

# A Sharp Concentration-Based Adaptive Segmentation Algorithm

Christophe Fiorio, André Mas

► **To cite this version:**

Christophe Fiorio, André Mas. A Sharp Concentration-Based Adaptive Segmentation Algorithm. G. Bebis. ISVC'2010: 6th International Symposium on Visual Computing, Nov 2010, Las Vegas, United States. Springer-Verlag Berlin Heidelberg, Advances in Visual Computing (6454), pp.85-96, 2010, LNCS. <<http://www.isvc.net/>>. <10.1007/978-3-642-17274-8\_9>. <lirmm-00532783>

**HAL Id: lirmm-00532783**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00532783>**

Submitted on 4 Nov 2010

**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.

# A sharp concentration-based adaptive segmentation algorithm

Christophe Fiorio, André Mas

No Institute Given

**Abstract.** We propose an adaptive procedure for segmenting images by merging of homogeneous regions. The algorithm is based on sharp concentration inequalities and is tailored to avoid over- and under-merging by controlling simultaneously the type I and II errors in the associated statistical testing problem.

## 1 Introduction

In this paper we focus on a strategy which belongs to region merging segmentation [1,2]. It consists in starting with pixel or very small groups of connected pixels as primary regions and then grouping them two by two when adjacent regions are considered to be the same. This decision is generally made through a statistical test to decide the merging of regions [3]. The segmentation process uses a merging predicate based on this test to decide, essentially locally, to merge or not regions. Difficulties of these approaches come from essentially two points. First, this locality in decisions has to preserve global properties such as those responsible for the perceptual units of the image. Second, since based on local decision, algorithms are generally greedy and so order-dependent. But conversely, these approaches can yield very efficient algorithms [4,5].

Our goal in this paper is to propose a local criterion based on a statistical test which ensures preserving global properties and stay efficient. The problem of order dependence will be mentioned in section 8 at the end of this document. To this aim we use concentration inequalities as first proposed by C. Fiorio and R. Nock [6,7] and then continued by R. Nock and F. Nielsen [8] to propose a statistical region merging algorithm. We propose to use another concentration inequality in order to try to improve quality of the result but also to improve control on the segmentation by defining more significant parameters and proposing to point out an indifference zone where we know that decision is not sure.

The paper is organized as follow. First in section 2 we present the basic theory we use to define our merging criterion. Then our framework is presented at section 3. Section 4 shows how to approximate our formula to be able to use them in practice. At last, section 5 presents our main result for region segmentation. Then we compare our results with previous works at section 6 and give some experimental results at section 7. At last we conclude this article at section 8.

## 2 Basics

Image segmentation by merging of regions or by any technique involving a decision rule (based for instance on a threshold) may be viewed as a traditional two (or more) samples testing problem. Usually some numerical features of two regions must be compared to decide whether they look alike or not. Suppose you are given two regions, say  $R$  and  $R'$ , made of pixels. Roughly speaking the general test could be written :  $\begin{cases} H_0 : \text{The regions } R \text{ and } R' \text{ must be merged} \\ H_1 : \text{The regions must not be merged} \end{cases}$ .

We remind the reader that two types of error may appear when implementing a test: rejecting  $H_0$  wrongly called type I error (we shall denote  $\alpha$  the probability of the type I error), and rejecting  $H_1$  wrongly called type II error (we shall denote  $\beta$  the probability of the type II error). It is a well known fact that  $\alpha$  and  $\beta$  cannot be made arbitrarily small together [9]

Usually authors propose merging predicates based on the control of the type I error (see e.g. [10,11]) without controlling the probability of the type II error. The main drawback of this procedure is clearly overmerging (one may not accept  $H_1$  often enough).

In this article we propose an adaptive procedure to control both error types. This procedure is adaptive in the sense that for each test the threshold depends on second order characteristics of the regions to be tested (the variance associated to the greylevels in fact). Besides the decision rule even depends on the size of the regions. In the following  $\mathbb{E}$ ,  $\mathbb{V}$  and  $\mathbb{P}$  denote expectation variance and probability respectively. Now we focus on the mathematical setting of our study. We consider that the region  $R$  is made of pixels viewed as independent discrete random variables denoted  $(X_i)_{i \in |R|}$  with values in  $\{0, \dots, 256\}$ . We assume that for all  $X_i \in R$ ,  $\mathbb{E}X_i = m$  and  $\mathbb{V}X_i = \sigma^2$ .

