A Convex Formulation for Global Histogram Based Binary Segmentation

Romain Yıldızoğlu, Jean-François Aujol, and Nicolas Papadakis

Univ. Bordeaux, Institut de Mathématiques de Bordeaux (IMB, UMR 5251),
351 Cours de la Libération F-33400 Talence, France
{romain.yildizoglu,jean-francois.aujol,
nicolas.papadakis}@math.u-bordeaux1.fr
http://www.math.u-bordeaux1.fr/imb

Abstract. In this paper, we present a general convex formulation for global histogram-based binary segmentation. The model relies on a data term measuring the histograms of the regions to segment w.r.t. reference histograms as well as TV regularization allowing the penalization of the length of the interface between the two regions. The framework is based on some l^1 data term, and the obtained functional is minimized with an algorithm adapted to non smooth optimization. We present the functional and the related numerical algorithm and we then discuss the incorporation of color histograms, cumulative histograms or structure tensor histograms. Experiments show the interest of the method for a large range of data including both gray-scale and color images. Comparisons with a local approach based on the Potts model or with a recent one based on Wasserstein distance also show the interest of our method.

Keywords: segmentation, global histogram, convex, structure tensor, non smooth optimization.

1 Introduction

Image segmentation has been the subject of active research for more than 20 years (see e.g. [2,11] and references therein). For instance, we can refer to the seminal work of Mumford and Shah [16], or to its very popular approximation with level sets developed by Chan and Vese in [9]. This last work provides a very flexible algorithm to segment an image into two homogeneous parts, each one being characterized by its mean gray level value.

In the case of textured images, a lot of extensions of [9] have been proposed to enhance the mean value image segmentation model. For instance, local histograms are used in [31,17], Gabor filters in [27] and [20], wavelet packets in [3] and textures are characterized thanks to the structure tensor in [6,26].

When considering the global histograms of the regions to segment, there also exist a large body of literature [1,7,15,14] also based on [9]. Recent works make use of the Bhattacharyya distance [30] or the Wasserstein distance [21] to compare globally the histograms. It is important to notice that this class of approaches involves complex shape gradient computations [12] for the level set

evolution equation. Moreover, as these methods all rely on the evolution of a level set function [19], it leads to non convex methods that are sensible to the initialization choice and only a local minimizer of the associated energy is computed. Other models as in [25,28,4,13] that use graph-based methods and also [24] obtain good results without level-sets, but these algorithms are not bound to give a global minimum of the original energy.

Recently, convexification methods have been proposed to tackle this problem, as in [18,23,8,5,29]. The original Chan-Vese model [9] can indeed be convexified, and a global solution can be efficiently computed, for instance with a primal-dual algorithm. A simple thresholding of this global solution provides then a global minimizer of the original non convex problem. Up to our knowledge, such approaches have not been developed yet for global histogram segmentation with length boundary regularization.

The contribution of this paper is thus to introduce a convex model to segment an image into two parts, each region being characterized by its global histogram. This convex model is minimized efficiently by using a non smooth convex optimization algorithm [8]. The model is first developed for the case of grayscale images, and then extended to the case of features based on the structure tensor, and finally to the case of color images.

With respect to the previously mentioned global histogram based segmentation methods relying on the evolution of a level set function, the approach here is much simpler. On the one hand, the proposed algorithm is faster than the ones based on Wasserstein distance [17,21], and on the other hand despite the simplicity of the approach the obtained segmentation results are very good while being independent of the initialization.

The plan of the paper is the following. We introduce our model for global histogram based binary segmentation in Section 2. Our functional relies on some l^1 norm for the data term, and we propose an efficient numerical scheme based on a non smooth convex optimization algorithm. We present in Section 3 numerous numerical examples to show the strength of the method, as well as its limitations. In Section 4, we illustrate the importance of global histogram comparisons to get a global image segmentation and we give some comparisons with the approach of [21]. We conclude in Section 5 and we present some future venues of research.

2   A Convex Variational Formulation

The problem tackled in this paper concerns the segmentation of an image into two parts through global histogram constraints. Our approach is motivated by two issues. The first one concerns the dependence to the initialization, the second is the problem of locality in the histogram comparison.

Notations

Let Ω be the image domain: we assume Ω to be a non empty open bounded subset of $\mathbb{R}^2$ with Lipschitz boundary. We note $|\Omega|$ the size of Ω and $\langle .,. \rangle$ the

standard inner product on $L^2(\Omega)$, that is $\langle u, v \rangle = \sum_{x \in \Omega} u(x)v(x)$ in the discrete setting. Let $I : \Omega \mapsto \Lambda \subset \mathbb{R}^k$ be the image, Λ finite, h^0 and h^1 two given reference histograms: $\Lambda \to [0;1]$, with $\sum_{\lambda \in \Lambda} h^i(\lambda) = 1$, $i = 0, 1$. We aim at estimating a binary segmentation represented by $u : \Omega \to \{0, 1\}$, where the histogram computed on the region $\Omega_0 := \{x \in \Omega, u(x) = 0\}$ is close to h^0 (resp. the histogram on $\Omega_1 := \{x \in \Omega, u(x) = 1\}$ should be close to h^1). The histogram on the region Ω_1 is then computed as:

$$h_u(\lambda) = \frac{1}{|\Omega_1|} \sum_{x \in \Omega} u(x) \mathbb{1}_{I=\lambda}(x) = \frac{1}{\sum_{x \in \Omega} u(x)} \sum_{x \in \Omega} u(x) \mathbb{1}_{I=\lambda}(x) \qquad (1)$$

