

Quadratic Markovian Probability Fields for Image Binary Segmentation

Mariano Rivera and Pedro P. Mayorga
Centro de Investigacion en Matematicas A.C.
Apdo Postal 402, Guanajuato, Gto. 36000 Mexico
mrivera@cimat.mx

Abstract

In this work we present a new Markov Random Field model for image binary segmentation that computes the probability that each pixel belongs to a given class. We show that if a real valued field is computed, instead of a binary one as in graph cuts based methods, then the resultant cost function has noticeable computational and performance advantages. The proposed energy function can be efficiently minimized with standard fast linear order algorithms as Conjugate Gradient or multigrid Gauss-Seidel schemes. Moreover, our formulation accepts a good initial guess (starting point) and avoids to construct from scratch the new solution accelerating the computational process. Then we naturally implement computationally efficient multigrid algorithms. For applications with limited computational time, a good partial solution can be obtained by stopping the iterations even if the global optimum is not yet reached. We performed a meticulous comparison (with state of the art methods: Graph Cut, Random Walker and GMMF) for the interactive image segmentation (based on trimaps). We compare the algorithms using cross-validation procedures and a simplex decent algorithm for learning the parameter set.

1. Introduction

Two-classes image segmentation [image binary segmentation (IBS)] is an important issue in image analysis and image editing tasks. There are many problems which the core solution algorithm is an IBS method; for instance: interactive image segmentation (trimap) [23], [5], [2], [22], [14] organs segmentation in medical imaging (e.g. skull stripping) [4], [10], foreground extraction (image matting) [22], [25], motion computation [7], [15], among others. Multiclass image segmentation is also commonly implemented by the successive applications of IBS methods [3], [7], [15]. Last listed applications shown that any improvement to IBS methods in the convergence ratio, reduction in memory requirements or error reduction will have an important im-

pact in many image processing and computer vision applications. In this paper we present a novel IBS method that improves the procedures of the state of the art with respect to the three above listed issues. Our method is based on a new Markov Random Field model and computes the probability of each pixel belongs to a given class. It is based on the minimization of a quadratic energy function; such a minimization corresponds to solve a linear system with standard iterative algorithms as Gauss-Seidel (GS) or Conjugate Gradient (CG) [20]. As it is well known, the convergence ratio of such algorithms can be improved by providing a good initial guess (starting point). Moreover gradient descent based algorithms (as GS or CG) produce a partial solutions sequence (a new partial solution at each iteration) that reduces successively the energy function. Thus, for applications with limited computational time, a good partial solution can be obtained by stopping the iterations even if the global optimum is not yet reached. These characteristics lead us to, naturally, implement computationally efficient multigrid algorithms [8].

We organize this paper as follows. In section 2 we presents a new derivation of the recently reported method by Rivera *et al.* [21] for soft (probabilistic) multiclass image segmentation. Our derivation is more accord with the Bayesian regularization framework. Such an algorithm is proposed as the minimization of a linearly constrained quadratic, does not necessarily positive definite, energy function. In section 3 we particularize the method in derived in section 2 for the IBS case. In our formulation the probabilities are represented by a single Markov Random Field (MRF) and our positive definite quadratic energy function incorporates effectively the constraints. In section 4 we present a discussion about related formulation for multilabel image segmentation. In section 5 we evaluate the performance of the proposed IBS method in the interactive color IBS task based on trimaps. For such purposes we follow the implementation by Boykov and Jolly [5] and then we only replace the IBS method. We used the popular Lasso's bench database, such database is used in [2] and available online [28]. The experimental



Figure 1. Image model generation. I_1 and I_2 are the original data, α is matting factor and g is the observed image.

results demonstrate a superior performance of our method compared with methods of the state of the art for IBS. We conduct a deeper evaluation based on a *cross-validation* technique was used for such purposes and the algorithms hyper-parameters were automatically adjusted. Then in section 6, we demonstrate (by numerical experiments in both real and synthetic data) the method capability for the simultaneous estimation task of the segmentation and the model parameters. Finally, section 7 presents our conclusions.

2. Quadratic Markov Measure Fields Models

Recently Rivera *et al.* [21] proposed the Entropy Controlled Gauss Markov Measure Fields (EC-GMMF) models for image multiclass segmentation. Such an algorithm is computationally efficient and produces “soft” segmentations of excellent quality. In [21] the algorithm is derived in the framework of Gauss Markov Measure Fields (GMMF) models [18], an early work of the EC-GMMF’s authors. In the GMMF framework the likelihood is constrained to be Gaussian and therefore quadratic in the measure fields [18]. However, the exact form of the quadratic potentials is not constrained and thus in [21] is chosen an “appropriated potential.” In this section we present an alternative derivation to the formulation in [21] rigorously based on the framework of Bayesian regularization (BR) with prior Markov random fields (MRF) models. We directly derive the formulation from an observation model instead of a simple conjecture. MRF models is a well-accepted and powerful approach for solving problems in early computer vision and image processing [1], [2], [3], [5], [7], [9], [14], [15], [16], [18], [21], [22], [23], [25].

