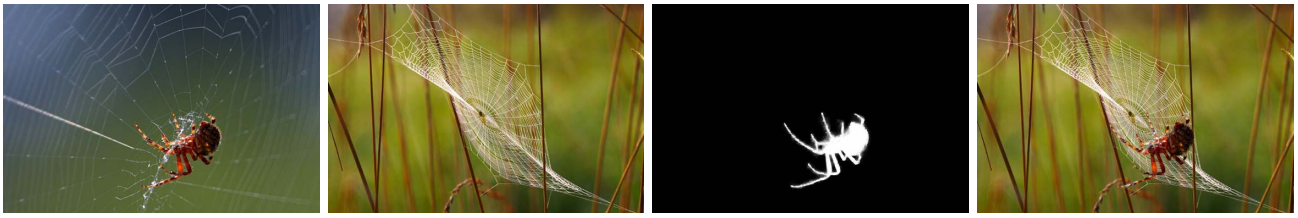
# Quadratic Markovian Probability Fields for Image Binary Segmentation

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



(a) $I_1$ (b) $I_2$ (c) $\alpha$ (d) $g = \alpha I_1 + (1 - \alpha) I_2$

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

## Abstract

*We present a Markov Random Field model for image binary segmentation that computes the probability that each pixel belongs to a given class. We show that the computation of a real valued field has noticeable computational and performance advantages with respect to the computation of binary valued field; the proposed energy function is efficiently minimized with standard fast linear order algorithms as Conjugate Gradient or multigrid Gauss-Seidel schemes. By providing a good initial guesses as starting point we avoid to construct from scratch a new solution, accelerating the computational process, and allow us to naturally implement 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 present a meticulous comparison with state of the art methods: Graph Cut, Random Walker and GMMF. The algorithms' performance are compared using a cross–validation procedure and an automatics algorithm for learning the parameter set.*

## 1. Introduction

Two–classes image segmentation [image binary segmentation (IBS)] is an important task in image analysis and in image editing tasks. There are many problems in which the core solution algorithm is an IBS method; for instance: interactive image segmentation (trimaps) [2, 5, 13, 23, 24], organs segmentation in medical imaging (*e.g.* skull stripping) [4, 9], foreground extraction (image matting) [23, 25], motion computation [6, 14], among others. Multiclass image segmentation is also commonly implemented by successive applications of IBS methods [3, 6, 14]. Last listed applications show that any improvement to the convergence ratio, memory requirements or error reduction of IBS methods will have an important impact 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 (MRF) model and computes the probability (memberships) that a pixel belongs to a given class. It is based on the minimization of a quadratic energy function archived by solving a linear system with standard iterative algorithms, as Gauss-Seidel (GS) or Conjugate Gradient (CG) [21]. 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 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 [7].

We organize this paper as follows. In section 2 we

present a new derivation of the recently reported method in [22] for soft (probabilistic) images multiclass segmentation. The here presented derivation is rigourously justified within the Bayesian regularization theoretical framework. In section 3 we particularize the method derived in section 2 for the IBS case. In this case the class memberships (probabilities) are represented by a single Markov Random Field (MRF) and our positive definite quadratic energy function incorporates effectively the constraints of being a probability field. In section 4 we present a discussion about related formulation for images multilabel segmentation. In section section 5 we evaluate the performance of the proposed IBS method in the interactive color IBS task based on trimaps. For such a purpose we follow the implementation by Boykov and Jolly [5] and only replacing the IBS method. We used the popular Lasso's bench database, such database is used in [2] and available online [28]. The experimental results demonstrate a superior performance of our method compared to popular methods of the state of the art for IBS. We conduct a careful performance evaluation based on a *cross-validation* technique. For such purposes the algorithms hyper-parameters were automatically adjusted (optimized) and a standard benchmark data were used. In section 6 we demonstrate (by numerical experiments in both real and synthetic data) the method capabilities for the simultaneously estimation of the segmentation and of the model parameters. Finally, section 7 presents our conclusions.