To realize the segmentation, we first require a metric between histograms, and we will thus consider a norm $\| \cdot \|$ on $\mathbb{R}^\Lambda$. In order to control the length of the interface between the two parts of the partition, a total variation regularization is also considered. The segmentation can therefore be obtained by minimizing the following non-convex energy :

$$J(u) = TV(u) + \left|\left|(h_u - h^1)_{\lambda \in \Lambda}\right|\right| + \left|\left|(h_{1-u} - h^0)_{\lambda \in \Lambda}\right|\right| \qquad (2)$$

Convexification

In the aim of defining a convex model, we propose to reformulate the data term comparing histograms. To that end, we can first observe that

$$\| h_u - h^1 \| = \left|\left|\left(\frac{1}{\sum_\Omega u(x)} \sum_\Omega u(x) \mathbb{1}_{I=\lambda}(x) - h^1(\lambda)\right)_{\lambda \in \Lambda}\right|\right| \qquad (3)$$

$$= \left|\left|\frac{1}{\sum_\Omega u(x)} \left(\sum_\Omega u(x) \mathbb{1}_{I=\lambda}(x) - \left(\sum_\Omega u(x)\right) h^1(\lambda)\right)_{\lambda \in \Lambda}\right|\right| \quad (4)$$

Assuming that the size of the area Ω_1 defined by $|\Omega_1| = \sum_\Omega u(x)$ is known, we can only keep the distance

$$\left|\left|\left(\sum_\Omega u(x) \left(\mathbb{1}_{I=\lambda}(x) - h^1(\lambda)\right)\right)_{\lambda \in \Lambda}\right|\right| \qquad (5)$$

which is convex in u. Note that a similar problem is tackled in [10] in the slightly different framework of cosegmentation with no total variation, and a solution is found without modifying this term. The very same convex reformulation can be done for the second term involving $1 - u$ and the histogram h^0, thus removing a factor $\frac{1}{\sum_\Omega (1-u(x))} = \frac{1}{|\Omega_0|}$.

With such normalizations, the data terms of the two partitions are no more balanced. Nevertheless, a weighting factor $\beta \in [0, 1]$ can be introduced. This factor represents the ratio $\beta = \sum_\Omega u(x)/|\Omega| = |\Omega_1|/|\Omega|$. As $|\Omega| = |\Omega_0| + |\Omega_1|$, one can see that $\sum_\Omega (1 - u(x))/|\Omega| = |\Omega_0|/|\Omega| = (|\Omega| - |\Omega_1|)/|\Omega| = 1 - \beta$, which gives the normalization factor of the second term.

Defining $g_\lambda^1(x) := \mathbb{1}_{I=\lambda}(x) - h^1(\lambda)$, $g_\lambda^0(x) := \mathbb{1}_{I=\lambda}(x) - h^0(\lambda)$, the final convex model reads:

$$J(u) = TV(u) + \frac{1}{\beta} \parallel (\langle u, g_\lambda^1 \rangle_\Omega)_{\lambda \in \Lambda} \parallel + \frac{1}{1-\beta} \parallel (\langle 1-u, g_\lambda^0 \rangle_\Omega)_{\lambda \in \Lambda} \parallel . \quad (6)$$

It now remains to choose the distance to compare histograms.

l^1 Data Term

Up to now, the general model has been designed for any data norm. In this section, we choose the l^1 norm: $\parallel h \parallel_{l^1} = \sum_{\lambda \in \Lambda} |h(\lambda)|$. We therefore have the following energy:

$$J(u) = TV(u) + \mu \frac{1}{\beta} \sum_{\lambda \in \Lambda} |\langle u, g_\lambda^1 \rangle_\Omega| + \mu \frac{1}{1-\beta} \sum_{\lambda \in \Lambda} |\langle 1-u, g_\lambda^0 \rangle_\Omega| \quad (7)$$

$$= TV(u) + \mu \frac{1}{\beta} \sum_{\lambda \in \Lambda} \left| \sum_{x \in \Omega} u(x) g_\lambda^1(x) \right| + \mu \frac{1}{1-\beta} \sum_{\lambda \in \Lambda} \left| \sum_{x \in \Omega} (1-u(x)) g_\lambda^0(x) \right| .$$

One can observe that our global data term depends on the whole state $\{u(x), x \in \Omega\}$, for each value $\lambda \in \Lambda$. We want to minimize $J(u)$ on the set $BV(\Omega, \{0,1\})$. As this binary set is not convex, we relax our problem by minimizing $J(u)$ on $BV(\Omega, [0,1])$. We will get back to the original domain by thresholding the solution of the relaxed problem. However, there is no guarantee that this strategy should lead to minimizers of the original problem, and different choices of thresholds may lead to different solutions.

Finally notice that if the original problem has many global minimizers, the result given by the convex minimization can be such that we do not know how to use it in order to get back to the original problem. Even if specific scenarios could be built to make such situation happen, this issue never occurred in our experiments with non-synthetic data.