Consider the example in Fig. 1 where the image g is generated by using the model

$$g(x) = \alpha(x)I_1(x) + (1 - \alpha(x))I_2(x), \quad (1)$$

where $x \in \mathcal{R} \subseteq \mathcal{L}$ denotes a pixel position in the region of interest, \mathcal{R} , into the regular lattice \mathcal{L} ; I_1 and I_2 are two general images and α is a matting factor [22], [25]. We can generalize the model (1) to the case of multiple regions as:

$$g(x) = \sum_k \alpha_k(x)I_k(x), \quad (2)$$

for $k = 1, 2, \dots, K$; where

$$\alpha_k(x)g(x) = \alpha_k(x)(I_k(x) + \eta_k(x)); \quad (3)$$

where η_k is a noise image with known distribution and the matting factors α ’s satisfy:

$$\sum_{k=1}^K \alpha_k(x) = 1, \quad x \in \mathcal{R}; \quad (4)$$

$$\alpha_k(x) \geq 0, \quad k = 1, \dots, K, x \in \mathcal{R}; \quad (5)$$

$$\alpha_i(x)\alpha_j(x) \approx 0 \quad \text{if } i \neq j, \quad (6)$$

$$\alpha(x) \approx \alpha(y), \quad x \in \mathcal{R}, y \in \mathcal{N}_x; \quad (7)$$

where \mathcal{N}_x denotes the set of first neighbors of x : $\mathcal{N}_x = \{y \in \mathcal{R} : |x - y| = 1\}$. Note that, because (4) and (5), α can be interpreted as a probability measure field where $\alpha_k(x)$ is understood as the probability of the pixel $I_k(x)$ is likely to produce the observation $g(x)$. Additionally (6) introduces the constraint on the probability vectors $\alpha(x)$ to have low entropy: together with (4) and (5), constraint (6) indicates that only one $\alpha(x)$ vector entry has a value close to one and the others are close to zero. The constraint (7) promotes the probability measure α to be spatially smooth.

The segmentation of the composed image g can be seen as the solution to ill-posed inverse problem stated in (2) and (3) subject to the *hard* constraints (4) and (5) and the *soft* constraints (6) and (7). This is, the computation of the matting factors α_k and the original images I_k , or at least the image fractions $\alpha_k I_k$. In the BR framework, with MRF model priors, one computes the solution (α^*, I^*) as an estimator of the posterior distribution $P(\alpha, I|g)$. Then by using the Bayes rule, the posterior distribution can be expressed as:

$$P(\alpha, I|g) = \frac{1}{Z} P(g|\alpha, I)P(\alpha, I). \quad (8)$$

where $P(g|\alpha, I)$ is conditional probability of the data by assuming given unknowns (α, I) , $P(\alpha, I)$ is the prior distribution and $Z = P(g)$ is a normalization constant (independent on (α, I)). In this framework, the conditional probability $P(g|\alpha, I)$ is derived from the noise distribution and the observation model [(2) and (3)]; the prior $P(\alpha, I)$ expresses that the parameters are Markovian.

The general inference of the images I_k from the data g is a complex inverse problem: even if α is given, we could recover the fraction $\alpha_k I_k$ of the whole image. Thus important assumptions (priors) need be used: for instance, we can consider that such images can be represented by a parametric function: $I_k(x) = \Phi(x, \beta)$, with parameters β . For simplify the notation we express our derivation in terms of I instead of the parameters β . Parametric forms, although limited, has been used and have shown be useful for defining layered models for segmenting gray scale images or optical flows [17], [24]. Other form for introducing prior (expert) knowledge is by interactively leading the segmentation. In this case the user label a subset of pixels and the unknown labels are estimated with a segmentation algorithm.

For deriving $P(g|\alpha, I)$ we start by assuming that the η_k is i.i.d. Gaussian noise with mean zero and variance σ_k^2 , i.e.:

$$P(\eta_k(x)) = G_{\sigma_k}(\eta_k(x)) \quad (9)$$

where we define

$$G_{\sigma}(z) \stackrel{def}{=} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{z^2}{2\sigma^2}\right]. \quad (10)$$

From (3) we have: $\alpha_k(x)\eta_k(x) = \alpha_k(x)(g(x) - I_k(x))$. As α_k is almost binary [because (6)], then for $\alpha_k(x) \approx 1$ one can expect a similar distribution for both $\alpha_k(x)\eta_k(x)$ and $\eta_k(x)$. Therefore (by defining $r_k(x) = g(x) - I_k(x)$):

$$\begin{aligned} P(\alpha_k g | \alpha_k I_k, \sigma_k^2) &= \prod_x G_{\sigma_k}(\alpha_k(x) r_k(x)) \\ &= \prod_x G_{\sigma_k}(r_k(x))^{\alpha_k^2(x)}. \end{aligned} \quad (11)$$