## 2. Quadratic Markov Measure Field Models

MRF models are a well-accepted and powerful approach for solving problems in early computer vision and image processing [1,2], [3,5,6,8,13,14], [16,18,22], [23–25]. Recently Rivera *et al*. [22] 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 [22] 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 [18], if no prior knowledge is available, the posterior probability marginals must be equal to the normalized likelihoods. However, in the derivation presented in [22] such a constraint neither it is satisfied nor its relaxation is theoretically justified. In this section we present an alternative derivation to the formulation in [22] rigourously based on the framework of Bayesian regularization (BR) with prior MRF models. We directly derive the formulation from the observation model illustrated in Fig. 1. We assume that the image $g$ is generated with the model

$$g(x) = \alpha(x)I_1(x) + (1-\alpha(x))I_2(x), \tag{1}$$

where $x \in \mathcal{R} \subseteq \mathcal{L}$ denotes a pixel position in the region of interest (the pixel set $\mathcal{R}$) into the regular lattice $\mathcal{L}$; $I_1$ and $I_2$ are two general images and $\alpha$ is a matting factor [23,25]. We can generalize the model (1) for the case of multiple regions as:

$$g(x) = \sum_k \alpha_k(x)g(x), \tag{2}$$

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

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

where $\eta_k$ is a noise image with known distribution and the matting factors satisfy:

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

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

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

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

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), $\alpha$ can be interpreted as a probability measure field where $\alpha_k(x)$ is understood as the probability of the observed pixel $g(x)$ is taken from the data $I_k(x)$. Additionally (6) introduces the constraint on the probability vectors $\alpha(x)$ to have a 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 entries are close to zero. The constraint (7) promotes the probability measure $\alpha$ to be spatially smooth.

The segmentation of the composed image, $g$, can be seen as the solution to the ill–posed inverse problem stated in (2) and (3) subject to the *hard* constraints (4) and (5) and to the *soft* constraints (6) and (7). This is, to compute the matting factors $\alpha_k$ 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 $(\alpha^*, I^*)$ as an estimator of the posterior distribution $P(\alpha, I|g)$. Then, by using the Bayes rule, the posterior distribution can be expressed as:

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

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

In general, the inference of the images $I_k$ from the data $g$ is a complex inverse problem: even if $\alpha$ 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 $\beta$.

For simplifying the notation, we express our derivation in terms of $I$ instead of the parameters $\beta$. Parametric forms, although limited, have successfully been used for defining layered models for segmenting gray scale images or optical flows [17, 22]. User interaction is a popular form for introducing prior (high level) knowledge for segmenting images with complex scenes. In that paradigm the user labels by hand a subset of pixels and then the unknown labels are estimated with a segmentation algorithm that takes into account the distribution of the labeled pixels and the smoothness of the spatial segmentation.

To derive $P(g|\alpha, I)$ we assume that $\eta_k$ is i.i.d. Gaussian noise with mean zero and variance $\sigma_k^2$, i.e.:

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

where we define $G_\sigma(z) \stackrel{def}{=} 1/\sqrt{2\pi}\sigma \exp\left[-|z|^2/2\sigma^2\right]$.

From (3) we have: $\alpha_k(x)\eta_k(x) = \alpha_k(x)(g(x) - I_k(x))$. As $\alpha_k$ 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 $r_k(x) = g(x) - I_k(x)$):

$$\begin{aligned} P(\alpha_k g|\alpha_k I_k, \sigma_k^2) \quad &= \quad \prod_x G_{\sigma_k}\big(\alpha_k(x) r_k(x)\big) \\ &= \quad \prod_x G_{\sigma_k}\big(r_k(x)\big)^{\alpha_k^2(x)}. \end{aligned} \tag{10}$$

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 [12]:

$$v_k(x, \theta_k) = \sum_{i=1}^{M} \pi_{ki} G_{\sigma_k}\left(r_k(x) - m_{ki}\right), \tag{11}$$

with $\theta_k = (\sigma_k, \pi_k, m_k)$; where $\pi_{ki} \geq 0$ are the mixture coefficients (with $\sum_i \pi_{ki} = 1$); where the mixture parameters are assumed known: the Gaussians centers $m_k = (m_{k1}, m_{k2}, \ldots, m_{kI})$ variances $\sigma_k$ and the number (maybe large) of Gaussians $M$. Then we have

$$P(\alpha_k g|\alpha_k I_k, \theta_k) \quad = \quad \prod_x \left[\sum_{i=1}^{M} \pi_{ki} G_\sigma\left(r_k(x) - m_{ki}\right)^{\alpha_k^2(x)}\right] \tag{12}$$

and in the low entropy limit we can approximate :

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

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

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

In particular, such an independency occurs if (6) is satisfied. In order to impose an explicit entropy control we introduce the Gini's potential $\mu\left(1 - \sum_k \alpha_k^2(x)\right)$, with $\mu > 0$ we promotes low entropy [12]. Additionally the region smoothness is promoted using a Gibbsian distribution based on MRF models. We finally obtain the prior $P(\alpha)$:

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

where $Z$ is a constant. If a prior uniform distribution on $I$ and independence among $I$ and $\alpha$ are assumed ($P(\alpha, I) \propto P(\alpha)$) then the posterior distribution takes the form $P(\alpha, \theta|g) \propto \exp\left[-U(\alpha, \theta)\right]$ and the MAP estimator is computed by minimizing the energy function:

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

subject to the constraints (4) and (5). This quadratic programming problem can efficiently be solved by incorporating the equality constraints (4) in a Lagrangian (in the Lagrange multipliers method) and using a projection strategy for the non-negativity constraint (5). The convergence of the algorithm is guaranteed to a local minima, see [22].

## 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 be the normalized likelihoods corresponding to the first and second classes:

$$\hat{v}_k(x, \theta) = \frac{v_k(x, \theta_k)}{s(x, \theta)}, \tag{17}$$

for $k = 1, 2$; with

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

then we define the distances:

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

Then our IBS method can be formulated as the minimization of the unconstrained quadratic cost function:

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

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 a unique global minimum that satisfies $\alpha(x) \geq 0, \forall x \in \mathcal{R}$.*

*(ii) Otherwise a local minima can be computed with an energy descend algorithm 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. $\alpha$ can be solved with the Gauss-Seidel (GS) scheme

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

with

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

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

where $\sharp \mathcal{N}_x$ 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 $\{\alpha^t(x)\}_{t=1,\ldots,T} \in [0,1]$, 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 $\alpha^0$.

*Proof of (ii).* It follows from the fact that any descent algorithm that produces a feasible sequence, $\{\alpha^t\}_{t=1,\ldots,T}$, for solving an indefinite quadratic (linearly constrained) problem converge to, at least, a local minima, [21]. □

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 $\{\alpha^t\}_{t=1,\ldots,T}$ such that: $Q(\alpha^0) \geq \ldots \geq Q(\alpha^i) \geq Q(\alpha^{i+1}) \geq \ldots \geq Q(\alpha^*) \geq 0$; 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 $b(x^t) \geq a(x^t) \geq 0, \forall t$). The feasibility of providing initial guesses and having partial solutions (by stopping the algorithm iterations before convergence) allow us to implement fast multigrid algorithms.

## 4. Relationship with other Markov Measure Fields Models for Soft Segmentation

A Markov measure field (MMF), $\alpha$, is a random vectorial field that satisfies (4) and (5) with a Gibbsian prior distribution $P(\alpha)$ in terms of MRF models [16]. In its original formulation, the image segmentation task is a combinatorial problem: to assign a class label to each pixel. Differently to hard segmentation schemes that directly compute the label map, the MMF paradigm propose to compute the probability (posterior marginals) that a pixel can be generated with a particular intensity model. In this study we discuses relationships between MMF models. Such algorithms are implemented by the minimization of posterior energies of the form:

$$U(\alpha) \quad = \quad D(\alpha, g) + \lambda R(\alpha). \tag{24}$$

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

*Gaussian MMF (GMMF)* [18]. In this framework the posterior marginals are directly modelled and estimated. Such framework construct of fact that if not prior knowledge is provided then Maximized of the Posterior Marginal, or MPM estimator, of a posterior distribution coincides with the Maximum Likelihood (ML) estimator [19] [$P(\alpha|g) = P(g|\alpha) \iff P(\alpha)$ is the uniform distribution]. In particular the GMMF potential,

$$\sum_k \sum_x (\alpha_k(x) - \hat{v}_k(x, \theta_k))^2 + \lambda R(\alpha), \tag{25}$$