Since our functional can be written as the sum of non-smooth convex terms, some of them being the composition of a l^1 norm and a linear operator, we will write the problem with dual variables. Let us write A and B the linear operators $A : u \mapsto (\langle u, g_\lambda^1 \rangle_\Omega)_{\lambda \in \Lambda}$, $B : u \mapsto (\langle u, g_\lambda^0 \rangle_\Omega)_{\lambda \in \Lambda}$. We can write the problem as a saddle point one

$$\arg \max_{q_1, q_2, q_3} \arg \min_u E(q_1, q_2, q_3, u) \quad (8)$$

where

$$E(q_1, q_2, q_3, u) = \langle q_1, \nabla u \rangle_\Omega + \langle q_2, Au \rangle_\Lambda + \langle q_3, B(1-u) \rangle_\Lambda \quad (9)$$
$$- \chi_{B(0,1)}(q_1) - \chi_{[-\mu\frac{1}{\beta}, \mu\frac{1}{\beta}]}(q_2) - \chi_{[-\mu\frac{1}{1-\beta}, \mu\frac{1}{1-\beta}]}(q_3) + \chi_{[0,1]}(u)$$
$$= \langle q, Ku \rangle_{\Omega \times \Lambda^2} - F^*(q) + G(u) \quad (10)$$

with $q = (q_1, q_2, q_3) \in (\mathbb{R}^\Omega \times \mathbb{R}^\Omega) \times \mathbb{R}^\Lambda \times \mathbb{R}^\Lambda$ of dimension $2|\Omega| + 2|\Lambda|$, $K : u \mapsto (\nabla u, Au, -Bu)$, $G(u) = \chi_{[0,1]}(u)$, $F^*(q) = \chi_{B(0,1)}(q_1) + \chi_{[-\mu\frac{1}{\beta}, \mu\frac{1}{\beta}]}(q_2) +$

$\chi_{[-\mu\frac{1}{1-\beta},\mu\frac{1}{1-\beta}]}(q_3) - \langle q_3, B1 \rangle$, $B(0,1)$ the closed unit disc of $\mathbb{R}^2$. In the previous expressions, χ_C denotes the characteristic function of the convex set C, i.e. $\chi_C(x) = 0$ if $x \in C$, $\chi_C(x) = \infty$ if $x \notin C$. We optimize this criterion with a primal-dual scheme as explained here-after.

Optimization with a Preconditioned Primal-Dual Algorithm (PPD)

We now give the general formulation of the PPD algorithm, and then write the algorithm applied to our functional.

Let X, Y be two finite-dimensional vector spaces. We write $\langle ., . \rangle$ the standard inner products, $K : X \to Y$ a linear operator and $G : X \to \mathbb{R} \cup \{\infty\}$ and $F^* : Y \to \mathbb{R} \cup \{\infty\}$ some convex functions (F^* being the Legendre Fenchel transform of F [22]). We want to solve

$$\min_{x \in X} \max_{y \in Y} \langle Kx, y \rangle + G(x) - F^*(y). \tag{11}$$

In [22], a preconditioned primal-dual algorithm has been proposed to solve such convex problem. The algorithm reads:

Algorithm 1. Preconditioned primal-dual algorithm ([22])

$u^{k+1} = (I + T\partial G)^{-1} \left(u^k - TK^t z^k \right)$
$z^{k+1} = (I + \Sigma\partial F^*)^{-1} \left(z^k - \Sigma K(2u^{k+1} - u^k) \right)$

---

with T and Σ symmetric positive definite matrices such that $\| \Sigma^{\frac{1}{2}} K T^{\frac{1}{2}} \|^2 < 1$ and $(I + T\partial G)^{-1}(\hat{u}) := \arg\min_{u \in U} G(u) + \frac{1}{2\tau} \langle T^{-1}(u - \hat{u}), u - \hat{u} \rangle$. Such process converges to a saddle point in $O(\frac{1}{k})$ (with k the number of iterations). As proposed in [22], one can take $T = \text{diag}(\tau_1, ..., \tau_n)$ and $\Sigma = \text{diag}(\sigma_1, ..., \sigma_m)$ with $\tau_j < \frac{1}{\sum_{i=1}^{m} |K_{i,j}|}$, $\sigma_i < \frac{1}{\sum_{j=1}^{n} |K_{i,j}|}$.

Application of PPD to Functional (8)

Introducing the following relations:

$$\begin{cases} Tu(x) = \tau_x u(x), \qquad \tau_x < \frac{1}{2 + \sum_{\lambda \in \Lambda}(|g_\lambda^1(x)| + |g_\lambda^0(x)|)}, \\ A^t q_2 = \sum_{\lambda \in \Lambda} q_2(\lambda) g_\lambda^1, \\ B^t q_3 = \sum_{\lambda \in \Lambda} q_3(\lambda) g_\lambda^0, \\ \sigma_1 < \frac{1}{4}, \\ \Sigma_2 q_2(\lambda) = \sigma_{2,\lambda} q_2(\lambda), \quad \sigma_{2,\lambda} < \frac{1}{\sum_{x \in \Omega} g_\lambda^1(x)}, \\ \Sigma_3 q_3(\lambda) = \sigma_{3,\lambda} q_3(\lambda), \quad \sigma_{3,\lambda} < \frac{1}{\sum_{x \in \Omega} g_\lambda^0(x)}, \\ b = B1, \end{cases} \tag{12}$$

the PPD Algorithm 1 applied to our functional (7) gives the Algorithm 2, where P_C is the orthogonal projection on the convex C:

Algorithm 2.

$$u^{k+1} = P_{[0,1]} \left(u_k - T \left(A^t q_2^k - B^t q_3^k - \mathrm{div}(q_1^k) \right) \right)$$

$$q_1^{k+1} = P_{B(0,1)} \left(q_1^k + \sigma_1 \nabla(2u^{k+1} - u^k) \right)$$

$$q_2^{k+1} = P_{[-\mu \frac{1}{\beta}, \mu \frac{1}{\beta}]} \left(q_2^k + \Sigma_2 A(2u^{k+1} - u^k) \right)$$

$$q_3^{k+1} = P_{[-\mu \frac{1}{1-\beta}, \mu \frac{1}{1-\beta}]} \left(q_3^k - \Sigma_3 \left(B(2u^{k+1} - u^k) - b \right) \right).$$

3   Applications

We now present some experiments on synthetic and natural images, assuming that the histograms of the two areas to segment are known. They are in fact given by manual scribbles selected by the user in our practical applications (they are displayed on the images with red rectangles for the first region and green rectangles for the second one).