Inspired by Nock and Fiorio's theory [6,12,8] of underlying perfect images we claim that two regions should be merged whenever the common expectation of pixels from both images are the same. In fact,  $H_0 : m = m'$  and  $H_1 : m \neq m'$ . Clearly  $m'$  is the expectation of the greylevel for region  $R'$ . This classical problem could be tackled through ANOVA techniques. But the usual assumption (gaussian distributions and common variance amongst the dataset) are unrealistic. Fortunately our method is based on concentration inequalities and makes independence on the pixels the only serious assumption in the model.

Now focusing on our segmentation problem and denoting  $X_i$  (resp.  $X'_j$ ) the value of the greylevel of pixel  $i$  (resp.  $j$ ) in region  $R$  (resp  $R'$ ), we set the empirical greylevel mean in  $R$  by  $S = \frac{1}{|R|} \sum_{i \in R} X_i$  and  $S' = \frac{1}{|R'|} \sum_{i \in R'} X'_i$  in region  $R'$ .

Let us denote by  $\sigma^2$  the common variance of the distributions of pixels within region  $R$  and  $\sigma'^2$  its counterpart for  $R'$ . If we deal with the type I error, the decision rule will be based on the difference  $|S - S'|$  and on a threshold  $t_\alpha$ . The latter being defined by

$$\mathbb{P}_{H_0} (|S - S'| > t_\alpha) \leq \alpha \tag{1}$$

The type II error will be treated below in the paper similarly to the type I error. The next section is devoted to introducing the probabilistic tools needed to handle inequalities such as (1).

### 3 Random framework and tools

#### 3.1 Concentration inequalities based on variance

Concentration inequality roughly claim that the probability for functional of independent random variables to deviate from their means rapidly decrease : these functionals are "concentrated" around a non random number (the expectation) with a high probability.

For instance the well-known Mc Diarmid's inequality [13] used by C. Fiorio and R. Nock to derive concentration bounds and thresholds for the segmentation algorithm in [6,12,8] is extremely general and powerful in many cases. Yet insofar as the procedure relies on a comparison of means (see 1), it could be fruitful to use concentration inequality more specifically designed for sums of independent random variables instead of general functionals. Amongst these are the famous Hoeffding's [14], Bennett's [15] and Bernstein's [16] inequalities... It is easily seen that Mc Diarmid's inequality generalizes Hoeffding's one since it involves a supremum bound for the variation of each coordinate of the function. On the other hand Bennett's inequality takes into account the variance of the random variables and we can expect it to be more precise in many situations. Here is this inequality in its general formulation (the notations are those of Ledoux-Talagrand (1991)[17]) :

$$\mathbb{P} \left( \left| \sum_i V_i \right| > t \right) \leq 2 \exp \left( \frac{t}{a} - \left( \frac{t}{a} + \frac{b^2}{a^2} \right) \log \left( 1 + \frac{at}{b^2} \right) \right) \quad (2)$$

Where the independent and **centered** real random variables  $V_i$ 's are such that  $|V_i| \leq a$  almost surely and  $b^2 = \sum_i \mathbb{E}V_i^2$ .

The preceding bound may be rearranged in a more practical way. Indeed if  $\mathbb{E}V_i^2 = \sigma^2$  for all  $i$ . Then  $b^2 = n\sigma^2$  and setting  $x = at/\sigma^2$

$$\mathbb{P} \left( \left| \sum_i V_i \right| > nt \right) \leq 2 \exp \left[ -\frac{n\sigma^2}{a^2} ((1+x) \log(1+x) - x) \right] \quad (3)$$

We have already seen above and will make more precise below that one of our goals would be to determine the smallest threshold  $t$  (or equivalently  $x$ ) such that the term of the right in the above display is less than  $\alpha$  ( $\alpha$  is the type I error). The solution is

$$x_{th}^* = \min \left\{ x > 0 : 2 \exp \left[ -\frac{n\sigma^2}{a^2} ((1+x) \log(1+x) - x) \right] \leq \alpha \right\} \quad (4)$$

