COMPARATIVE STUDY ON 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

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.

Keywords: Image binary segmentation, Segmentation comparison, Interactive computer vision, Markov random fields, energy minimization, Image matting.

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) Ruzon & Tomasi (27); Boykov & Jolly (5); Blake at al. (2); Rother et al. (26); Juan & Keriven (13) organs segmentation in medical imaging (*e.g.* skull stripping) Boykov & Jolly (6); Grady et al. (10), foreground extraction (image matting) Rother et al. (26); Wabg & Cohen (28), motion computation Boykov etal. (4); Kolmogorov et al. (14), among others. Multiclass image segmentation is also commonly implemented by the successive applications of IBS methods Bouman & Shapiro (3); Boykov etal. (4); Kolmogorov et al. (14).

In this paper we present a novel IBS method 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) Nocedal & Wright (22). As it is well known, the convergence ration 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 Brigg et al. (7).

We organize this paper as follows. In section 2 we presents a new derivation of the recently reported method by Rivera *e*t al.(23) 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. MRF models are a well-accepted and powerful approach for solving problems in early computer vision and image processing Besag (1); Blake at al. (2); Bouman & Shapiro (3); Boykov & Jolly (5); Boykov etal. (4); Geman & Geman (8); Juan & Keriven (13); Kolmogorov et al. (14); LI (16); Marroquin et al. (19); Rivera et al. (23); Rother et al. (26); Ruzon & Tomasi (27); Wabg & Cohen (28). In subsection 3.2 we present a discussion about related formulation for multilabel image segmentation. Then in section 3.3, 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.

In section section 4 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 & Jolly (5) and then we only replace the IBS method. We used the popular Lasso's bench database, such database is used in Blake at al. (2) and available online in Ref. (30). The experimental results demonstrate a superior performance of our method compared with methods of the state of the art for IBS. We conducts a deeper evaluation based on a *cross-validation* technique was used for such purposes and the algorithms hyper-parameters were automatically adjusted. Finally, section 5 presents our conclusions.

2. Entropy–Controlled Quadratic Markov Measure Field Models

Recently, Rivera et al. (23); Rivera at al. (25) proposed the Entropy Controlled Quadratic Markov Measure Fields (EC-QMMF) models for image multiclass segmentation. Such an algorithm is computationally efficient and produces "soft" segmentations of excellent quality. In subsection 2.1 we present a variant derivation EC-QMMF models for the case of Gaussian noise (Gaussian Likelihoods). The presented variant can formally been generalized to distinct noise models than Gaussian, see subsection 2.2.

2.1. **Revised Gaussian EC-QMMF models derivation.** We assume that the image g is generated with the model

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

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 α is a matting factor Rother et al. (26); Wabg & Cohen (28) and η is a possible noise. We can generalize the model (1) for the case of multiple regions as:

$$(2) \qquad g(x) = \sum_k \alpha_k(x)g(x) + \eta_k(x),$$

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

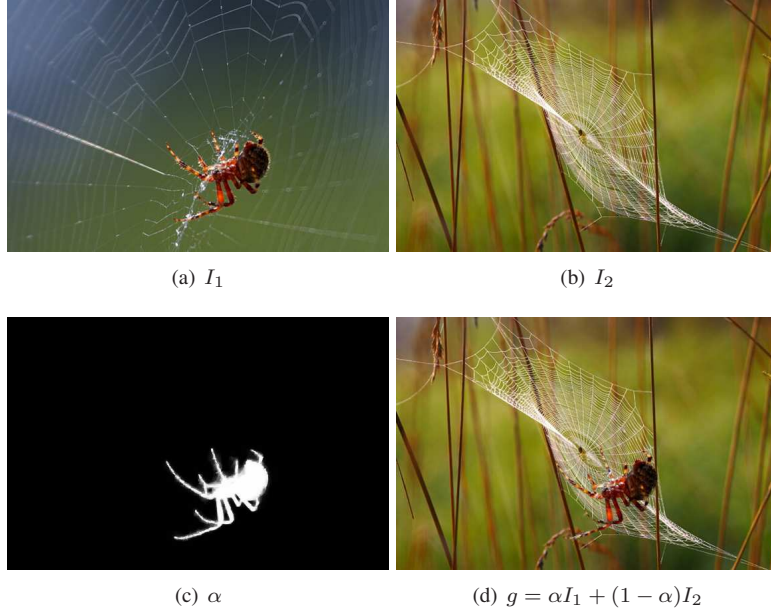$$(3) \qquad \alpha_k(x)g(x) = \alpha_k(x)(I_k(x) + \eta_k(x));$$