The balance between the data term and the regularization term is tuned by the user.

The parameter β is always fixed at 0.5, we show in Figure 4 that changing it can improve the result, but its value did not seem critical in our experiments. As our functional is convex, the process is simply initialized with $u = 0.5$ in all our experiments, which is a main difference with existing approaches based on global histograms [17,21,30]. The final solution is currently obtained by thresholding the estimated solution u^* with $u^* \geq 0.5$, the choice of the threshold leading to slightly different results.

3.1   Grayscale Images

In practical applications, images are always corrupted by some noise. In order to get a more robust algorithm, we can work with cumulative histograms instead of histograms. Situations may happen where a pixel value is represented in none of the reference distributions h^0 and h^1. For instance, let us consider that the reference histograms have been learned from a clean image. If this histogram is finely quantified and sparse, adding some noise (or an offset) to the image may lead to such issue. In this case, the values of a lot of noised pixels may belong to empty bins of both reference histograms and the segmentation will fail.

Cumulative histograms are one way to tackle this problem.

All the derivations we have made so far in the paper for histogram based segmentation hold true for cumulative histograms and we get the following functional:

$$J(u) = TV(u) + \mu \sum_{\Lambda} \left(\frac{1}{\beta} \left| \langle u, g_{C,\lambda}^1 \rangle_\Omega \right| + \frac{1}{1-\beta} \left| \langle 1 - u, g_{C,\lambda}^0 \rangle_\Omega \right| \right) \qquad (13)$$

with $g_{C,\lambda} = \sum_{\gamma \leq \lambda} g_\lambda^0 = \mathbb{1}_{I \leq \lambda} - h_C^0(\lambda)$, and the cumulative histogram is obtained as $h_C^0(\lambda) = \sum_{\gamma \leq \lambda} h^0(\gamma)$. The associated numerical schemes are therefore similar.

In order to show the robustness of cumulative histograms to noise, we segment in Figure 1 a synthetic image (a) with two areas that have the same color mean but different histograms with two non zero values, and we add a Gaussian noise to the image (b). We first estimate the reference histograms on the clean image. The objective is here to see if it is possible to retrieve the objects in the noisy image, knowing the histograms on the clean image. We show the results obtained with the data terms based on (c) histograms and (d) cumulative histograms. As expected, the simple histogram comparison fails, whereas the cumulative version of the model is able to give good segmentations of the objects.
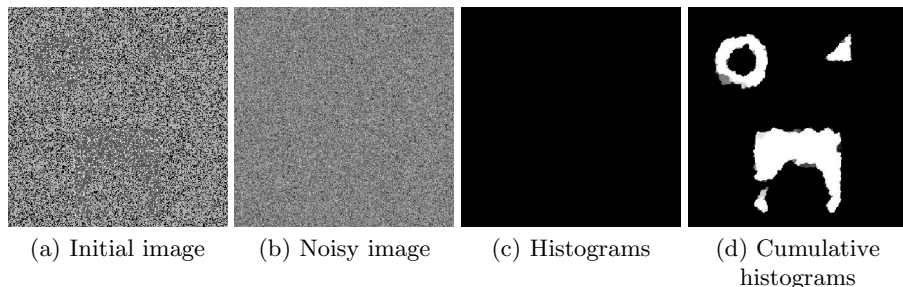

(a) Initial image   (b) Noisy image   (c) Histograms   (d) Cumulative histograms

Fig. 1. Segmentation of a noisy image with two zones that have the same color mean but different histograms, with $\beta = 0.5$, $\lambda = 0.07$ and 100 bins. The cumulative histograms version produces a good estimation, whereas the simple histogram comparison fails. The results are here not thresholded in order to show that the failure of the histogram model is not due to the choice of the threshold.

We will then use the cumulative histogram version of our data term, since it is well adapted to gray-valued images and it is more robust to noise. Notice that the l^1 data term between cumulative histograms of grayscale images corresponds to the l^1 Wasserstein distance. A segmentation of a natural image is presented in Figure 2. The animal is accurately segmented, even if the tail is missed.
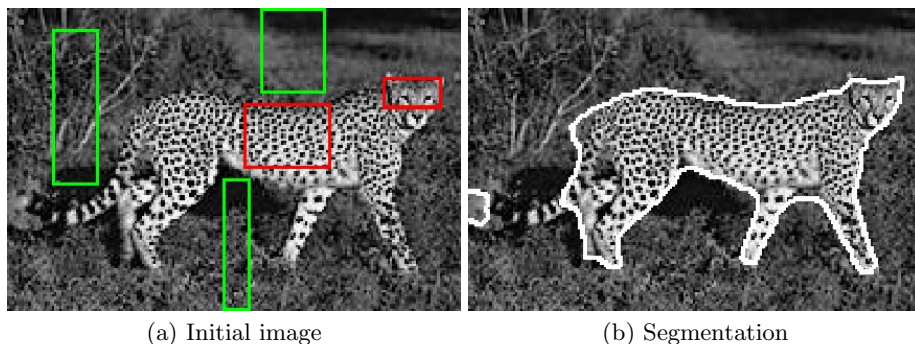

(a) Initial image   (b) Segmentation

Fig. 2. Segmentation of a natural image, with $\beta = 0.5$, $\lambda = 0.15$ and 10 bins. The animal is mainly found.