Let us denoted  $\varphi$  the positive non-decreasing, convex, one to one function defined on the set of positive real numbers by  $\varphi(x) = (1+x) \log(1+x) - x$ . This

function will be referred to as the "concentration function", even if this term does not match the usual probabilistic meaning. If the inverse of function  $\varphi$  was explicitly known an obvious solution would be  $x_{th}^* = \varphi^{-1} \left( \frac{a^2}{n\sigma^2} \log \left( \frac{2}{\alpha} \right) \right)$ . But  $\varphi^{-1}$  cannot be explicitly computed. We will propose an explicitly known function  $\varphi^\dagger$  such that  $\varphi^\dagger \geq \varphi^{-1}$  and we will choose  $x^* = \varphi^\dagger \left( \frac{a^2}{n\sigma^2} \log \left( \frac{2}{\alpha} \right) \right)$ . Obviously the closer function  $\varphi^\dagger$  is from  $\varphi^{-1}$ , the closer  $x^*$  is from  $x_{th}^*$  ( $x^* \geq x_{th}^*$ ) and the better our threshold. Before tailoring  $\varphi^\dagger$  we underline important features of the bound (3).

### 3.2 From Gaussian to Poisson bounds

It is often noted in the literature on concentration inequalities (or on large deviation theory) that, depending on the value of  $t$ , exponential bounds such as (2) may be approximated the following way:

- when  $t$  is small, namely if  $at/b^2 \leq 1/2$

$$\mathbb{P} \left( \left| \sum_i V_i \right| > t \right) \leq 2 \exp(-t^2/4b^2) \quad (5)$$

which illustrates the gaussian behavior of the random sum.

- when  $t$  is large the following inequality is sharper

$$\mathbb{P} \left( \left| \sum_i V_i \right| > t \right) \leq 2 \exp \left( -\frac{t}{a} \left( \log \left( 1 + \frac{at}{b^2} \right) - 1 \right) \right) \quad (6)$$

and is referred to as a Poisson type behavior.

The main problem relies on the missing link between the two displays above: what should we do when  $t$  is neither large nor small? What does even mean "large" and "small" with respect to our problem. We answer these questions in the next section. However we notice that Bernstein [16] proved that, uniformly in  $x \in \mathbb{R}$ ,  $\varphi(x) \geq \frac{x^2}{2(1+x/3)}$ . This bound is sharp for small  $x$  -and would lead to an explicit  $\varphi^\dagger$  -but not for large  $x$ . The function  $\varphi^\dagger$  we propose below strictly reflects both "domains": Poisson and Gaussian and even define an intermediate area.

## 4 Approximating the inverse concentration function

### 4.1 Tailoring the function

We set

$$\varphi^\dagger(y) = \begin{cases} \varphi_1^\dagger(y) = \frac{y}{3} + \sqrt{\frac{y^2}{9} + 2y} & \text{if } 0 \leq y \leq 3(\log 3) - 2 \\ \varphi_2^\dagger(y) = \frac{y + 2.3}{\log 3.3} - 1 & \text{if } 3(\log 3) - 2 \leq y \leq 4(\log 4) - 3 \\ \varphi_3^\dagger(y) = \frac{y}{\log\left(\frac{y}{\log y} + 1\right)} \left(1 + \frac{1}{\log y}\right) - 1 & \text{if } 4(\log 4) - 3 \leq y \end{cases} \quad (7)$$

The functions  $\varphi_1^\dagger$  (resp.  $\varphi_3^\dagger$ ) lies in the Gaussian (resp Poissonian) domain and  $\varphi_2^\dagger$  features an intermediate domain as announced sooner in the paper.

Now we quickly explain how function  $\varphi^\dagger$  was derived :

- Bernstein bound directly provides  $\varphi_1^\dagger$ : Indeed if  $\varphi(x) \geq \frac{x^2}{2(1+x/3)}$ ,  $\varphi^{-1}(y) \leq \varphi_1^\dagger(y)$
- Function  $\varphi_3^\dagger$  is obtained as a by product of Newton algorithm applied to (4). Indeed we seek  $\min\{x > 0 : ((1+x)\log(1+x) - x) \geq M\}$ , where  $M$  is known and (4) is rewritten a simpler way. Straightforward calculations prove that  $x_0(M) = M/\log(M)$  is a good estimate for large  $M$ . Then iterating once Newton's algorithm from  $x_0(M)$ , we seek the intersection of the tangent at  $x_0(M)$  to function  $\varphi$  and the line  $y = M$ . Denote this point  $(\kappa(M), M)$ . We just set  $\varphi_3^\dagger(y) = \kappa(y)$ . The good behavior of this estimate is due to the slow rate of increase of  $\varphi'(x) = \log(1+x)$ .
- Function  $\varphi_2^\dagger$  linearly interpolates between  $\varphi_1^\dagger$  and  $\varphi_3^\dagger$ .