Now we remove the assumption of Gaussian noise. First we consider that any smooth density distribution v_k can be expressed with a Gaussian mixture model [13]:

$$v_k(x, \theta_k) = \sum_{i=1}^I \pi_{ki} G_{\sigma_k}(r_k(x) - m_{ki}), \quad (12)$$

with $\theta_k = (\sigma_k, \pi_k, m_k)$; where $\pi_{ki} \geq 0$ are the mixture coefficients (with $\sum_i \pi_{ki} = 1$), the Gaussians centers $m_k = (m_{k1}, m_{k2}, \dots, m_{kI})$ and variances σ_k are known (this is equivalent to assume a known noise distribution) and I is the number (maybe large) of Gaussians in the mixture. Then in the low entropy limit we have:

$$\begin{aligned} P(\alpha_k g | \alpha_k I_k, \theta_k) &= \prod_x \left[\sum_{i=1}^I \pi_{ki} G_{\sigma_k}(r_k(x) - m_{ki})^{\alpha_k^2(x)} \right] \\ &\approx \prod_x \left[\sum_{i=1}^I \pi_{ki} G_{\sigma_k}(r_k(x) - m_{ki}) \right]^{\alpha_k^2(x)} \\ &= \prod_x v_k(x, \theta_k)^{\alpha_k^2(x)}. \end{aligned} \quad (13)$$

If independency between $\alpha_i I_i$ and $\alpha_j I_j$ (for $i \neq j$) is assumed, then the likelihood of the observed (composed) image g is given by

$$P(g|\alpha, I, \theta) = \prod_k P(\alpha_k g | \alpha_k I_k, \theta_k). \quad (14)$$

In particular, such an independency occurs if (6) is satisfied. The entropy of discrete distributions can be reduced by increasing its variance: $E^2[\alpha_k] - E[\alpha_k^2]$. That leads us to introduce the potential $\mu (1 - \sum_k \alpha_k^2(x))$ with $\mu > 0$, also named Gini's coefficient in [13]. Additionally a Gibbsian distribution based on MRF models controls the granularity of the regions (i.e. for promoting smooth regions). We finally obtain the prior $P(\alpha)$:

$$\frac{1}{Z'} \exp \left[\sum_k \sum_{x \in \mathcal{R}} \left(\mu \alpha_k^2(x) - \lambda \sum_{y \in \mathcal{N}_x} |\alpha(x) - \alpha(y)|^2 \right) \right]; \quad (15)$$

where Z' is a normalization constant. Therefore, we use $P(\alpha, I) \propto P(\alpha)$ as prior probability by assuming independence among I and α , and an uniform distribution on I . Thus the posterior distribution takes the form $P(\alpha, \theta | g) \propto \exp[-U(\alpha, \theta)]$ and the maximum a posteriori (MAP) estimator is computed by minimizing the cost function:

$$\begin{aligned} U(\alpha, \theta) &= \sum_{x \in \mathcal{R}} \left\{ \sum_{k=1}^K \alpha_k^2(x) [-\log v_k(x, \theta) - \mu] \right. \\ &\quad \left. + \frac{\lambda}{2} \sum_{y \in \mathcal{N}_x} |\alpha(x) - \alpha(y)|^2 \right\}, \end{aligned} \quad (16)$$

subject to the constraints (4) and (5).

This quadratic programming problem can efficiently be solved by using the Lagrange multiplier procedure for the equality constraint (4) and a projection strategy for the non-negativity constraint (5). The convergence is guaranteed to a local minima. See [21] for details.

3. Quadratic Markov Probability Fields

For the particular case of IBS [i.e. for the case in model (1)], the resultant energy function has remarkable computational and performance advantages over standard IBS methods. Let the normalized likelihoods corresponding to the first and second classes:

$$\hat{v}_k(x, \theta) = \frac{v_1(x, \theta_1)}{s(x, \theta)}, \quad (17)$$

for $k = 1, 2$; with

$$s(x, \theta) \stackrel{def}{=} \sum_k v_k(x, \theta_k). \quad (18)$$

Thus we define the distances:

$$d_k(x) \stackrel{def}{=} -\log \hat{v}_k(x, \theta_k). \quad (19)$$

That leads us to formulate our IBS method as the minimization of the unconstrained quadratic cost function:

$$Q(\alpha) = \sum_{x \in \mathcal{R}} \left\{ \alpha^2(x) [d_1(x) - \mu] + (1 - \alpha(x))^2 [d_2(x) - \mu] + \lambda \sum_{y \in \mathcal{N}_x} (\alpha(x) - \alpha(y))^2 \right\}. \quad (20)$$

The minimization convergence properties of (20) are established in the following theorem.

Theorem 1. *Binary QMPF convergence.*

- (i) If $\mu < \min_{k,x} d_k(x)$, then (20) has unique global minimum that satisfies $\alpha(x) \geq 0, \forall x \in \mathcal{R}$.
- (ii) Otherwise an energy descend algorithm convergence, at least, to a local minima if the additional constraint $\alpha(x) \geq 0, \forall x \in L$, is enforced.