3.2   Color Images

We show in Figure 3 the segmentation of a zebra, using color histograms. Here again the global model produces accurate results. Next we present in Figure 4 the segmentation of a boat. In this example, the size of the two wanted partitions is unbalanced, as the boat is small compared to the rest of the image. A slight increase of the β parameter then allows one to obtain a more accurate estimation, as illustrated in the image (c).
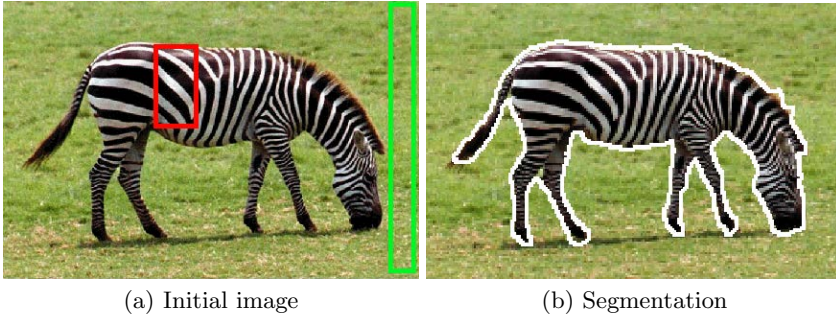

(a) Initial image   (b) Segmentation

Fig. 3. Segmentation of a natural image, with $\beta = 0.5$, $\lambda = 0.2$ and 7^3 bins. The animal is mainly found.

(a) Initial image   (b) Segmentation with   (c) Segmentation with
                                $\beta = 0.5$ $\beta = 0.59$
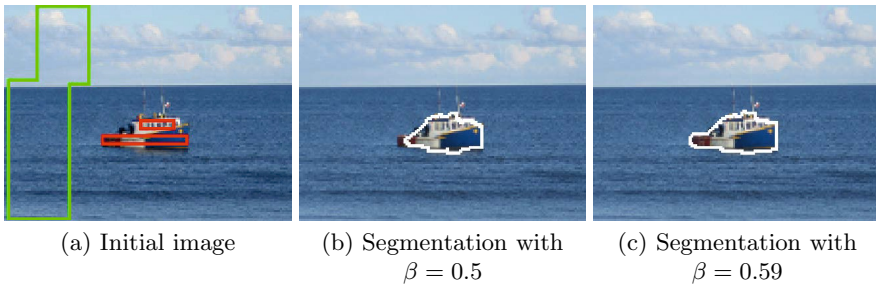
Fig. 4. Segmentation of a boat, with $\beta = 0.5$, $\lambda = 0.13$ and 5^3 bins. The boat is found, but choosing $\beta = 0.59$ gives a more accurate segmentation (c) than the one obtained with the default parameter (b).

3.3   Structure Tensor

The proposed model is also designed to deal with texture, using for instance the structure tensor [6] defined as follows. If I is a grayscale image, the structure tensor is the symmetric matrix $(\nabla I)(\nabla I)^t$. Since the tensor is symmetric, it has only 3 independent dimensions. A $3D$ histogram is therefore sufficient to represent the distribution of the structure tensor values.

We can therefore apply the algorithm on the three channels of the structure tensor or enhance the model with a 4 dimensional histogram (3 dimensions for

the structure tensor and 1 for the grayscale image I). Here we present some examples just involving the components of the structure tensor. We choose to linearly normalize all channels between 0 and 1 before taking the histograms. We give an example on a synthetic image with two regions that have the same histograms of gray levels, and with the same orientations. The results are given in Figure 5. The model here again performs well, even if the ring is recovered as a disc. We also present a segmentation of two images with two textures in Figures 6 where our global model is able to estimate accurate segmentations of the different textures.
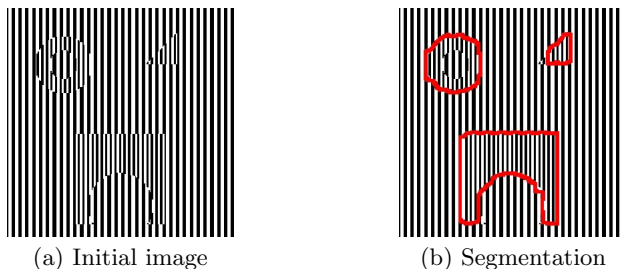

(a) Initial image   (b) Segmentation

Fig. 5. Segmentation of a synthetic image, with $\beta = 0.5$, $\lambda = 0.025$ and 3^3 bins

(a) Initial image   (b) Segmentation   (c) Initial image   (d) Segmentation
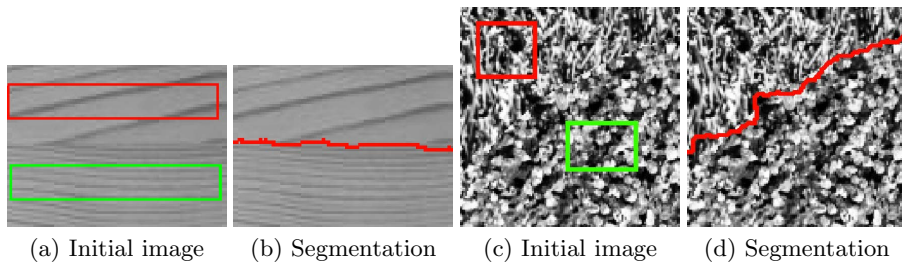
Fig. 6. Segmentation of an image of wood, with $\beta = 0.5$, $\lambda = 0.075$ and 7^3 bins, and segmentation of an image composed of two textures, with $\beta = 0.5$, $\lambda = 0.1$ and 7^3 bins.