(a) I_1

(b) I_2

(c) α

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

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

η_k is a noise image with known distribution and the matting factors satisfy:

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

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

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

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

$\mathcal{N}_x$ denotes in (7) 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 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 α 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 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:

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

3

where $P(g|\alpha, I)$ is the conditional probability of the data by assuming given the unknowns (α, I), $P(\alpha, I)$ is the prior distribution and $Z = P(g)$ is a normalization constant (independent on (α, I)).

For the general case illustrated in Fig. 1, a more realistic posterior distribution could take the form $P(\alpha, \{\alpha_k I_k\}|g) \propto P(g|\alpha, \{\alpha_k I_k\})P(\alpha, \{\alpha_k I_k\})$. Moreover, there are cases in which one can recover the entire images $\{I\}$ if such images can be represented by parametric functions: $I_k(x) = \Phi(x, \beta)$, with parameters β. Thus the posterior distribution is in terms of the parameters β instead of the images $\{I\}$. Parametric forms, although limited, have successfully been used for defining layered models for segmenting gray scale images or optical flows Marroquin et al. (17, 20); Rivera et al. (23); Rivera at al. (25). In this work, we will use the form (8) for simplicity.

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

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

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 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 $r_k(x) = g(x) - I_k(x)$):

$$(10) \qquad P(\alpha_k g|\alpha_k I_k, \sigma_k^2) = \prod_x G_{\sigma_k}\big(\alpha_k(x)r_k(x)\big) = \prod_x G_{\sigma_k}\big(r_k(x)\big)^{\alpha_k^2(x)}.$$

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

$$(11) \qquad P(g|\alpha, I, \sigma) = \prod_k P(\alpha_k g|\alpha_k I_k, \sigma_k).$$

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 Hastie et al. (12). Additionally the region smoothness is promoted using a Gibbsian distribution based on MRF models. We finally obtain the prior:

$$(12) \qquad P(\alpha) = \frac{1}{Z} \exp\left[\sum_{x \in \mathcal{R}}\left(\mu\|\alpha(x)\|^2 - \frac{\lambda}{2}\sum_{y \in \mathcal{N}_x}\|\alpha(x) - \alpha(y)\|^2\right)\right];$$

where Z is a constant. If a uniform prior distribution on I and independence among I and α 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:

$$(13) \quad U(\alpha, \theta) = \sum_{x \in \mathcal{R}}\left\{\sum_{k=1}^K \alpha_k^2(x)\left[-\log G_{\sigma_k}\big(r_k(x)\big) - \mu\right] + \frac{\lambda}{2}\sum_{y \in \mathcal{N}_x}\|\alpha(x) - \alpha(y)\|^2\right\},$$

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):

$$(14) \qquad \min_x \max_\pi \mathcal{L}(x, \pi) = U(\alpha, \theta) - \sum_{x \in \mathcal{R}} \pi(x)\left(\left(1 - \sum_k \alpha_k(x)\right)\right.$$

subject to $\alpha_k(x) \geq 0$. The convergence of the algorithm is guaranteed to a local minima, see Rivera et al. (23).

