting homogeneous 20x20-patches from the two images. We find $\sigma = 5.91$ between the non-noisy frames #1 and #2, and $\sigma = 16.5$ between the frame #1 and the noisy frame #2. The rejected patches are the ones which present significant translations of the car hood inside the patch. We can notice that the computed threshold adapts itself automatically to the additional noise in frame #2. Indeed, using the same threshold for non-noisy sequences would lead to the rejection of all patches in noisy sequences. With this adaptive threshold, the discrimination between rejected and accepted patches is still valid.

4 Application to motion estimation and tracking

The framework developed above can be efficiently applied for motion estimation and tracking (especially in the case of noisy images). This leads us to the design of a robust variable size block matching (VSBM) method that is then used for rigid object tracking. For both methods, we present some experimental results on real video sequences and we focus our attention on two main points: robustness to noise and automatic detection of false matching.

4.1 Motion estimation using variable size block matching

Block matching (BM) algorithms were originally described by Jain and Jain [10] and are now widely adopted for motion estimation especially in video compression (MPEG-2,MPEG-4). Each frame is divided into a fixed number of usually square blocks. For each block in the frame, we search for the "best" matching block in the next frame. The dissimilarity measure is usually taken as the SSD measure defined in equation 5. Since a fixed size for the blocks is not adapted to the granularity of both motion and objects, VSBM algorithms have been introduced and are now widely used in new compression standards (H264, MPEG-4 AVC) [18]. If the best matching error for a block is above some threshold, the block is divided into four smaller blocks, until the maximum number of blocks or locally minimum errors are obtained, see [3] for example.

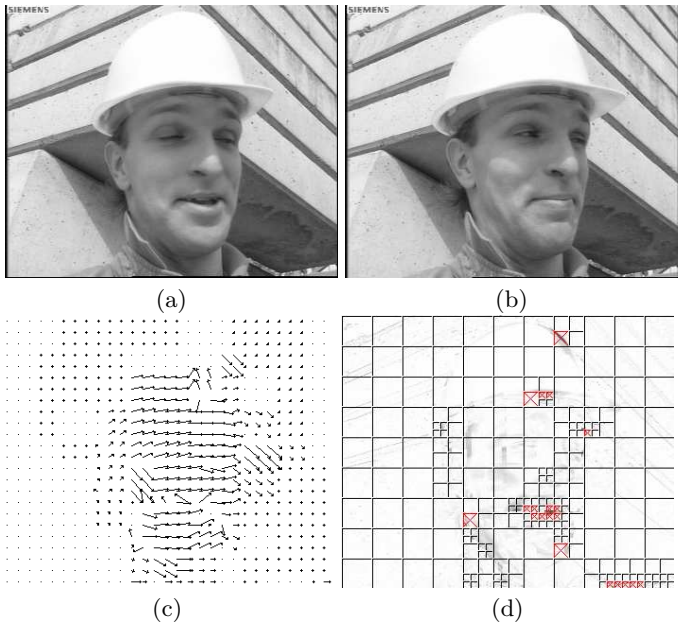


Fig. 2. "Foreman" sequence (a) and (b), the computed motion estimation (c) by our VSBM. The variable size block division, the residual between the reconstruction of (b) from (a) and rejected blocks represented by a cross are superposed on (d).

Using *a contrario* approaches, we have designed a statistical decision rule for the dissimilarity measure used in BM. Our decision rule allows us to detect false matching and is robust to noise. It is thus adapted to the design of VSBM algorithms. Fig. 2 presents the results of our VSBM between frames (a) and (b). Motion estimation errors are visualized through the inverse difference (d) between (b) and its reconstruction. Such a reconstruction is computed using (a) and the estimated motion field (c) obtained by our VSBM algorithm. These results have been obtained by choosing 32x32-blocks which can be recursively cut in half size blocks while the SSD is above the threshold $T(\delta)$ and until 8x8-blocks are reached. The threshold T given by (7) is computed by fixing N to 65

and δ to 0.1. We can remark that the algorithm gives smaller blocks when the motion is more complex (the mouth or the helmet). Rejected blocks (represented by a cross on (d)) are also (qualitatively) well detected. Such blocks correspond to strong and complex motions, such as the mouth which is opened on frame (a) and closed on frame (b).

4.2 Application to tracking in noisy sequences

The validity of our algorithm can be further attested for rigid object tracking in noisy video sequences. In this application, we assume as in [12] that the object of interest undergoes an affine deformation from one frame to another which is estimated by minimizing the following convex criterion according to the six parameters of the affine transformation $\mathbf{A}$:

$$E(\mathbf{A}) = \sum_{\mathbf{p} \in \Omega_I / \Theta} ||\mathbf{d}(\mathbf{p}) - \mathbf{A}(\mathbf{p})||^2$$

where $\mathbf{p} = [x_1, x_2]^T$ is the pixel, $\mathbf{d}(\mathbf{p}) = [u(\mathbf{p}), v(\mathbf{p})]^T$ the motion vector estimated as explained in the previous section, $\mathbf{A}$ the affine transformation applied on $\mathbf{p}$, and Ω_I the region domain. The set Θ represents the set of pixels within the rejected blocks (section 4.1). Such blocks are represented by a cross in Fig. 2. The estimated transformation $\mathbf{A}$ is then used to deform an initial segmentation.