4   Comparison with Local and Non Convex Approaches

In this section, we present comparisons with algorithms of the two main class of approaches that can be found in the literature of the histogram-based segmentation problem. We first compare our approach with an approach based on local histogram comparisons through the convex Potts model. As will be shown in the numerical computations, such a method cannot compete with the global framework introduced in this paper. We also present some comparisons with a recent non-convex approach based on the Wasserstein distance between global histograms. The different experiments demonstrate the capacity of our method to deal with global constraint while being independent of the initialization.

Comparisons with a Local Histogram-Based Approach

In order to illustrate the importance of global histogram comparisons over local ones, we now introduce a local histogram segmentation algorithm based on the Potts model. A pointwise estimation of the probability of a pixel to belong to a class can be formulated for instance with the functional:

$$J_{pointwise}(u) = TV(u) + \langle u(x), h_0(I(x)) - h_1(I(x)) \rangle_\Omega, \tag{14}$$

where the data model will enforce $u(x) = 1$ when $h_1(I(x)) > h_0(I(x))$. In the aim of having a more robust data term, we can realize a local estimation of the probability of a pixel to belong to a class. This can be formulated in the same way with the functional:

$$J_{local}(u) = TV(u) + \langle u(x), \| h_0 - h_{V(x)} \|_{l^1} - \| h_1 - h_{V(x)} \|_{l^1} \rangle_\Omega, \tag{15}$$

where $h_{V(x)}$ is the histogram estimated on a neighborhood $V(x)$ of $x \in \Omega$. Notice that in these two functionals, the data term can be computed pointwise (once the map of distances with local histograms has been calculated for the local model). This is a main difference with our data term which values depend on the whole state u. On the other hand, it is important to underline that such functionals can be globally minimized on $u \in BV(\Omega, \{0; 1\})$, using convexification approaches (see[18,23,8,5,29]) for more details). We illustrate the need to use global histogram comparisons instead of local ones on the two examples given in Figures 7 and 8.
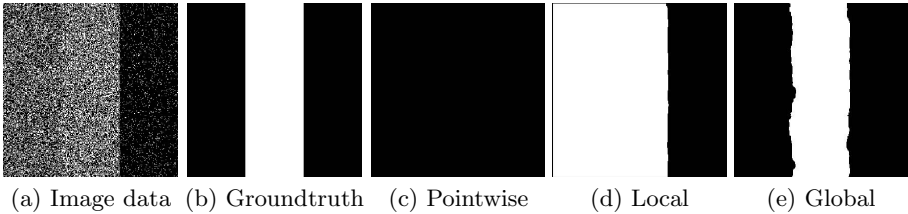

(a) Image data   (b) Groundtruth   (c) Pointwise   (d) Local   (e) Global

Fig. 7. Segmentation of a binary image with three regions. The respective probability of 0 are 0.65, 0.5, 0.95. We fix two histograms $h_0 = [0.8, 0.2]$ and $h_1 = [0.55, 0.45]$. The local estimation on neighborhoods of size 9×9 with $\lambda = 0.5$, gives a classification where the estimated histograms on the labeled images are not as close as the segmentation given in (c) with our model (with parameters $\beta = 0.5, \lambda = 0.5$).

The synthetic data presented in Figure 7 (a) is composed of three different regions of the same area. Each region is filled with independent realizations of a binomial distribution with the probability of 0 depending on the region: this probability is equal to 0.65 for the left region, 0.55 for the middle one and 0.95 for the right region. We fix two histograms, $h_0 = [0.8, 0.2]$ (i.e. the frequency of 0 is 0.8 and the frequency of 1 is 0.2) and $h_1 = [0.55, 0.45]$. Hence, the first histogram (h_0) exactly matches the expected histogram of the union of the left

and the right regions, while the second one (h_1) matches the one of the middle region. The ground truth segmentation is illustrated in image (b) of Figure 7.

When realizing a pointwise estimation of the data term, the pixels with value 0 are always more probable in the class defined by h_0. With no regularization the black pixels will therefore be classified in the h_0 class, and the white pixels in the h_1 class. Since in each model the probability of 0 is greater than the one of 1, increasing the regularization weight will increase the number of pixels labeled in h_0. This kind of classification is obtained with the functional (14) and illustrated in Figure 7 (c).

If we consider local histograms through the functional (15), the problem will be partially solved. Indeed, if the neighborhood used to compute the local histograms is sufficiently large, the data estimation will be good enough and the central region will be classified in h_1. However, if the metric used to compare histograms is the l^1 one, then the local histograms computed in the first region will be closer to h_1 and will thus be misclassified, as illustrated in image (d).

With both local models, the global histograms of the segmented regions are not close to the given reference histograms. This is the main difference with our modeling, as we aim at globally segmenting the image, in the sense that the histograms of the global segmented zones should be close to the reference distributions. We show in figure 7 (e) the aforementioned example, its segmentation with local histograms and the one with our model.

A second comparison is presented in Figure 8. The reference color distributions are here not homogeneous in the sense that the histogram of the whole region is not similar to the ones computed in local neighborhoods. The histograms are estimated in the regions given in the image (a). Here the orange colors are more probable in the region related to the butterfly, so in small neighborhoods the flowers are classified as the butterfly, and the darker regions are segmented as being in the background. This example illustrates again the importance of global histogram comparisons to get a global segmentation of an image. Indeed, our global model (c) is able to recover the butterfly , whereas the local approach (b) completely fails. Local approaches are therefore only relevant when the local histograms correctly approximate the global ones.