2.2. General EC-QMMF models derivation.
For removing the Gaussian noise assumption, we first consider that any smooth density distribution v_k can be expressed with a Gaussian mixture model Hastie et al. (12):

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

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 $m_k = (m_{k1}, m_{k2}, \ldots, m_{kI})$, the variances σ_k and the number (maybe large) of Gaussians M. Then we have

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

and in the low entropy limit we can approximate :

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

Following the reasoning in previous subsection, the likelihood of the observed (composed) image g is given by

$$
(17) \qquad P(g | \alpha, I, \theta) = \prod_k P(\alpha_k g | \alpha_k I_k, \theta_k)
$$

and we have the general posterior energy form:

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

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

3. QUADRATIC MARKOV PROBABILITY FILED MODELS

3.1. Mathematical Development.
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:

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

for $k = 1, 2$; with

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

then we define the distances:

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

5

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

(22)
$$Q(\alpha) = \sum_{x \in \mathcal{R}} \left\{ \alpha^2(x) \left[d_1(x) - \mu \right] + (1 - \alpha(x))^2 \left[d_2(x) - \mu \right] + \lambda \sum_{y \in \mathcal{N}_x} (\alpha(x) - \alpha(y))^2 \right\}.$$

The minimization convergence properties of (22) are established in the following theorem.

Theorem 1. *QMPF global convergence conditions.*

Let be μ choose such that $\mu < \min_{k,x} d_k(x)$, then (22) has a unique global minimum that satisfies $\alpha(x) \geq 0, \forall x \in \mathcal{R}$.

Proof. 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 (22) w.r.t. α can be solved with the Gauss-Seidel (GS) scheme

(23)
$$\alpha(x) = \frac{a(x)}{b(x)}$$

    with

(24)
$$a(x) \stackrel{def}{=} d_2(x) - \mu + \lambda \sum_{y \in \mathcal{N}_x} \alpha(y),$$

(25)
$$b(x) \stackrel{def}{=} d_1(x) + d_2(x) - 2\mu + \lambda \sharp \mathcal{N}_x;$$

    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 (23) 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 $[0,1]$.

    (c) 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.

    $\square$

Otherwise (if the condition in Theorem 1 is not satisfied) converge to, at least, a local minima can be guaranteed if the non-negativity constraint $\alpha_k(x) \geq 0$ is enforced. That can be implemented in a gradient projection kind algorithm that produces a feasible sequence, $\{\alpha^t\}_{t=1,\ldots,T}$, for solving an indefinite quadratic (linearly constrained) problem Nocedal & Wright (22).

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 (23). 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.

3.2. **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 LI (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 discusses relationships between MMF models. Such algorithms are implemented by the minimization of posterior energies of the form:

$$(26) \qquad U(\alpha) = D(\alpha, g) + \lambda R(\alpha).$$

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) Marroquin et al. (19). In this framework the posterior marginals are directly modelled and estimated. Such framework constructs on 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 Marroquin et al. (18) $[P(\alpha|g) = P(g|\alpha) \iff P(\alpha)$ is the uniform distribution]. In particular the GMMF potential,

$$(27) \qquad \sum_k \sum_x (\alpha_k(x) - \hat{v}_k(x, \theta_k))^2 + \lambda R(\alpha),$$

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

$$(28) \qquad \alpha_k(x) = \hat{v}_k(x, \theta_k) \overset{def}{=} \frac{v_k(x, \theta_k)}{s(x)}$$

is satisfied, see (19).

Random Walker (RW) Grady et al. (11); Grady (9). Although introduced in terms of random walks of particles, RW is a variant of the GMMF formulation (see the diffusion process in Marroquin et al. (19)). In that framework, the *consistence condition* is reformulated as:

$$(29) \qquad s(x)\alpha_k(x) = v_k(x, \theta_k).$$

Then the corresponding potential is a quadratic one such that the minimum for $\lambda = 0$ results in (29) and consequently satisfies the GMMF *consistence condition* (28). The image coloring procedure proposed by (15) is close related with the GMMF diffusion process with space-varying weights Marroquin et al. (19).

Quadratic MMF [This work]. Differently to GMMF models, the minimum of the QMPF potential (18) for the case of $\lambda = 0$ corresponds to:

$$(30) \qquad \alpha_k(x) = \frac{1}{K} \frac{H(d(x))}{d_k(x)};$$

where $H(d) \overset{def}{=} K(\sum_{i=1}^{K} d^{-1})^{-1}$ is the *harmonic mean* of d. As the GMMF–*consistence condition* is not satisfied by (18), it does not corresponds to a GMMF model.