### 4.2 Measuring the error

Many procedures could be proposed to measure the error between the theoretical threshold  $x_{th}^*$  and our estimate  $x^*$ . We propose the following which is quite intuitive and global.

Take an  $x$ , whose image by  $\varphi$  is  $y = \varphi(x)$ . The approximation error by  $\varphi^\dagger$  is clearly  $|x - \varphi^\dagger(\varphi(x))|$ .

**Definition 1.** *As a criterion for measuring the goodness of our approximating function we take  $\mathcal{C} = \sup_{x \in \mathbb{R}} \frac{|x - \varphi^\dagger(\varphi(x))|}{x}$ .*

The next Proposition enlightens the choice of the numerical constants in (7)

**Proposition 1.** *If  $\varphi^\dagger$  is chosen as in (7), then  $\mathcal{C} \leq 0.055$*

In other words, our method provides a threshold which approximates the optimal one just inducing a 5% error (about 5% in fact...).

*Remark 1.* Obviously the bound on  $\mathcal{C}$  given above may be improved (and made smaller than the 0.05 threshold) by adding a fourth function, say  $\varphi_{2.5}^\dagger$ , between  $\varphi_2^\dagger$  and  $\varphi_3^\dagger$ .

## 5 Main result

### 5.1 Type I&II errors

Obviously the greylevel of any region may be viewed as a discrete random variables almost surely bounded by  $a = 255$ . Assume first that the variances of both regions (denoted  $\sigma^2$  and  $\sigma'^2$ ) are non null. Here is the first result related to error of type I : Set  $M = \frac{a^2}{|R| \sigma^2} \log\left(\frac{4}{\alpha}\right)$ ,  $M' = \frac{a^2}{|R'| \sigma'^2} \log\left(\frac{4}{\alpha}\right)$ , and  $t_I(\alpha) = \frac{\sigma^2}{a} \varphi^\dagger(M) + \frac{\sigma'^2}{a} \varphi^\dagger(M')$ .

**Theorem 1.** *When  $m = m'$ ,  $\mathbb{P}(|S - S'| > t_I(\alpha)) \leq \alpha$ .  
In other words if  $|S - S'| \leq t_I(\alpha)$  we accept the merging of both regions*

*Remark 2.*  $M$  and  $M'$  both depend on  $\sigma^2$  and  $\sigma'^2$  which are unknown and will be approximated by  $\frac{1}{|R|} \sum_{i=1}^{|R|} (X_i - \bar{X})^2$  and  $\frac{1}{|R'|} \sum_{i=1}^{|R'|} (X'_i - \bar{X}')^2$  respectively.

Now we focus on the error of type II whose probability was denoted  $\beta$  : accepting  $H_0$  wrongly. At this point we need to address a crucial issue. We try to avoid overmerging and we should first define what we mean by "distinct" regions. We guess that if we take two perfectly homogeneous regions says  $R$  with all pixels taking values  $v$  and  $R'$  with all pixels taking values  $v + 1$ , the human eye may not consider them as "different". Consequently before going further we should define precisely what we mean by distinct regions. We introduce the parameter  $\Delta \in \mathbb{N}$ . as the minimum greylevel difference between two perfect regions beyond which the segmentation procedure should always "refuse" a merging.

In other words and with mathematical symbol : if  $|\mathbb{E}X - \mathbb{E}X'| > \Delta$  we assume that the average greylevel is too different for both regions to be merged. Obviously  $\Delta$  is a tuning parameter that will first of all depend on the "universal human eye" but also on the type of images that have to be analyzed.

We are ready to state the next Theorem providing the threshold for type II error.