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

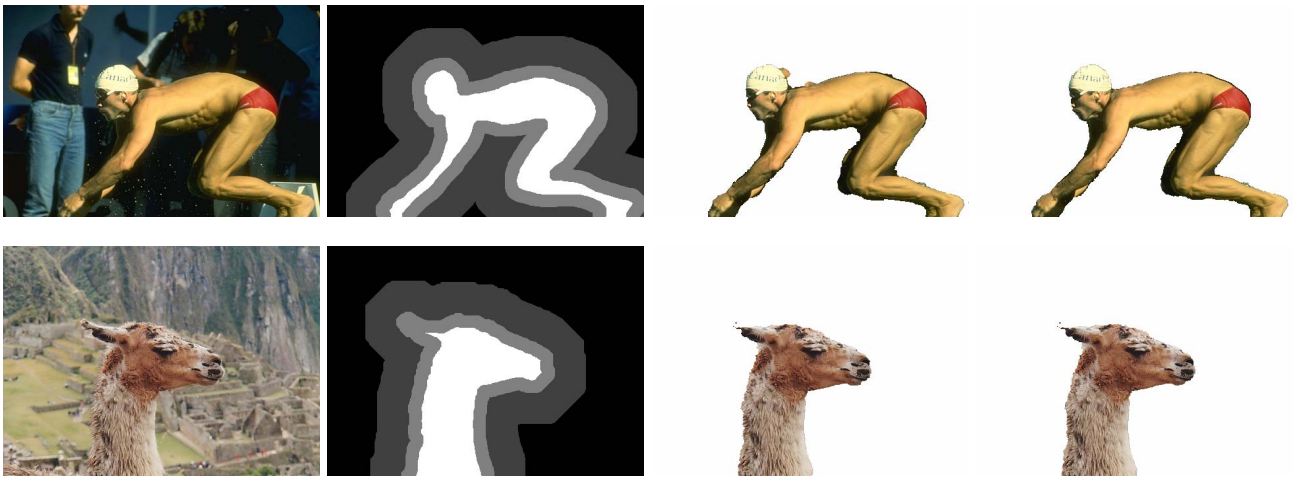
$$\alpha_k(x) = \hat{v}_k(x, \theta_k) \stackrel{def}{=} \frac{v_k(x, \theta_k)}{s(x)} \tag{26}$$

is satisfied, see (17).

*Random Walker (RW)* [10, 11]. 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:

$$s(x)\alpha_k(x) = v_k(x, \theta_k). \tag{27}$$

Then the corresponding potential is a quadratic one such that the minimum for $\lambda = 0$ results in (27) and consequently



(e) Original (f) Trimap (g) QMPF (h) QMPF+EC

Figure 2. Selected trimaps segmentations.

satisfies the GMMF *consistence condition* (26). The image coloring procedure proposed Levin *et al*. [15] 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 (16) for the case of $\lambda = 0$ corresponds to:

$$\alpha_k(x) = \frac{1}{K}\frac{H(d(x))}{d_k(x)}; \tag{28}$$

where $H(d) \stackrel{def}{=} K(\sum_{i=1}^{K} d^{-1})^{-1}$ is the *harmonic mean* of $d$. As the GMMF–*consistence condition* is not satisfied by (16), it does not corresponds to a GMMF model.

| | Params. | AIC | Training | Testing |
|---|---|---|---|---|
| Graph cut | $\lambda, \gamma$ | 8.58 | 6.89% | 6.93% |
| Rand. Walk. | $\lambda, \gamma$ | 6.50 | 5.46% | 5.50% |
| GMMF | $\lambda, \gamma$ | 6.49 | 5.46% | 5.49% |
| QMPF | $\lambda, \gamma$ | 6.04 | 5.02% | 5.15% |
| QMPF+EC | $\lambda, \gamma, \mu$ | 5.39 | 3.13% | 3.13% |

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

## 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 methods generalization capabilities [12]. The benchmark data is the set of 50 images in the Lasso's database used in [2] and available online in [28]. Such a database contains a natural images set with their corresponding trimaps and the ground truth segmentations. Actually, a Lasso's trimap is an 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 a 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 the above class enumeration. In this case, the region to process is labeled as "unknown" and the boundary conditions are imposed by the foreground and background labeled regions. The regularization term in (20) is replaced by:

$$\lambda \sum_{y \in \tilde{\mathcal{N}}_x} [\alpha(x) - \alpha(y)]^2 \, l_{xy}, \tag{29}$$

where $\tilde{\mathcal{N}}_x = \{y \in \mathcal{R} \cup \mathcal{B} \cup \mathcal{F} : |x - y| = 1\}$ and

$$l_{xy} = \frac{\gamma}{\gamma + \|g(x) - g(y)\|^2} \tag{30}$$

is an affinity measure that takes a value close to one if the neighbor pixels $x$ and $y$ have similar colors and close to zero otherwise. This affinity measure leads the border regions (classes) to follow the color edges and $\gamma$ is a method's hyper–parameter that controls the edge sensibility. We noted that the color image, $g$, is previously transformed to the CIE–Lab color space with the Ruzon's C-code library [27]. 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 (30) [23] [9]. However, in our experiments, we use the simple form (30) for comparing directly the methods performance.