7

	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 1. Computed parameters for Fig. 3.

3.3. **Model parameters estimation.** In the QMPF model, if Gaussian likelihoods are assumed, the parameters $\theta_k = [m_k, \sigma_k]$ (mean and standard deviation, respectively) can been efficiently estimated by using an alternated minimization scheme of the cost function (22) *w.r.t.* the MMF, α, and the parameters, θ. In such case

$$(31) \qquad -\log v_k(x, \theta) = \frac{1}{2\sigma_k^2} |g(x) - I_k|^2 + \log \sqrt{2\pi}\sigma_k.$$

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

$$(32) \qquad I_k = \frac{\sum_x \alpha_k^2(x) g(x)}{\sum_x \alpha_k^2(x)}$$

and

$$(33) \qquad \sigma_k^2 = \frac{\sum_x \alpha_k^2(x) |g(x) - I_k|^2}{\sum_x \alpha_k^2(x)}.$$

Such formulas, (32) and (33), 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, see experiments in Rivera & Mayorga (24), Section 6.

For illustrating this capability, we consider the task of computing a binarization of a synthetic image (Fig. 2a) 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 $\theta_k = \left[I_k, \sigma_k^2\right]^T$ for $k = 1, 2$.

Figure 2 shows the pair of images used this experiment. The synthetic binary image, in Fig. 2(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 2(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. 3. The computed α field with $\mu = 0$ (without entropy control) and the corresponding binarization are shown in Figs. 3(a) and 3(b). Figs. 3(c) and 3(d) show the results computed with $\mu = 0.5$. Table I 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.
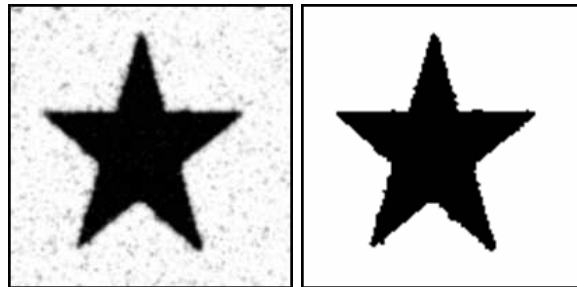
Fig. 4 shows the results corresponding to the speckle image. In Fig. 4(a) we show the computed α field with the proposed QMPF method (with $\mu = 0$ and $\lambda = 1 \times 10^3$) and Fig. 4(b) the corresponding segmentation. Second row shows the computed results with GMMF. The computed α field with the GMMF algorithm has, evidently, larger entropy
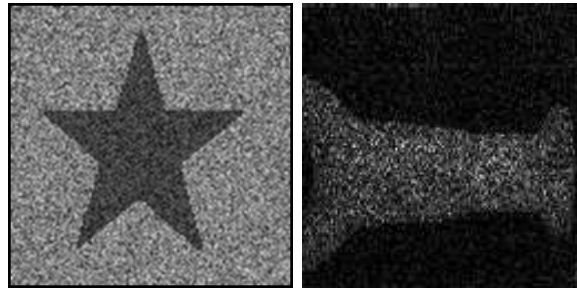
(a) Synthetic.

(b) Speckle.

FIGURE 2. Test Images.

(a) α field with $\mu = 0.0$.

(b) Segmentation.

(c) α field with $\mu = 0.5$.

(d) Segmentation.