**Theorem 2.** *When  $|\mathbb{E}X - \mathbb{E}X'| \geq \Delta$  (i.e. when  $R$  and  $R'$  should be considered as distinct regions),*

$$\mathbb{P}(|S - S'| < \Delta - t_I(\beta)) < \beta$$

*and if  $|S - S'| \geq \Delta - t_I(\beta)$  we reject the merging.*

*Remark 3.* In fact the previous Theorem bounds  $\sup_{|\mathbb{E}X - \mathbb{E}X'| \geq \Delta} \mathbb{P}(|S - S'| \leq t)$  by  $\beta$ . This probability is nothing but the maximal probability of error for merging two distinct regions whereas they should not.

We assumed that  $\sigma \neq 0$  and  $\sigma' \neq 0$ . In practice it may happen that either one or both variances are null. This situation is encountered when all pixels of one (or both) of the regions have the same greylevel (say  $m$  and  $m'$ ). If both variance are equal we propose to merge them iff  $|m - m'| \leq \Delta$ . Now if only one of the variances is null (say  $\sigma' = 0$  and  $\sigma \neq 0$ ) then the region  $R'$  will be seen as a perfect region. Our concentration inequalities still hold but in a one sample instead of the two sample framework above. A quick inspection of the proof of [Theorem 1](#) shows that the bound that appears there must be replaced by

$$\tilde{t}_I(\alpha) = \frac{\sigma^2}{a} \varphi^\dagger(M) \text{ and } \tilde{t}'_I(\alpha) = \frac{\sigma'^2}{a} \varphi^\dagger(M') \quad (8)$$

with  $M$  like at [subsection 5.1](#).

## 5.2 The segmentation process

From what was done above the segmentation process decision could be the following :

- If  $|S - S'| \leq t_I(\alpha)$  merge  $R$  and  $R'$
- If  $|S - S'| \geq \Delta - t_I(\beta)$  do not merge  $R$  and  $R'$

We know that in the first case (resp. the second), the probability of error is less than  $\alpha$  (resp.  $\beta$ ). But for both conditions not to be contradictory the following must hold :

$$\Delta \geq t_I(\alpha) + t_I(\beta) \quad (9)$$

An ultimate question arises : what should we do if  $t_I(\alpha) < |S - S'| < \Delta - t_I(\beta)$ ? The interval  $[t_I(\alpha), \Delta - t_I(\beta)]$  is sometimes referred to as the indifference zone in statistical testing theory.

We propose to base this last step on an homogeneity test for the second moment (we could also have investigated a randomized test but this approach is less intuitive). Since we cannot make a decision based on means we look at the second order moments : if they are close enough we accept the merging of both regions. Set  $V = \frac{1}{|R|} \sum_{i=1}^{|R|} X_i^2$  and  $V' = \frac{1}{|R'|} \sum_{i=1}^{|R'|} (X'_i)^2$ . Applying Mc Diarmid's (see [\[18,13\]](#)) inequality to  $V$  and  $V'$  we get when  $\mathbb{E}X^2 = \mathbb{E}X'^2$  :

$$\mathbb{P}(|V - V'| > s) \leq \exp \left( - \frac{2s^2}{\frac{a^4}{|R|} + \frac{a^4}{|R'|}} \right)$$



This time the threshold on the second moment is

$$t_I(\gamma) = \sqrt{\frac{1}{2} \left( \frac{a^4}{|R|} + \frac{a'^4}{|R'|} \right) \log \left( \frac{1}{\gamma} \right)} \quad (10)$$

where  $\gamma$  is a prescribed probability.

To summarize the test procedure, we have four parameters  $(\alpha, \beta, \Delta)$  and  $\gamma$  where

- $\Delta$  is the value of intensity discrepancy that makes two regions being considered as belonging to two different objects in the image ; this threshold would be adapted to the dynamics of the image.
- $\alpha$  is the probability not to merge the two regions whereas we should.
- $\beta$  is the probability to merge the two regions whereas we should not.
- $\gamma$  is used in the indifference zone and is the probability that the difference of second order moments of the two regions is too large so that the regions should not be merge.

In order to achieve an optimal time segmentation, we base our algorithm on algorithms described in [19] and used by [6,12,8]. These algorithms are based on the following tests:

1. If  $|S - S'| \leq t_I(\alpha)$  merge  $R$  and  $R'$  where  $t_I(\alpha)$  is given in [Theorem 1](#).
2. If  $|S - S'| \geq \Delta - t_I(\beta)$  do not merge  $R$  and  $R'$  where  $t_I(\beta)$  is computed using same formula as  $t_I(\alpha)$  given in [Theorem 1](#).
3. If  $t_I(\alpha) < |S - S'| < \Delta - t_I(\beta)$  compute  $t_I(\gamma)$ ,  $V$  and  $V'$  and if  $|V - V'| > t_I(\gamma)$ , do not merge  $R$  and  $R'$  otherwise merge  $R$  and  $R'$ , where  $t_I(\gamma)$  is given by [Equation 10](#).