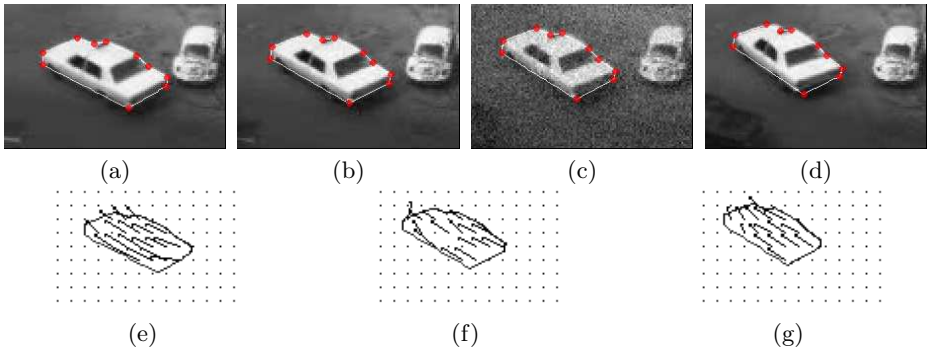


Fig. 3. First line - tracking of the white car in a sequence which presents a noisy frame (c). Second line - the motion between frames is estimated by our VSBM algorithm (section 4.1).

Fig. 3 illustrates the effectiveness of our method on the video sequence "Taxi" where an additional Gaussian noise $\mathcal{N}(0, 15)$ has been added on frame (c). The white car has been manually segmented in the first frame (a). This segmentation is then deformed according to the affine transformation computed using the motion field estimated by our VSBM (section 4.1). The estimated motion is presented on the second line of Fig. 3 and the tracking results on the first line. The set of parameters used here for the VSBM is the same as the one used in the previous experiment (section 4.1) and is kept all along the sequence. This result shows that we can track the car with the same set of parameters all along the sequence even in its noisy part.

5 Conclusion

In this paper, we propose to combine *a contrario approaches* and the recent statistical inequality of McDiarmid to design significance tests for region matching. This approach is illustrated by designing an improved decision rule for patch matching that is robust to noise. This decision rule has been applied to motion estimation with a VSBM algorithm and tracking. Our experimental results show the effectiveness of our approach on noisy video sequences. Our on going research is directed towards the design of other decision rules for region merging algorithms so as to extend the work proposed in [9]. Moreover, we plan to extend our method to other noise models such as Rayleigh or Poisson which are encountered in several application domains (e.g. medical image).

6 Appendix

6.1 Statistical notations and definitions

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a family of r.v. with Y_k taking values in a set A_k, and let f be a real-valued function defined on $\prod A_k$. Let $y_i \in A_i$ for each $i = 1, \ldots, k$ and let B_k denotes the set of events $\{Y_i = y_i\}_{i=1,\ldots k}$. For $y_k \in A_k$, let:

$$g(y_k) = \mathbf{E}\left(f(\mathbf{Y})|B_{k-1}, Y_k = y_k\right) - \mathbf{E}\left(f(\mathbf{Y})|B_{k-1}\right) \tag{8}$$

be the function that measures how much the expected value of $f(\mathbf{Y})$ changes if it is revealed that the r.v. Y_k takes the value y_k. By these notations, the range of $g(Y_k|B_{k-1})$ is defined as follows:

$$ran(y_1, \ldots, y_{k-1}) = \sup\left\{|g(y) - g(x)| : x, y \in A_k\right\} \tag{9}$$

For $\mathbf{y} \in \prod A_k$, the sum of squared range and its maximum are defined as:

$$R^2(y) = \sum_{k=1}^{n} (ran(y_1, \ldots, y_{k-1}))^2, \quad \hat{r}^2 = \sup_{y \in \prod A_k} \left(R^2(y)\right) \tag{10}$$

6.2 Proof of Proposition 1

The r.v. $\{Y_k\}_{k=1\ldots n}$ are independent and so the function g defined in equation (8) that measures how much the expected value of SSD changes can be written as:

$$\begin{aligned} g(y) &= \mathbf{E}\left(f(y_1, \ldots, y_{k-1}, y_k, Y_{k+1}, \ldots, Y_n)\right) \\ &\quad - \mathbf{E}\left(f(y_1, \ldots, y_{k-1}, Y_k, Y_{k+1}, \ldots, Y_n)\right) \\ g(y) &= \frac{1}{n}\left(y_k - \mathbf{E}\left(Y_k\right)\right) \end{aligned} \tag{11}$$