Proof of (i). Assuming $\mu < \min_{k,x} d_k(x)$.

- (a) $Q(\alpha)$ a convex quadratic potential with a unique global minima. Thus the linear system that results of equaling to zero the gradient of (20) w.r.t. α can be solved with the Gauss-Seidel (GS) scheme

$$\alpha(x) = \frac{a(x)}{b(x)} \quad (21)$$

with

$$a(x) \stackrel{def}{=} d_2(x) - \mu + \lambda \sum_{y \in \mathcal{N}_x} \alpha(y), \quad (22)$$

$$b(x) \stackrel{def}{=} d_1(x) + d_2(x) - 2\mu + \lambda \#\mathcal{N}_x; \quad (23)$$

where $\#\mathcal{N}_x$ denotes the cardinality of \mathcal{N}_x .

- (b) If $\alpha^0(x) \in [0, 1], \forall x$, is provided as initial guess then the sequence generated by the GS scheme (21) satisfies $\{\alpha^t(x)\}_{t=1, \dots, T} \in [0, 1]$, for any iteration number t (given that $b(x) \geq a(x) \geq 0$). Therefore the unique global minimizer is also in the interval $[0, 1]$.

Finally, from (a) and (b), any minimization algorithm converge to the unique global minima, $\alpha^*(x) \in [0, 1]$, independently of the initial point α^0 .

Proof of (ii). It follows from the fact that any descent algorithm that produces a feasible sequence, $\{\alpha^t\}_{t=1, \dots, T}$, for solving an indefinite quadratic (linearly constrained) problem converge to, at least, a local minima, [20]. \square

The formulation of the IBS problem as the minimization of an unconstrained positive definite quadratic energy function

has the advantage of being achieved by computational efficient algorithms, as CG or a multigrid implementation of the GS scheme in (21). Although an initial guess does not determine the convergence to the global minima, a good starting point can accelerate the convergence rate. For instance, we initialize $\alpha(x) = \hat{v}_1(x, \theta_1)$ in this work. Moreover, descend algorithms produce sequences $\{\alpha^t\}_{t=1, \dots, T}$ such that:

$$Q(\alpha^0) \geq \dots \geq Q(\alpha^i) \geq Q(\alpha^{i+1}) \geq \dots \geq Q(\alpha^*) \geq -\frac{1}{2}; \quad (24)$$

where the superscripts i and $i + 1$ indicate consecutive iteration numbers. GS can be seen as a particular case of a coordinate descent that converges if $b(x^t) \geq a(x^t) \geq 0, \forall t$. These features (the possibilities of to provide initial guesses and to stop the algorithm iterations for having partial solutions) allows us to naturally implement fast multigrid algorithms.

4. Relationship with other Markov Measure Fields Models

A Markov measure field (MMF), α , is a random vectorial field that satisfies (4) and (5) with a Gibbsian prior distribution $P(\alpha)$ in terms of MRF models [16]. In this comparative study we focus in segmentation algorithms that are implemented as the computation of the MAP estimator, i.e. by minimizing a posterior energy of the general form:

$$MMF(\alpha) = D(\alpha, g) + \lambda R(\alpha). \quad (25)$$

The potential D corresponds to the negative log-likelihood of the data given the labels and is determined by the observation model and the noise distribution. The potential R is the negative log-prior, also known as regularization term. We will focus our discussion in variants for the potential D :

Binary MMF. The image segmentation task is a combinatorial problem in its original formulation: to assign to each pixel the label of the class that is more likely to belong. The prior knowledge of the labels field smoothness is introduced in form of an MRF [9] [1]. Among the most computationally efficient algorithms for solving such a problem is the graph-cut (GC) algorithm that essentially computes the MAP estimator from a posterior distribution as (8) [7], [6]. The GC algorithm guarantees convergence to a global minima for the IBS case.

Gaussian MMF (GMMF) [18]. Such a work proposed to compute the probability that a pixel can be generated by a particular intensity model instead of directly the label map, i.e. differently to hard segmentation schemes based on graph cuts methods. The method models (with a MRF) directly the posterior marginals field, α , of the hard-labels.

The GMMF potential,

$$\sum_k \sum_x (\alpha_k(x) - \hat{v}_k(x, \theta_k))^2 + \lambda R(\alpha), \quad (26)$$

is chosen such that, for $\lambda = 0$, the posterior marginals are equal to the likelihoods, *i.e.* the *consistence condition*:

$$\alpha_k(x) = \hat{v}_k(x, \theta_k) \stackrel{\text{def}}{=} \frac{v_k(x, \theta_k)}{s(x)}. \quad (27)$$

is satisfied.