For the sake of efficiency we use algorithms proposed in [19] based on Union-Find data structure [20]. To ensure linear complexity, [19] claims that computing the criteria has to be done in constant time. Moreover, updating parameters of regions should also be done in constant time. If we look at formula of [Theorem 1](#) we can easily check that, whenever  $\sigma$  and  $|R|$  are known, computing  $t_I(\alpha)$  can be considered done in constant time (we suppose that in practice arithmetic operations, even log are done in constant time, which is not true in theory but almost true in practice with regard to other operations).

Knowing  $|R|$  is easy, it suffices to keep this number as parameter of each region. When a merging is done, we just add the two corresponding values to obtain  $|R|$  of the newly created region.

As noted in [remark 2](#),  $\sigma^2$  is approximated by  $\frac{1}{|R|} \sum_{i=1}^{|R|} (X_i - \bar{X})^2$  that can be set to  $\sigma^2 = \frac{\sum_{i=1}^{|R|} X_i^2}{|R|} - \bar{X}^2$  and so we only need to keep updated  $\sum_{i=1}^{|R|} X_i^2$  and  $\sum_{i=1}^{|R|} X_i$  which is as easy to maintain as  $|R|$  by the same process.

When we are located in the indifference zone, we must be able to compute in constant time  $V$ ,  $V'$  of and [Equation 10](#). It is simple to check, according

---

**Function Oracle( $R, R'$ )**

---

**Input:**  $R$  and  $R'$  two regions to be checked  
**Result:** **true** if they should be merged, else **false**

- 1 **if**  $(\sigma^2 = 0)$  **and**  $(\sigma'^2 = 0)$  **then return**  $|m - m'| \leq \Delta$ ;
- 2 **else if**  $\sigma^2 = 0$  **then return**  $|S' - m| \leq \tilde{t}_I(\alpha)$ ;
- 3 **else if**  $\sigma'^2 = 0$  **then return**  $|S - m'| \leq \tilde{t}_I(\alpha)$ ;
- 4 **else if**  $|S - S'| \leq t_I(\alpha)$  **then return true**;
- 5 **else if**  $|S - S'| \geq \Delta - t_I(\beta)$  **then return false**;
- 6 **else** /\* **indifference zone:**  $t_I(\alpha) < |S - S'| < \Delta - t_I(\beta)$  \*/
- 7   **return**  $|V - V'| \leq t_I(\gamma)$ ;

---

to previous remarks, that these equations can also be computed in practical constant time. So we can define a function `Oracle`, based on the 3 tests described at end of [subsection 5.2](#), which decides if two region  $R$  and  $R'$  should be merge or not. According to this `Oracle` (see [function Oracle](#)), `Scanline` and `MergeSquare` algorithms of [19] perform in linear times and so are very fast.

## 6 Discussion and comparison with previous works

Our threshold cannot always be compared with the "theoretical" one proposed by Nock and Nielsen (2005) [10]. Since it is based on a different concentration inequality. However for the sake of completeness it may be of interest to compare both threshold. The analytic comparison is really uneasy, so

we plotted this function for  $|R| = |R'| = r$ , ranging from 1 to 40 (there was no need to go beyond since  $\theta(|r|)$  decreases quickly) and for different values of  $\sigma^2$  ranging from 1000 to 16000. It turns out that our threshold is less sharp than the one by Nock and Nielsen only for very small regions and works better for small  $\sigma$ . We check that introducing the variance in the concentration inequality provides adaptivity and an extra amount of sharpness.

## 7 Experimental results and practical improvements

Our main test is  $t_I(\alpha)$  but it requires that  $\sigma \neq 0$  and  $\sigma' \neq 0$ . If we look an image with low dynamic such as *cornouaille* image (see [figure 1a](#)) and perform our segmentation process directly we obtain image at [figure 1b](#) which clearly shows effect of scanning order. In fact this effect is amplified by the fact that dynamic is low and so a lot of regions have initially null variance. To avoid this bias, we propose to pre-segment image by grouping very similar pixels into small regions by using same algorithm but with an Oracle looking only at greyvalue of pixels.