FIGURE 3. Entropy control.

than the QMPF solution. This is consistent with the results reported by Marroquin et al. (17); Rivera et al. (23); Rivera at al. (25). 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 Marroquin et al. (17). As it is expected, we observed a similar behavior for the Random Walker algorithm than for GMMF.
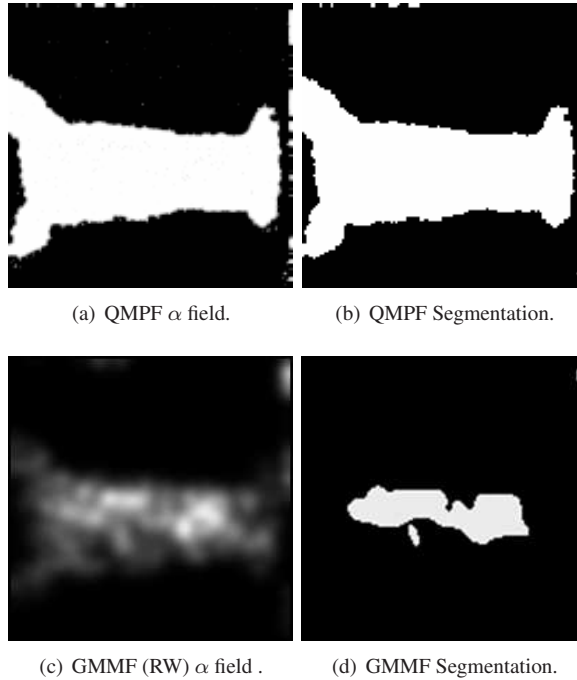
(a) QMPF α field.

(b) QMPF Segmentation.

(c) GMMF (RW) α field .

(d) GMMF Segmentation.

FIGURE 4. Effect of the data term.

4. IMAGE BINARY INTERACTIVE SEGMENTATION

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 labelled pixels and the smoothness of the spatial segmentation. Fig. 5 illustrate the interactive IBS process. These results were computed with the proposed algorithm.

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 Hastie et al. (12). The benchmark data is the set of 50 images in the Lasso's database used by Blake at al. (2) and available online in Ref. (30). 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. 6 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
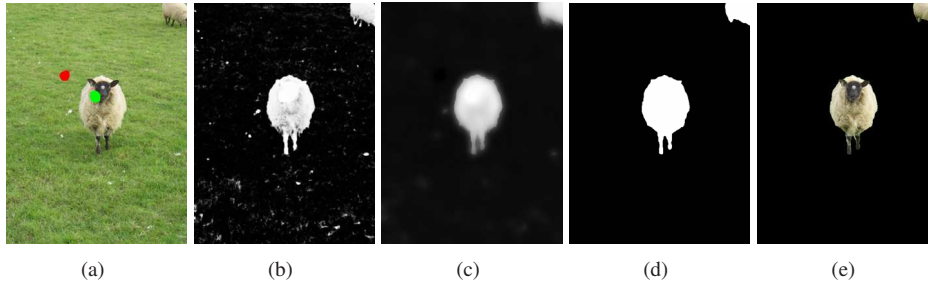
10

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

FIGURE 5. Interactive binary segmentation process illustration: (a) Pixels labelled by hand (scribbles), (b) Likelihoods computed from the empirical class distributions, (c) computed α map, (d) label map (maximum α) and (e) segmented image.

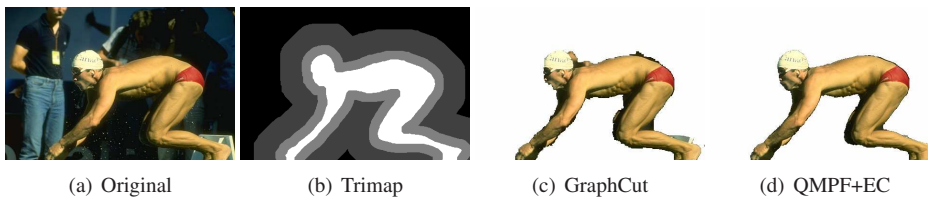(a) Original   (b) Trimap   (c) GraphCut   (d) QMPF+EC

FIGURE 6. Segmentation example from the Lasso's data set (image file 153077).

background labeled regions. The regularization term in (22) is replaced by:

$$\lambda \sum_{y \in \tilde{\mathcal{N}}_x} [\alpha(x) - \alpha(y)]^2 \, l_{xy}, \tag{34}$$

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

$$l_{xy} = \frac{\gamma}{\gamma + \|g(x) - g(y)\|^2} \tag{35}$$

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 γ 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 in Ref. (29). 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 (35) (26; 10). However, in our experiments, we use the simple form (35) for comparing directly the methods performance.