Random Walker (RW) [11], [12]. Although introduced in terms of random walks of particles, RW is a variant of the GMMF formulation (see the diffusion process in [18]). The *consistence condition* is reformulated as:

$$s(x)\alpha_k(x) = v_k(x, \theta_k). \quad (28)$$

Then the corresponding potential is a quadratic one such that the minimum for $\lambda = 0$ results in (28) and consequently satisfies the GMMF *consistence condition* (27). The image coloring procedure proposed Levin *et al.* is close related with the GMMF diffusion process with space-varying weights [18].

Quadratic MMF [This work]. Differently to GMMF models, the minimum of the QMPF potential (4) for the case of $\lambda = 0$ corresponds to:

$$\alpha_k(x) = \frac{1}{K} \frac{H(d(x))}{d_k(x)}; \quad (29)$$

where $H(z)$ is the *harmonic mean* of z . Because the GMMF-*consistence condition* is not satisfied by , it does not corresponds to a GMMF model.

5. Image Binary Interactive Segmentation

In this section we compare the performance of the proposed probabilistic method (based on QMPF models) with of popular IBS segmentations methods: maximum flow (minimum graph cut), GMMF and Random Walker. The task is the binary interactive segmentation of color images (segmentation by trimaps). A *cross-validation* procedure was implemented for comparing the generalization capabilities of the methods [13]. The benchmark data is the set of 50 images in the Lasso’s database used in [2], available online in [28]. Such a database contains a natural images set with the corresponding trimap and the ground truth. Actually, a Lasso’s trimap is image of class labels: no-process mask (\mathcal{M}), definitively background (\mathcal{B}), unknown (\mathcal{R}) and definitively foreground (\mathcal{F}). Note that each pixel $x \in \mathcal{L}$ has an unique label. First column in Fig. 2 shows images in the Lasso’s database and second column the corresponding trimaps (the gray scale corresponds with above class enumeration). In this case, the region to process is labelled as

“unknown” and the boundary condition are imposed by the foreground and background labelled regions. The regularization term in (20) is replaced by:

$$\lambda \sum_{y \in \mathcal{N}_x \cap \mathcal{B} \cap \mathcal{F}} [\alpha(x) - \alpha(y)]^2 l_{xy}, \quad (30)$$

for leading the border regions to follow the color edges; where the color border information is computed with

$$l_{xy} = \frac{\gamma}{\gamma + |g(x) - g(y)|^2}; \quad (31)$$

and the color image, g , is previously transformed to the CIE-Lab color space with the Ruzon’s C-code library [27] and γ is a method’s hyper-parameter that controls the edge sensibility. l_{xy} is an affinity measure that takes a value close to one if the neighbor pixels x and y have similar colors and its is close to zero for too different colors. Recent reported matting computation methods have focused in variants of the intra-pixel affinity measure with improved results w.r.t. the basic one in (31) [22] [10]. However, in our experimental, we use the simple affinity measure (31) for comparing directly the performance of the methods.

In this task, empirical likelihoods are computed from the histogram of the labelled by hand pixels. Following [5], the empirical likelihoods are computed from the smoothed (with 10 iterations of a homogeneous diffusion filter) color histograms of the foreground, h_1 , and background, h_2 , labelled pixels. Then the normalized likelihoods are computed with:

$$\hat{v}_k(x) = \frac{h_k(g(x)) + \epsilon}{h_1(g(x)) + h_2(g(x)) + 2\epsilon}, \quad (32)$$

for $k = 1, 2$; where $\epsilon = 10^{-3}$ is a small positive constant that introduces a contaminant uniform distribution that stabilizes the likelihoods and it avoids the undefined computation of $\log 0$. We initialize α with (29) and set $d_k(x) = -\log \hat{v}_k(x)$. The normalized histograms can be seen as 3D Look-Up-Table with $50 \times 100 \times 100$ dimensions for the L , a and b coordinates, respectively.

The data set was partitioned in 5 groups of 10 images. The parameters set were trained by minimizing the mean of the segmentation error (computed according to [2]) in

	Params.	AIC	Training	Testing
Graph cut	λ, γ	8.58	6.89%	6.93%
Rand. Walk.	λ, γ	6.50	5.46%	5.50%
GMMF	λ, γ	6.49	5.46%	5.49%
QMPF	λ, γ	6.04	5.02%	5.15%
QMPF+EC	λ, γ, μ	5.39	3.13%	3.13%

Table 1. Cross-validation results. Parameters, Akaike information criterion, training and testing error.