Result can be seen at [figure 1c](#). It can be compare to result with original scanline process of [19] with same value of  $\Delta$  presented [figure 1d](#).

We have also done some comparisons with work of [6] which first introduce concentration inequality for segmentation and from which are derived numerous

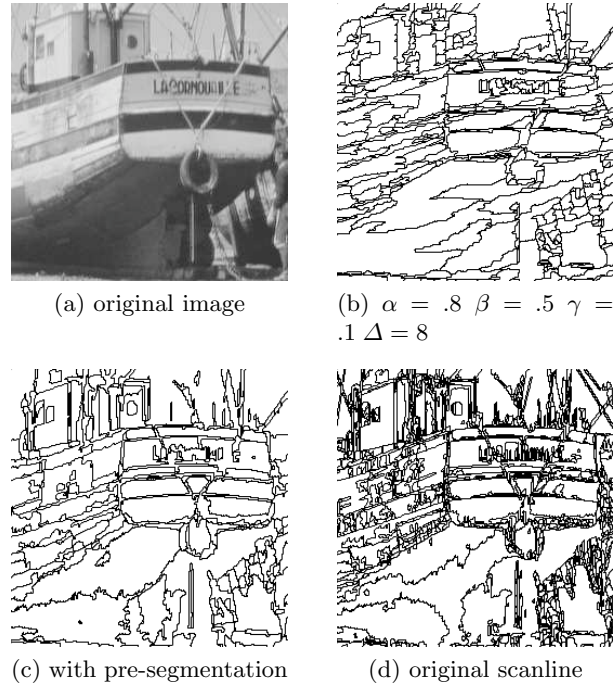


Fig. 1: Comparison with original scanline and without pre-segmentation.

papers of R. Nock [12,8]. For that, we use a very colorful image and perform segmentation by applying our algorithm to each channel and comparing them as R. Nock in [5]. As algorithm is basically designed for greylevel images, we need to relax condition on  $\alpha$  to have consistent results on the three RGB channels. So we set  $\alpha = \beta = \gamma = 0.10$  and compare the result (see figure 2b with Mc Diarmid segmentation criterion in [6,12,8] with  $Q$  parameter set to 8 (see figure 2c).

## 8 Conclusion and further work

We proposed in this paper a merging criterion based on a statistical test using local properties but preserving global properties. Moreover an efficient region segmentation algorithm, based on this test, has been given. The originality of this work lies in the fact that we take into account the two classical types of error for statistical tests. This leads to an adaptive multi-test which decides if regions should be merged, should not, or if we are in indifference zone where both decisions could be accepted. For this last case, we have proposed a way to finally decide by introducing another parameter and looking at second order moments.

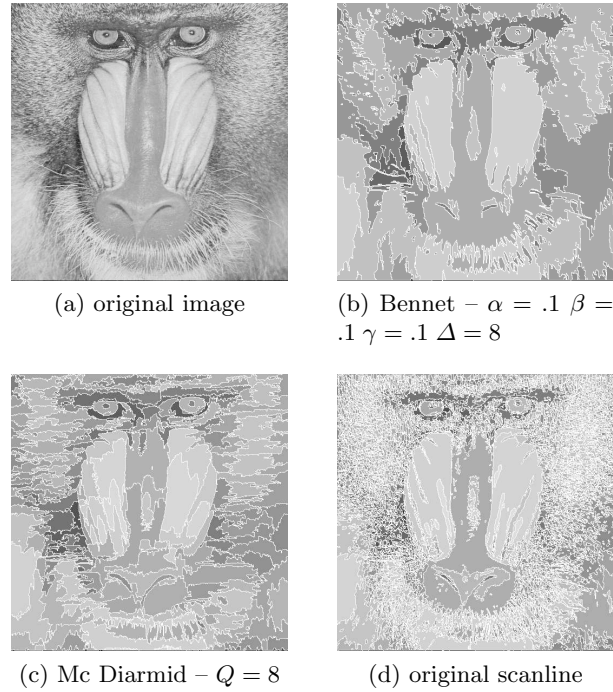


Fig. 2: Comparison with previous works.