Table 2: Comparative performance on the Lasso's data set. Adjusting the best parameters to the entire data set.

Filename	Group	Algorithms' MSE				
		GraphCut	GMMF	RW	QMPF	QMPF+EC
21077	1	7.82%	3.84%	3.77%	3.84%	4.01%
						Continued on next page

Table 2 – continued from previous page

Filename	Group	Algorithms MSE				
		GraphCut	GMMF	RW	QMPF	QMPF+EC
24077	2	16.68%	10.44%	10.48%	11.00%	4.21%
37073	3	8.84%	5.40%	5.43%	5.06%	1.44%
65019	4	6.49%	2.32%	2.20%	3.68%	0.27%
69020	5	7.99%	6.29%	6.28%	5.80%	2.95%
86016	1	7.49%	3.39%	3.37%	3.01%	1.99%
106024	2	10.05%	8.89%	8.93%	8.20%	7.55%
124080	3	3.57%	5.14%	5.22%	3.40%	3.43%
153077	4	13.04%	3.99%	3.97%	6.32%	1.65%
153093	5	4.75%	4.30%	4.35%	3.70%	4.08%
181079	1	13.42%	13.16%	13.14%	12.19%	7.41%
189080	2	11.66%	10.79%	10.77%	10.44%	6.22%
208001	3	3.53%	3.21%	3.23%	2.27%	1.50%
209070	4	5.98%	3.86%	3.85%	3.82%	2.25%
227092	5	3.99%	2.40%	2.37%	3.90%	3.46%
271008	1	4.68%	3.43%	3.48%	3.47%	2.33%
304074	2	16.69%	6.10%	6.12%	8.29%	10.90%
326038	3	11.30%	14.73%	14.87%	9.00%	7.53%
376043	4	15.95%	11.06%	11.05%	11.47%	6.14%
388016	5	1.64%	3.45%	3.51%	2.39%	1.50%
banana1	1	12.21%	10.45%	10.44%	10.31%	3.91%
banana2	2	2.81%	5.08%	5.06%	3.41%	1.49%
banana3	3	5.69%	8.04%	8.10%	5.53%	1.91%
book	4	8.24%	9.99%	10.01%	7.99%	3.52%
bool	5	4.01%	2.00%	2.00%	2.17%	1.74%
bush	1	13.02%	9.12%	9.04%	11.28%	7.86%
ceramic	2	4.92%	6.18%	6.18%	4.97%	1.73%
cross	3	3.30%	2.35%	2.35%	2.36%	1.75%
doll	4	2.53%	2.92%	2.91%	2.28%	1.03%
elefant	5	8.51%	6.19%	6.13%	5.83%	2.05%
flower	1	1.23%	0.71%	0.71%	0.58%	0.61%
fullmoon	2	0.95%	0.88%	0.88%	0.80%	0.27%
grave	3	1.84%	1.86%	1.89%	1.44%	1.27%
llama	4	18.30%	14.65%	14.69%	13.86%	4.32%
memorial	5	9.93%	4.75%	4.75%	5.72%	1.49%
music	1	3.43%	2.21%	2.19%	2.70%	2.26%
person1	2	3.06%	4.55%	4.69%	2.33%	1.16%
person2	3	1.35%	4.48%	4.54%	1.24%	0.71%
person3	4	6.62%	4.81%	4.81%	3.63%	0.87%
person4	5	8.46%	6.27%	6.23%	5.88%	3.27%
person5	1	7.51%	4.62%	4.62%	4.77%	2.48%
person6	2	9.87%	7.91%	7.88%	8.37%	5.19%
person7	3	1.41%	2.21%	2.26%	1.31%	0.96%
person8	4	6.15%	3.92%	3.88%	3.92%	0.93%
scissors	5	5.89%	6.38%	6.32%	6.52%	2.87%
Continued on next page						