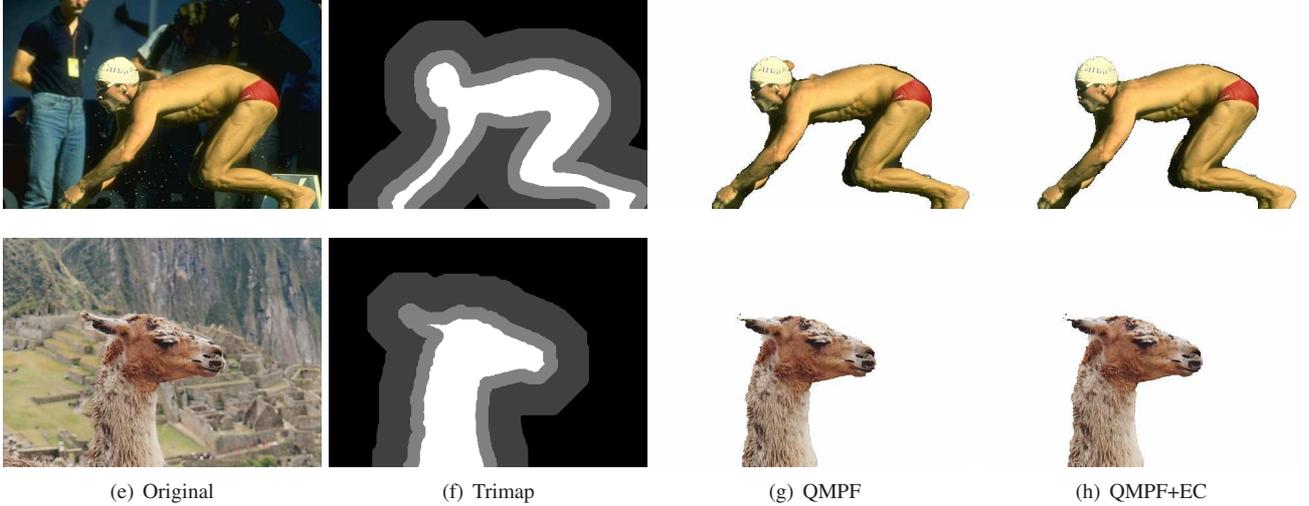


Figure 2. Selected trimaps segmentations.

groups of 40 images by using the Nelder and Mead simplex descent [19]. Table 1 shows the training and testing error averages. Figure 2 shows examples of the segmented images. Additionally, the Akaike information criterion (AIC) is computed for the parameters optimized (trained) with the 50 image in the database [13], such an AIC estimation of the prediction error is consistent with the cross-validation results. Note that the QMPF algorithm has the best performance in the group. For our implementation, the learned parameters were: $(\lambda = 4.7 \times 10^3, \gamma = 9 \times 10^{-6})$ and $(\lambda = 3.8 \times 10^{-4}, \gamma = 1.3 \times 10^{-6}, \mu = -123)$ for QMPF and QMPF+EC, respectively. We note that such parameters are appropriated for the trimap segmentation task and should not produce the expected results in other tasks as, for instance, matting computation. For instance we note that the learned parameter μ for QMPF+EC promotes large entropy. For illustrate last statement, we compute the matting factors for the example illustrated in Fig. 1. The results are shown in Fig. 3. In particular the matting factor shown Fig. 1 was computed with QMPF with $\mu = 0$.

6. Model parameters estimation

The method has a noticeable advantage if Gaussian likelihoods are assumed and the parameters, $\theta_k = [m_k, \sigma_k]$ (mean and standard deviation, respectively), are unknown. Then the parameters can be efficiently estimated by using an alternated minimization scheme of the cost function (20) w.r.t. the MMF, α , and the parameters, θ . For illustrating this capability, we consider the task of computing a binarization of a synthetic image (Fig. 4a) generated with model (1); where I_k are constant values for all x (actually white and black in gray values), and $\eta_k(x) \sim \mathcal{N}(0, \sigma_k^2)$ (i.i.d. Gaussian noise. Such a segmentation task (i.e. the

estimation of the indicator variables $b_k(x) \in \{0, 1\}$ with $\alpha_k(x) \approx b_k(x)$) requires of the simultaneously estimation of α and $\theta_k = [I_k, \sigma_k^2]^T$ for $k = 1, 2$. As $\eta_k(x)$ is Gaussian, we have:

$$-\log v_k(x, \theta) = \frac{1}{2\sigma_k^2} |g(x) - I_k|^2 + \log \sqrt{2\pi} \sigma_k. \quad (33)$$

Then, by assuming an uniform distribution as prior for θ , from the partial derivatives w.r.t. the parameters, we have:

$$I_k = \frac{\sum_x \alpha^2(x) g(x)}{\sum_x \alpha^2(x)} \quad (34)$$

and

$$\sigma_k^2 = \frac{\sum_x \alpha^2(x) |g(x) - I_k|^2}{\sum_x \alpha^2(x)}. \quad (35)$$

Such formulas, (34) and (35), are similar to the ones obtained in an Expectation-Maximization (EM) procedure; except by the $\alpha^2(x)$ weighting factor instead of $\alpha(x)$. Such a factor is also changed for estimating the covariance matrix of multivariated Gaussian models.