In this task, empirical likelihoods are computed from the histogram of the labeled 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, $h_2$, labeled pixels. Then the normalized likelihoods are computed with:

$$\hat{v}_k(x) = \frac{h_k(g(x)) + \epsilon}{h_1(g(x)) + h_2(g(x)) + 2\epsilon}, \tag{31}$$

for $k = 1, 2$; where $\epsilon = 10^{-4}$ 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 $\alpha$ with (28). The normalized histograms can be seen as 3D Look-Up-Table with $50 \times 100 \times 100$ dimensions for the *Lab* coordinate space.

A hard segmentation can be computed by labelling each pixel $x$ with the class 1 if $\alpha(x) > 0.5$, otherwise with the class 2. For implementing the cross validation we follow the recommendation in [12] and part the data set in 5 groups of 10 images. The parameters set were trained by minimizing the mean of the segmentation error (computed according to [2]) in groups of 40 images by using the Nelder and Mead simplex descent [20]. 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 optimized (trained) parameters with the 50 image in the database [12]. Note that the AIC is consistent with the cross-validation results: the order in the methods performance is preserved. 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 the learned parameter $\mu$ for QMPF+EC promotes large entropy, such parameter was appropriated for the trimap segmentation task and should not produce the expected results in other tasks. However the entropy control allows one to adapt the algorithm for different tasks, for instance, we compute the matting factors for the example illustrated in Fig. 1 and 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 efficiently been estimated by using an alternated minimization scheme of the cost function (20) *w.r.t.* the MMF, $\alpha$, and the parameters, $\theta$. 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 $\alpha$ and $\theta_k = \left[I_k, \sigma_k^2\right]^T$ for $k = 1, 2$. As $\eta_k(x)$ is Gaussian, we have:

$$-\log v_k(x, \theta) = \frac{1}{2\sigma_k^2}|g(x) - I_k|^2 + \log\sqrt{2\pi}\sigma_k. \tag{32}$$

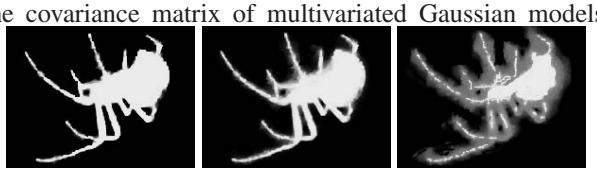
Then, by assuming a uniform distribution as prior for $\theta$, from the partial derivatives w.r.t. the parameters, we have:

$$I_k = \frac{\sum_x \alpha^2(x) g(x)}{\sum_x \alpha^2(x)} \tag{33}$$

and

$$\sigma_k^2 = \frac{\sum_x \alpha^2(x)|g(x) - I_k|^2}{\sum_x \alpha^2(x)}. \tag{34}$$

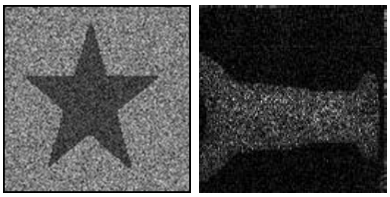
Formulas (33) and (34) are similar to the ones obtained in an Expectation-Maximization (EM) procedure; except by the $\alpha^2(x)$ weighting factor instead of $\alpha(x)$. Such a factor is also changed for estimating the covariance matrix of multivariated Gaussian models.



(a) QMPF $\mu = 10$. (b) $\mu = 0$. (c) $\mu = -123$.

(d) GMMF. (e) Rand. Walk. (f) Graph Cut.

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.

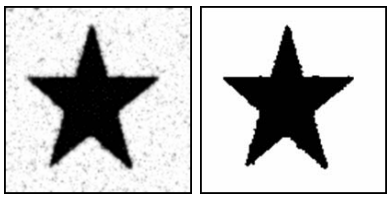


(a) Synthetic. (b) Speckle.

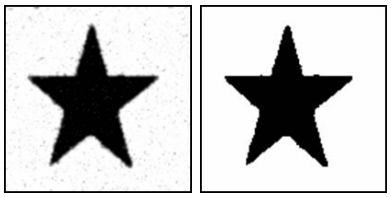
Figure 4. Test Images.

| | $I_1$ | $\sigma_1$ | $I_2$ | $\sigma_2$ |
|---|---|---|---|---|
| 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.