Thus the algorithm we propose requires the setting of four parameters  $\alpha, \beta, \gamma, \Delta$ . In fact, the two first parameters need not to be tuned precisely. They even can be fixed whatever is the image since they are related to the probability of making error and so denote the trust you want in your test. Classical value in statistical applications is 5%.  $\gamma$  and  $\Delta$  are more sensitive parameters but are far more intuitive than the  $Q$  parameter of [6,12,8] since they describe the perceptual separability property of grey levels. Then, it has to be related to the image analyzed. We have proposed a first improvement by doing a pre-segmentation in order to decrease impact of  $\Delta$  parameter and scanning order. Another way to improve results is to sort regions before merging them in order to reduce the order dependence as done in [12,8]. Instead of sorting, we could also carry out seeded region merging. According to [19], the algorithm will loose its linear time complexity but will stay near linear as proved by R.E. Tarjan in [20] and so keep almost all of its efficiency. In the present paper, criteria are based on first and second order moment of the regions, but as `Scanline` algorithm scans regions starting from pixels, we could also take into account local pixel criteria to improve the test in the case of indifference zone.

## References

1. Pavlidis, T.: Segmentation of pictures and maps through functional approximation. *Computer Graphics and Image Process.* **1** (1972) 360–372
2. Zucker, S.W.: Survey: Region growing: Childhood and adolescence. *Computer Vision, Graphics, and Image Processing.* **5** (1976) 382–399
3. Forsyth, D.A., Ponce, J.: *Computer Vision: A Modern Approach.* Prentice Hall (2003)
4. Monga, O.: An optimal region growing algorithm for image segmentation. *International Journal of Pattern Recognition and Artificial Intelligence* **1** (1987) 351–375
5. Nock, R.: Fast and reliable color region merging inspired by decision tree pruning. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)* **1** (2001) I-271–I-276 vol.1
6. Fiorio, C., Nock, R.: Image segmentation using a generic, fast and non-parametric approach. In: *Tools with Artificial Intelligence, 10th IEEE International Conference on, Taipei, Taiwan, R.O.C., IEEE Computer Society* (1998) 450–458
7. Fiorio, C., Nock, R.: A concentration-based adaptive approach to region merging of optimal time and space complexities. In: *British Machine Vision Conference. Volume 02., Bristol, England* (2000) 775–784
8. Nock, R., Nielsen, F.: Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intell.* **26** (2004) 1452–1458
9. Lehmann, E.L.: *Testing statistical hypotheses.* Second edition edn. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons, New York, USA (1986)
10. Nock, R., Nielsen, F.: Semi-supervised statistical region refinement for color image segmentation. *Pattern Recognition: Image Understanding for Photographs* **38** (2005) 835–846
11. Demaine, E.D., Emanuel, D., Fiat, A., Immorlica, N.: Correlation clustering in general weighted graphs; approximation and online algorithms. *Theoretical Computer Science* **361** (2006) 172–187
12. Fiorio, C., Nock, R.: Sorted region merging to maximize test reliability. In: *International Conference on Image Processing. Volume 01., Vancouver, Canada, IEEE* (2000) 808–811
13. Mc Diarmid, C.: Concentration for independent permutations. *Comb. Probab. Comput.* **11** (2002) 163–178
14. Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of American Statistical Association* **58** (1963) 13–30
15. Bennett, G.: Probability inequalities for the sum of independent random variables. *Journal of American Statistical Association* **57** (1962) 33–45
16. Bernstein, S.: On a modification of chebyshev’s inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math* **1** (1924)
17. Ledoux, M., Talagrand, M.: *Probability in banach spaces. isoperimetry and processes.* *Ergebnisse der Mathematik und ihrer Grenzgebiete* **3** (1991) xii+480
18. Mc Diarmid, C.: Concentration. In: *Habib, Diarmid, M., Ramirez-Alfonsin, Reed, eds.: Probabilistic methods for algorithmic discrete mathematics, New-York, Springer* (1998) 195–248
19. Fiorio, C., Gustedt, J.: Two linear time Union-Find strategies for image processing. *Theoretical Computer Science* **154** (1996) 165–181
20. Tarjan, R.E.: Efficiency of a good but not linear set union algorithm. *J. of the Association for Computing Machinery* **22** (1975) 215–225