Table 2 – continued from previous page

Filename	Group	Algorithms MSE				
		GraphCut	**GMMF**	**RW**	**QMPF**	**QMPF+EC**
sheep	1	2.48%	3.08%	3.14%	1.83%	4.53%
stone1	2	1.08%	0.72%	0.72%	0.79%	0.73%
stone2	3	0.56%	0.33%	0.32%	0.49%	0.78%
teddy	4	2.19%	2.91%	2.95%	2.23%	1.91%
tennis	5	8.69%	7.55%	7.61%	8.34%	7.31%
mean	–	6.84%	5.47%	5.47%	5.08%	3.03%
median	–	6.07%	4.59%	4.66%	3.87%	2.15%
stddev	–	4.66%	3.57%	3.58%	3.46%	2.40%

In this task, empirical likelihoods are computed from the histogram of the labeled by hand pixels. Following Boykov & Jolly (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:

$$(36) \qquad \hat{v}_k(x) = \frac{h_k(g(x)) + \epsilon}{h_1(g(x)) + h_2(g(x)) + 2\epsilon},$$

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 α with (30). The normalized histograms can be seen as 3D Look-Up-Table with $50 \times 100 \times 100$ dimensions for the *Lab* coordinate space. The hard segmentation is computed by labeling each pixel x with the class 1 if $\alpha(x) > 0.5$, otherwise with the class 2.

The parameters set were trained by minimizing the mean of the segmentation error in the image set by using the Nelder and Mead simplex descent Nelder & Mead (21). In Table 2 we show the means square error (in percents) by automatically fitting the best value parameters to the whole data set. For our implementation, the learned parameters are reported in Table 3. Additionally, we implement a cross–validation procedure following the recommendation by Hastie et al. (12) and part the data set in 5 groups of 10 images. Figure 6 shows an example of the segmented images. Table 4 shows the resume of the training (Table 5) and testing (Table6) error and the Akaike information criterion (AIC). The AIC was computed for the optimized (trained) parameters with the 50 image in the database Hastie et al. (12), Table 1. 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. We note that the learned parameter μ 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

Parameter	QMPF	QMPF+EC
λ	4.7×10^3	2.28×10^5
γ	9.14×10^{-6}	5.75×10^{-3}
μ	0.0	-5.75×10^5

TABLE 3. Adjusted parameters for the results in table 2.

Algorithm	Params.	AIC	Training	Testing
Graph cut	λ, γ	8.58	6.82%	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	3.58	3.13%	3.13%

TABLE 4. Cross-validation results. Parameters, Akaike information criterion, training and testing error.

Training Set	Training Error				
	Graph cuts	RW	GMMF	QMPF	QMPF+EC
1	6.71%	5.49%	5.48%	4.96%	2.96%
2	6.55%	5.29%	5.29%	4.83%	2.92%
3	7.51%	5.62%	5.61%	5.39%	3.30%
4	6.37%	5.32%	5.32%	4.86%	3.33%
5	6.95%	5.60%	5.59%	5.06%	3.12%
mean	**6.82%**	**5.46%**	**5.46%**	**5.02%**	**3.13%**

TABLE 5. MSE of the training stage. The k*t*h training set correspond to Lasso's data set when the files in the k*t*h group are removed (see table 2).

Testing Set	Testing Error				
	Graph cuts	RW	GMMF	QMPF	QMPF+EC
1	7.35%	5.41%	5.41%	5.33%	3.80%
2	7.96%	6.17%	6.16%	5.85%	3.96%
3	4.15%	4.89%	4.88%	3.79%	2.44%
4	8.78%	6.07%	6.07%	5.87%	2.30%
5	6.39%	4.95%	4.95%	4.94%	3.14%
mean	**6.93%**	**5.50%**	**5.49%**	**5.15%**	**3.13%**

TABLE 6. The k*t*h testing set corresponds to Lasso's files in the k*t*h group (see table 2).

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. 7. In particular the matting factor shown Fig. 1 was computed with QMPF with $\mu = 0$.

5. CONCLUSIONS

We have generalized the Quadratic Markov Measure Filed models Rivera et al. (23); Rivera at al. (25) for any data smooth probability distribution, in particular the empirical distributions estimated by histogram techniques or kernel methods.