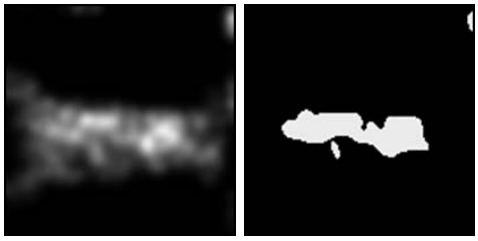
(a) $\alpha$ with $\mu = 0.0$. (b) Segmentation.



(c) $\alpha$ with $\mu = 0.5$. (d) Segmentation.

Figure 5. Entropy control.

(a) $\alpha$ with QMPF. (b) QMPF Segmentation.



(c) $\alpha$ with GMMF (RW). (d) GMMF Segmentation.

Figure 6. Effect of the data term.

Figure 4 shows the pair of images used in 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, $\mu$, is showed in Fig. 5. The computed $\alpha$ 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 robust to the exact value of the entropy control. The models ($I_1$ and $I_2$) where initialized with the maximum and minimum image gray values, respectively.

Fig. 6 shows the corresponding results to the speckle image. Fig. 6(a) shows the computed $\alpha$ field with the proposed QMPF method (with $\mu = 0$ and $\lambda = 1 \times 10^3$) and Fig. 6(b) shows the corresponding segmentation. Second row shows the computed results with GMMF. The computed $\alpha$ field with the GMMF algorithm has, evidently, larger entropy than the QMPF solution. This is consistent with the results reported in Refs. [17] and [22]. If such a high–entropy $\alpha$ 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.

## 7. Discussion and Conclusions

We started our paper by presenting a new derivation of the Markovian models for multi–class image segmentation presented in [22]. Our derivation is accord with the Bayesian Regularization framework. We have named Quadratic Markovian Measure Fields (QMMF) such models and have exposed the relationship (and difference) with the GMMF models (or Random Walker). The algorithm computes a low entropy (almost binaries) and regularized (smooth) vector field, $\alpha$. Such that $\alpha_k(x)$ that can be interpreted as the probability of the pixel $x$ were generated with the distribution $k$. We have theoretically proof and experimentally demonstrate that the QMMF models accept generic likelihood distribution. For instance empirical likelihoods were used in the trimaps assisted segmentation.

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 replacing our algorithm with state of the art methods: Graph Cut, Random Walker and GMMF. In such a evaluation the remaining implementations details were unaltered. As benchmark data 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 in [22]. 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 accelerates the interactive process by avoiding to construct from scratch the new solution.

Our algorithm demonstrated a superior performance than

graph cuts in the IBS task and is easier to implement. Although graph cuts algorithms are computationally efficient (low polynomial order algorithms), our approach is substantially faster given that can be implemented with algorithm of lineal order [say conjugate gradient (CG)] or multigrid Gauss-Seidel (MGS) schemes. In particular, we found that the Lassos trimaps are faster segmented with MGS than CG. The reason maybe that the cardinality of the set of unknown labeled pixels is relatively small w.r.t. the size of the image. But for a trimap provided for a common user (only few pixels are labeled) CG shown a faster performance. In our experiments, the Lasso's database was processed an about a half of the time with MGS than with graph cut. We can expect that, given the computational order of such algorithms, as the size of the unknown label pixels set grows, such difference would be more noticeable. While in this work we have evaluate the algorithms' accurateness, in future work we will focus in evaluate the computational time efficiency.

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 be computed with fast methods based on graph cuts [26].

## References


[1] J. Besag. On the statistical analisys 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, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE PAMI*, 23(11):1222–1239, 2001.

[7] W. L. Briggs, S. McCormick, and V. Henson. *A Multigrid Tutorial*. SIAM Publications, second edition, 2000.

[8] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE PAMI*, 6:721–741, 1984.

[9] 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.

[10] 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.

[11] L. Grady. Multilabel random walker image segmentation using prior models. In *CVPR*, volume 1 of *CVPR*, pages 763–770, 2005.

[12] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2001.

[13] O. Juan and R. Keriven. Trimap segmentation for fast and user-friendly alpha matting. In *VLSM, LNCS 3752*, pages 186–197, 2005.

[14] 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.

[15] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *CVPR (1)*, pages 61–68, 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. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill–posed problems in computational vision. *J. Am. Stat. Asoc.*, 82:76–89, Nov 1987.

[20] J. A. Nelder and R. Mead. A simplex method for function minimization. *Comput. J.*, 7:308–313, 1965.

[21] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operation Research, 2000.

[22] 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.

[23] 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.

[24] M. A. Ruzon and C. Tomasi. Alpha estimation in natural images. In *CVPR (1)*, pages 18–25, 2000.

[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.