Now, let us consider $C = \left\{\mathbf{Y} = \mathbf{y} : \mathbf{y} \in \left[N^2; M^2\right]^n\right\}$ the subset of all the outliers such that the vector of r.v. $\mathbf{Y} = \mathbf{y}, \mathbf{y} \in C$ is not acceptable, then the range defined by equation (9) becomes:

$$ran(y_1, \ldots, y_{k-1}) = \sup\left\{\frac{1}{n}\left|y_k - y'_k - \mathbf{E}\left(Y_k - Y'_k\right)\right| : y_k, y'_k \in \left[0; N^2\right]\right\} = \frac{N^2}{n} \tag{12}$$

where y'_k is another observation of the r.v. Y_k. Consequently:

$$\mathbf{E}\left(Y_k - Y'_k\right) = \mathbf{E}\left(Y_k\right) - \mathbf{E}\left(Y'_k\right) = 0$$

So the *maximum sum of squared ranges* defined in (10) is equal to: $\hat{r}^2 = \frac{N^4}{n}$.
By applying equation (3) and from theorem 1, we obtain the proposition 1:

$$\mathbf{P}\left\{f(\mathbf{Y}) \geq \alpha\right\} \leq \exp\left(-\frac{2n(\alpha - \mu)^2}{N^4}\right) + K \tag{13}$$

where $K = \mathbf{P}\left\{\mathbf{Y} = \mathbf{y} : \mathbf{y} \in C\right\}$ and $\mu = \mathbf{E}\left(f(\mathbf{Y})\right)$.

References

1. P. Aschwanden and W. Guggenbül. Experimental results from a comparative study on correlation type registration algorithms. In W. Förstner and S. Ruwiedel, editors, *Robust computer vision: Quality of Vision Algorithms*, pages 268–282. Wichmann, Karlsruhe, Allemagne, mars 1992.
2. N. Burrus, T.M. Bernard, and Jolion J.M. Bottom-up and top-down object matching using asynchronous agents and *a contrario* principles. In Antonios Gasteratos, Markus Vincze, and John K. Tsotsos, editors, *Computer Vision Systems*, volume 5008 of *LNCS*, pages 343–352. Springer, 2008.
3. M.H. Chan, Y.B. Yu, and A.G. Constantinides. Variable size block matching motion compensation with applicationsto video coding. *IEEE Communications, Speech and Vision*, 137(4):205–212, 1990.
4. D. Coupier, A. Desolneux, and B. Ycart. Image denoising by statistical area thresholding. *Journal of Mathematical Imaging and Vision*, 22(2-3):183–197, 2005.
5. A. Desolneux, L. Moisan, and J.-M. Morel. Meaningful alignments. *International Journal of Computer Vision*, 40(1):7–23, 2000.
6. A. Desolneux, L. Moisan, and J.-M. Morel. Computational Gestalts and perception thresholds. *Journal of Physiology*, 97(2-3):311–324, 2003.
7. A. Desolneux, L. Moisan, and J.-M. Morel. Maximal meaningful events and applications to image analysis. *Annals of Statistics*, 31(6):1822–1851, 2003.
8. R.L. Felip, X. Binefa, and J. Diaz Caro. A new parameter estimator based on the helmholtz principle. In *International Conference on Image Processing*, pages 1306–1309, 2005.
9. M. El Hassani, S. Jehan-Besson, L. Brun, and al. Time-consistent video segmentation algorithm designed for real-time implementation. *VLSI Design*, 2008.
10. J.R. Jain and A.K. Jain. Displacement measurement and its application in interframe image coding. *IEEE Transactions on Communications*, pages 1799–1808, 1981.
11. C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics, Springer*. 1998.
12. F. Meyer and P. Bouthemy. Region-based tracking using affine motion models in long image sequences. *CVGIP : Image Understanding*, 60(2):119–140, 1994.
13. P. Musé, F. Sur, F. Cao, and Y. Gousseau. Unsupervised thresholds for shape matching. In *International Conference on Image Processing*, pages 647–650, 2003.
14. P. Musé, F. Sur, F. Cao, Y. Gousseau, and J.-M. Morel. An a contrario decision method for shape element recognition. *International Journal on Computer Vision*, 69(3):295–315, 2006.
15. R. Nock and F. Nielsen. Statistical region merging. *IEEE Pattern Analysis and Machine Intelligence*, 26(11):1452–1458, 2004.
16. S. Tilie, L. Laborelli, and I. Bloch. Blotch Detection for Digital Archives Restoration based on the Fusion of Spatial and Temporal Detectors. In *Fusion*, Florence, Italy, 2006.
17. T. Veit, F. Cao, and P. Bouthemy. An a contrario decision framework for region-based motion detection. *International Journal on Computer Vision*, 68(2):163–178, 2006.
18. T. Wiegand, G.J. Sullivan, G Bjntegaard, and A. Luthra. Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7):560 – 576, 2003.