Figure 4 shows the pair of images used this experiment. The synthetic binary image, in Fig. 4(a), was precluded with Gaussian noise with zero mean and $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$ for the white and black regions, respectively. Fig 4(b) shows a metallic real piece illuminated with laser (coherent) light and thus corrupted with speckle (multiplicative) noise. The effect of the entropy control parameter, μ , is showed in Fig. 5. The computed α field with $\mu = 0$ (without entropy control) and the corresponding binarization are shown in Figs. 5(a) and 5(b). Figs. 5(c) and 5(d) show the results computed with $\mu = 0.5$. Table 2 summarizes the experiment results. We noted that, for the IBS case, the results (segmentation and the estimated parameters) are

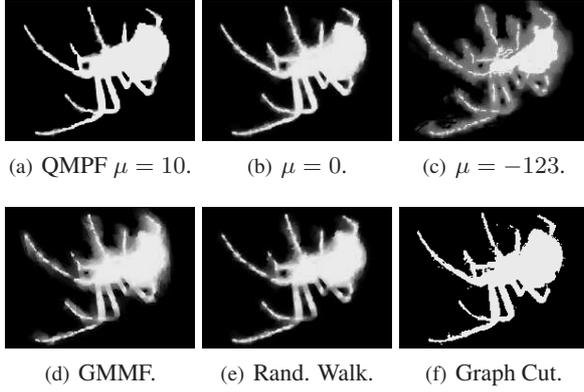


Figure 3. First row, results computed with the proposed method with a) low-entropy, b) without entropy control and c) high entropy. Second row, results computed with methods of the state of the art.

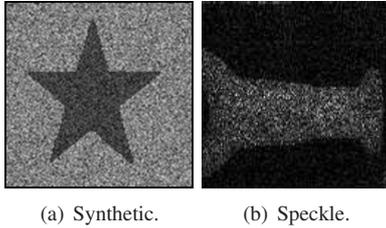


Figure 4. Test Images.

robust to the exact value of the entropy control. The models (I_1 and I_2) were initialized with the maximum and minimum image gray values, respectively.

	I_1	σ_1	I_2	σ_2
Real values	1.000	0.500	0.000	0.300
Initial condition	2.760	0.100	-1.013	0.100
$\lambda = 4, \mu = 0.0$	1.002	0.477	0.005	0.325
$\lambda = 4, \mu = 0.3$	0.999	0.488	0.002	0.308

Table 2. Computed parameters for Fig. 5.

Fig. 6 shows the results corresponding to the speckle image. In Fig. 6(a) we show the computed α field with the proposed QMPF method (with $\mu = 0$ and $\lambda = 1 \times 10^3$) and Fig. 6(b) the corresponding segmentation. Second row shows the computed results with GMMF. The computed α field with the GMMF algorithm has, evidently, larger entropy than the QMPF solution. This is consistent with the results reported in Refs. [17] and [21]. If such a high-entropy α field were used in an EM kind scheme for estimating the model parameters then the algorithm may converge to a single value. Such a limitation of the GMMF model is discussed in [17]. As it is expected, we observed a similar behavior for the Random Walker algorithm than for GMMF.

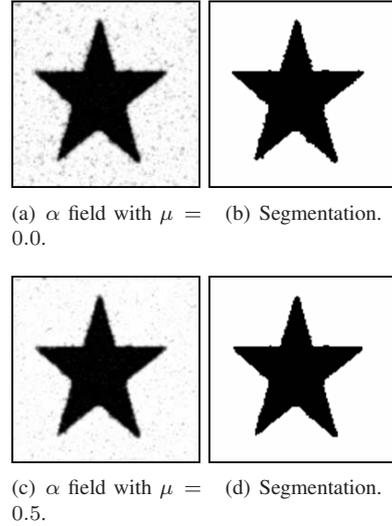


Figure 5. Entropy control.

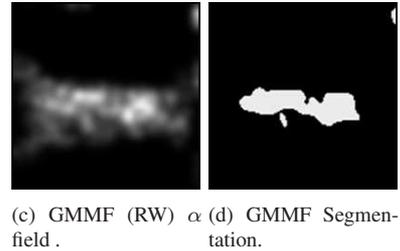
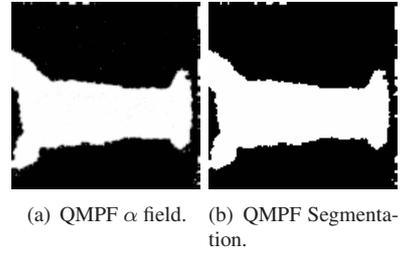


Figure 6. Effect of the data term.

7. Conclusions

We started our paper by presenting a new derivation of the Markovian models for multi-class image segmentation presented by Rivera *et al.* [21]. Our derivation is accord with the Bayesian Regularization framework. We have named such a models Quadratic Markovian Measure Fields (QMMF) and have exposed the relationship (and difference) with the Gauss Markov random fields (GMMF) and/or Random Walker. The algorithm computes a low entropy (almost binaries) and regularized (smooth) vector field, α . Such that $\alpha_k(x)$ that can be interpreted as the probability of the pixel x were generated with the distribution k . We have shown that the QMMF models accept any probability distribution as the empirical distributions estimated

by histogram techniques or kernel methods.