Comparisons with a Shape Gradient Approach

We finally illustrate the advantage of having a convex model that does not depend on the initialization. We compare our results with the ones obtained with the Wasserstein Active Contour method proposed in [21][1]. Such approach consists in deforming a level set function in order to minimize globally the Wasserstein distance between the reference histograms and the one induced by the segmentation. To make the level set evolve, this formulation requires complex shape gradients computations. In Figure 9, we present the results obtained with this approach on the synthetic three regions example (build as in Figure 7).

---

[1] We want to thank the authors of [21], Gabriel Peyré, Jalal Fadili and Julien Rabin, for their code.
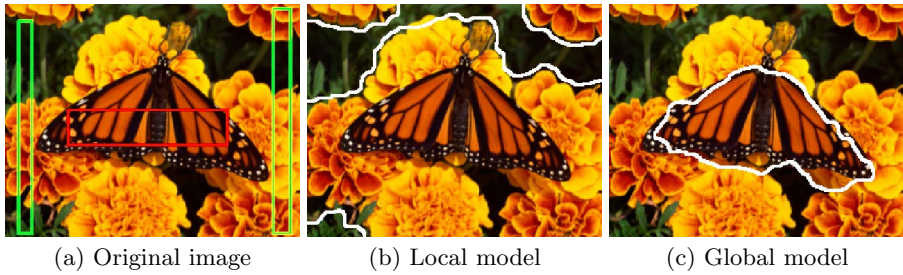
(a) Original image   (b) Local model   (c) Global model

Fig. 8. Segmentation of an image where local histograms (9×9 neighborhoods) are different from the global ones. The segmentation fails for the local histogram model (with $\lambda = 10$) as it classifies the orange areas in the first class and the darker ones in the second class. The global histograms on the segmented zones are not close to the given ones, contrary to our model parameterized with $\beta = 0.5$ and $\lambda = 0.22$. The histograms were composed of 4^3 bins.

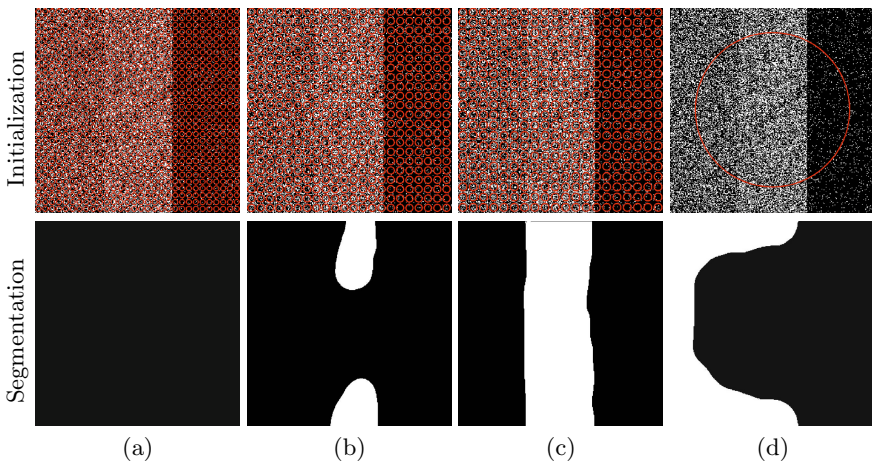

(a)   (b)   (c)   (d)

Fig. 9. We illustrate on the three regions synthetic example the problems arising with an algorithm that is very sensitive to the chosen initialization. The Wasserstein active contours method [21] have been initialized in different ways (first row, a-d). As illustrated in the second row, it leads to very different segmentations, even with really similar initializations (a-c). When carefully parameterized, such approach can lead to good segmentations as illustrated in column (c), which is very close to the one obtained with our global approach (see Figure 7 (c)).

An experiment on a natural image is finally shown in Figure 10. Even if this model can give good segmentations that are close to the ones we obtained in Figures 7 (c) and 8 (c), we illustrate that its initialization may be a critical step as really different segmentations are obtained with very similar initializations.
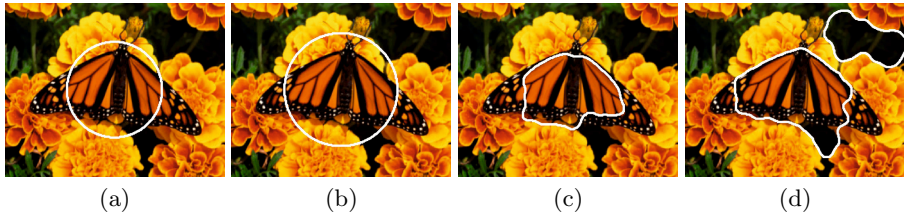
(a)   (b)   (c)   (d)

Fig. 10. The Wasserstein active contours method [21] have been initialized in two different ways (a-b), the corresponding segmentations being presented in (c-d). When carefully parameterized, it leads to a segmentation close to the one obtained with our global approach (see Figure 8 (c)).

5   Conclusions and Perspectives

In this paper, we have presented a general convex framework based on global histograms to segment images into two parts, and we experimented a special case defined by the l^1 data term. Our approach gives good results in the set of experiments realized on gray images, color images and also when using the structure tensor. The additional use of the cumulative histograms also permits to be robust to noise. Comparisons with local histogram data terms and non-convex approaches also demonstrate the capacity of our model to deal with complex configurations.

Even if the convex reformulation changes the nature of the problem by normalizing the data terms with respect to their prior area, the value of the area ratio parameter β is not an issue. Namely, it has been set to 0.5 in almost all of our experiments, which shows that our model is robust in practice.