We have presented a new quadratic energy function for Image Binary Segmentation. Let be problem of labelling a pixel x as belonging to the classes set $\{C_0, C_1\}$ then the algorithm computes a low entropy (almost binaries) and regularized (smooth) field, α. Such that $\alpha(x) \in [0, 1]$ that can be interpreted as the probability of the pixel x were generated with the distribution C_0.

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

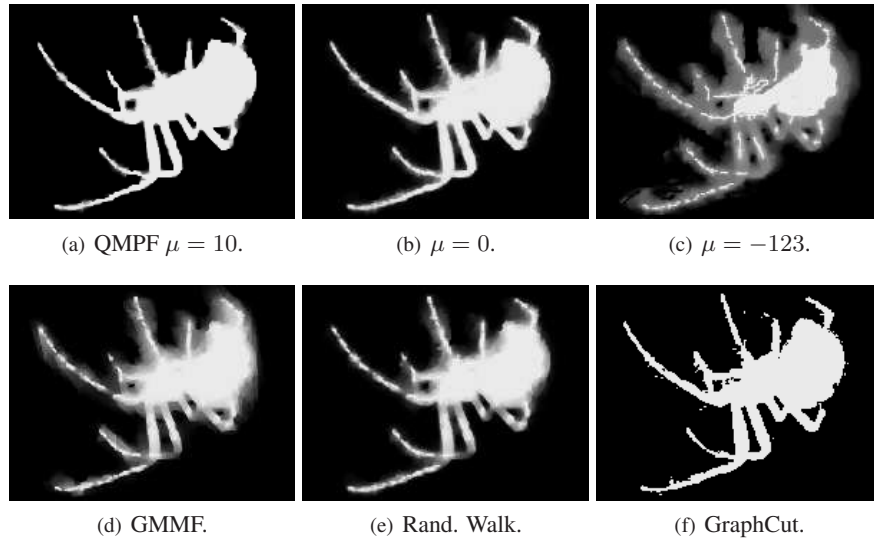(d) GMMF.   (e) Rand. Walk.   (f) GraphCut.

FIGURE 7. 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 8. Label maps corresponding to Fig. 7, same order. The QMPF method algorithm produces, in all the cases, better segmentation with smooth boundaries than GMMF, Rand. Walk. and GraphCut.

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 test data set we used the Lasso's trimap set of 50 natural images. We have achieved the meticulous algorithms comparison 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. (23); Rivera at al. (25). 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 a new solution.

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, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE PAMI*, 23(11):1222–1239, 2001.

[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 M. P. Jolly. Interactive organ segmentation using graph cuts. In *MICCAI, LNCS 1935*, pages 276–286, 2000.

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

[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] T. Hastie, R. Tibshirani and J. Friedman. *The elements of statistical learning*. Springer Series in Statistics, 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. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill–posed problems in computational vision. *J. Am. Stat. Asoc.*, 82:76–89, Nov 1987.

[19] 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.

[20] J. L. Marroquin, E. Arce, and S. Botello. Hidden Markov measure field models for image segmentation. *IEEE PAMI*, 25:1380–1387, 2003.

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

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

[23] 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.

[24] M. Rivera and P. P. Mayorga. Quadratic markovian probability fields for image binary segmentation. In *accepted to ICV'07*, 2007.

[25] M. Rivera, O. Ocegueda, and J. L. Marroquin. Entropy-controlled quadratic Markov measure field models for efficient image segmentation. *To appear in IEEE Trans. Image Processing*, 2007.

[26] 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.

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

[28] J. Wang and M. Cohen. An interactive optimization approach for unified image segmentation and matting. In *ICCV*, number 3, pages 936–943, 2005.

[29] M. A. Ruzon. http://ai.stanford.edu/∼ ruzon/software/rgblab.html.

[30] http://research.microsoft.com/vision/cambridge/i3l/segmentation/GrabCut.htm.

E-mail address, M. Rivera: `mrivera@cimat.mx`

E-mail address, P. Mayorga: `mayorga@cimat.mx`