We have presented a new quadratic energy function for IBS. For evaluating the proposed model performance, we implemented an interactive binary segmentation tool (segmentation by trimaps) and compare the results by substituting our algorithm with state of the art methods: Graph Cut, Random Walker and GMMF. The remaining implementations details were unaltered. As bench data set we used the Lasso's trimap set of 50 natural images.

We have achieved a meticulous comparison of the algorithm by using a cross-validation procedures and a simplex decent algorithm for learning the parameter set. Such a comparison showed that our proposal have a superior performance than the compared methods and illustrate the importance of the entropy control introduced by Rivera *et al.* [21]. According with our experiments the interactive IBS task is better achieved with high entropy probabilities, however, the matting computation (as the simultaneous estimation of the segmentation and parameter) requires of low-entropy fields.

In the interactive IBS task is common that once a solution is computed then the user refine such a solution by retouching the initial trimap. Our method can use as initial guess for a subsequent refining the previous final solution (a feasible point for the next problem). That accelerate the interactive process by avoiding to construct from scratch the new solution.

Future work consider to extend our trimap based comparison to other IBS methods as, for instance, the based on the Maximum of the Posterior Marginal (MPM) estimator. Such estimator is now possible to estimate by fast methods based on graph cuts [26].

References

- [1] J. Besag. On the statistical analysis of dirty pictures. *J. R. Stat. Soc., Ser. B, Methodol.*, 48:259–302, 1986.
- [2] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr. Interactive image segmentation using an adaptive gmmrf model. In *ECCV*, volume 1, pages 414–427, 2004.
- [3] C. A. Bouman and M. Shapiro. A multiscale random field model for bayesian image segmentation. *IEEE Trans. Image Processing*, 3(2):162–177, 1994.
- [4] Y. Boykov and M.-P. Jolly. Interactive organ segmentation using graph cuts. In *MICCAI, LNCS 1935*, pages 276–286, 2000.
- [5] Y. Boykov and M.-P. Jolly. Interactive graph cut for optimal boundary & region segmentation of objects in N-D images. In *ICIP (1)*, pages 105–112, 2001.
- [6] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Int'l J. Computer Vision*, 70(2):109–131, 2006.
- [7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE PAMI*, 23(11):1222–1239, 2001.
- [8] W. L. Briggs, S. McCormick, and V. Henson. *A Multigrid Tutorial*. SIAM Publications, second edition, 2000.
- [9] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE PAMI*, 6:721–741, 1984.
- [10] L. Grady, T. Schiwietz, S. Aharon, and R. Westermann. Random Walks for interactive organ segmentation in two and three dimensions: Implementation and validation. In *MICCAI (2), LNCS 3750*, pages 773–780, 2005.
- [11] L. Grady, Y. Sun, and J. Williams. Interactive graph-based segmentation methods in cardiovascular imaging. In N. P. et al., editor, *Handbook of Mathematical Models in Computer Vision*, pages 453–469. Springer, 2006.
- [12] L. Grady. Multilabel random walker image segmentation using prior models. In *CVPR*, volume 1 of *CVPR*, pages 763–770, 2005.
- [13] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2001.
- [14] O. Juan and R. Keriven. Trimap segmentation for fast and user-friendly alpha matting. In *VLSM, LNCS 3752*, pages 186–197, 2005.
- [15] V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, and C. Rother. Probabilistic fusion of stereo with color and contrast for bi-layer segmentation. *IEEE PAMI*, 28(9):1480–1492, 2006.
- [16] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, Tokyo, 2001.
- [17] J. L. Marroquin, S. Botello, F. Calderon, and B. C. Vemuri. MPM-MAP algorithm for image segmentation. In *ICPR*, 2000.
- [18] J. L. Marroquin, F. Velazco, M. Rivera, and M. Nakamura. Probabilistic solution of ill-posed problems in computational vision. *IEEE PAMI*, 23:337–348, 2001.
- [19] J. A. Nelder and R. Mead. A simplex method for function minimization. *Comput. J.*, 7:308–313, 1965.
- [20] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operation Research, 2000.
- [21] M. Rivera, O. Ocegueda, and J. L. Marroquin. Entropy controlled Gauss-Markov random measure fields for early vision. In *VLSM, LNCS 3752*, pages 137–148, 2005.
- [22] C. Rother, V. Kolmogorov, and A. Blake. Interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics*, number 23 (3), pages 309–314, 2004.
- [23] M. A. Ruzon and C. Tomasi. Alpha estimation in natural images. In *CVPR (1)*, pages 18–25, 2000.
- [24] The Authors.
- [25] J. Wang and M. Cohen. An interactive optimization approach for unified image segmentation and matting. In *ICCV*, number 3, pages 936–943, 2005.
- [26] Measuring uncertainty in graph cut solutions: Efficiently computing min-marginal energies using dynamic graph cuts. In *ECCV*, pages 30–43, 2006.
- [27] <http://ai.stanford.edu/~ruzon/software/rgblab.html>.
- [28] <http://research.microsoft.com/vision/cambridge/i3l/segmentation/GrabCut.htm>.