Future works will be to study the use of other descriptors for texture. We will also be interested in investigating the connections between minimizers of the original problem and the ones of the relaxed problem for different norms in order to make the best possible choice when finally thresholding the estimated solution.

References

1. Aubert, G., Barlaud, M., Faugeras, O., Jehan-Besson, S.: Image segmentation using active contours: Calculus of variations or shape gradients? SIAM Applied Mathematics 63, 2003 (2002)
2. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing. Applied Mathematical Sciences, vol. 147. Springer (2002)
3. Aujol, J.-F., Aubert, G., Blanc-Féraud, L.: Wavelet-based level set evolution for classification of textured images. IEEE Transactions on Image Processing 12, 1634–1641 (2003)
4. Ayed, I., Chen, H., Punithakumar, K., Ross, I., Li, S.: Graph cut segmentation with a global constraint: Recovering region distribution via a bound of the bhattacharyya measure. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3288–3295. IEEE (2010)

5. Brown, E., Chan, T.F., Bresson, X.: Completely Convex Formulation of the Chan-Vese Image Segmentation Model. In: IJCV, pp. 1–19 (2011)
6. Brox, T., Rousson, M., Deriche, R., Weickert, J.: Unsupervised segmentation incorporating colour, texture, and motion. In: Petkov, N., Westenberg, M.A. (eds.) CAIP 2003. LNCS, vol. 2756, pp. 353–360. Springer, Heidelberg (2003)
7. Brox, T., Weickert, J.: Level set segmentation with multiple regions. IEEE Transactions on Image Processing 15(10), 3213–3218 (2006)
8. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. JMIV 40, 120–145 (2011)
9. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Transactions on Image Processing 10(2), 266–277 (2001)
10. Collins, M.D., Xu, J., Grady, L., Singh, V.: Random walks based multi-image segmentation: Quasiconvexity results and gpu-based solutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), pp. 1656–1663. IEEE (2012)
11. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. International Journal of Computer Vision 72, 215 (2007)
12. Delfour, M.C., Zolésio, J.-P.: Shapes and geometries: analysis, differential calculus, and optimization. Society for Industrial and Applied Mathematics, Philadelphia (2001)
13. Gorelick, L., Schmidt, F.R., Boykov, Y., Delong, A., Ward, A.: Segmentation with non-linear regional constraints via line-search cuts. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part I. LNCS, vol. ECCV 2012, pp. 583–597. Springer, Heidelberg (2012)
14. Herbulot, A., Jehan-Besson, S., Duffner, S., Barlaud, M., Aubert, G.: Segmentation of vectorial image features using shape gradients and information measures. Journal of Mathematical Imaging and Vision 25(3), 365–386 (2006)
15. Kim, J., Fisher, J.W., Yezzi, A., Cetin, M., Willsky, A.S.: A nonparametric statistical method for image segmentation using information theory and curve evolution. IEEE Transactions on Image Processing 14, 1486–1502 (2005)
16. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. 42, 577–685 (1989)
17. Ni, K., Bresson, X., Chan, T.F., Esedoglu, S.: Local histogram based segmentation using the wasserstein distance. International Journal of Computer Vision 84(1), 97–111 (2009)
18. Nikolova, M., Esedoglu, S., Chan, T.F.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM Journal on Applied Mathematics 66(5), 1632–1648 (2006)
19. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. JCP 79(1), 12–49 (1988)
20. Paragios, N., Deriche, R.: Geodesic active regions for supervised texture segmentation. In: ICCV (1999)
21. Peyré, G., Fadili, J., Rabin, J.: Wasserstein active contours. In: IEEE International Conference on Image Processing (ICIP 2012) (2012)
22. Pock, T., Chambolle, A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In: IEEE International Conference on Computer Vision (ICCV 2011), pp. 1762–1769 (2011)
23. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: Global solutions of variational models with convex regularization. SIAM JIS 3, 1122–1145 (2010)

24. Punithakumar, K., Yuan, J., Ben Ayed, I., Li, S., Boykov, Y.: A convex max-flow approach to distribution-based figure-ground separation. SIAM Journal on Imaging Sciences 5(4), 1333–1354 (2012)
25. Rother, C., Minka, T., Blake, A., Kolmogorov, V.: Cosegmentation of image pairs by histogram matching-incorporating a global constraint into mrfs. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2006), vol. 1, pp. 993–1000. IEEE (2006)
26. Rousson, M., Brox, T., Deriche, R.: Active unsupervised texture segmentation on a diffusion based feature space. In: CVPR (2003)
27. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the mumford and shah model. International Journal of Computer Vision 50, 271–293 (2002)
28. Vicente, S., Kolmogorov, V., Rother, C.: Joint optimization of segmentation and appearance models. In: IEEE International Conference on Computer Vision (ICCV 2009), pp. 755–762. IEEE (2009)
29. Yildizoglu, R., Aujol, J.-F., Papadakis, N.: Active contours without level sets. In: IEEE International Conference on Image Processing (ICIP 2012) (2012)
30. Yuan, Y., Ukwatta, E., Tai, X.C., Fenster, A., Schnörr, C.: A fast global optimization-based approach to evolving contours with generic shape prior. In: submission in IEEE TPAMI, also UCLA Tech. Report CAM 12-38 (2012)
31. Zhu, S.C., Lee, T.S., Yuille, A.L.: Region competition: unifying snakes, region growing, energy/bayes/mdl for multi-band image segmentation. In: IEEE International Conference on Computer Vision (ICCV 1995), pp. 416 –423 (June 1995)