

## *Review Article*

# **Markov Models for Image Labeling**

**S. Y. Chen,<sup>1</sup> Hanyang Tong,<sup>1</sup> and Carlo Cattani<sup>2</sup>**

<sup>1</sup> College of Computer Science, Zhejiang University of Technology, Hangzhou 310023, China

<sup>2</sup> Department of Mathematics, University of Salerno, Via Ponte Don Melillo, 84084 Fisciano (Sa), Italy

Correspondence should be addressed to S. Y. Chen, sy@ieee.org

Received 2 May 2011; Accepted 18 May 2011

Academic Editor: Gani Aldashev

Copyright © 2012 S. Y. Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Markov random field (MRF) is a widely used probabilistic model for expressing interaction of different events. One of the most successful applications is to solve image labeling problems in computer vision. This paper provides a survey of recent advances in this field. We give the background, basic concepts, and fundamental formulation of MRF. Two distinct kinds of discrete optimization methods, that is, belief propagation and graph cut, are discussed. We further focus on the solutions of two classical vision problems, that is, stereo and binary image segmentation using MRF model.

## **1. Introduction**

Many tasks in computer vision and image analysis can be formulated as a labeling problem where the correct label has to be assigned to each pixel or clique. The label of a pixel represents some property in the real scene, such as the same object or the disparity. Such problems can be naturally represented in Markov random field (MRF) model. MRF is firstly introduced into vision by S. Geman and D. Geman [1] in 1984 and has been widely used in both low-level and high-level vision perception in recent years.

Basically, humans understand a scene mainly by using the spatial and visual information which is assimilated through their eyes. Inversely, given an image or images, this information such as boundary or object, mainly based on the contextual constraints, is extremely necessary for scene interpretation. We hope to model the vision problem to capture the full interaction between pixels. On the other hand, due to the sensor noise and complexity of the real world, exact interpretation is rather difficult for computers. As a result, researchers have realized that the solution of vision problems should be solved by using optimization methods. As the most popular models for gridded image-like data, the MRF

provides a series of mathematical theories to find such optimal solutions under the contextual visual information in the images. The context-dependent object in digital images can be modeled in a convenient and consistent way through MRF theory. It is achieved through characterizing mutual influences among such entities using conditional MRF distributions [2]. Besides, the images we captured are always piecewise smooth, which can be encoded as a prior distribution. Thus, we can use the MRF model whose negative log-likelihood is proportional to a robustified measure of image smoothness [3]. Moreover, we may know that some premise or external knowledge of the scene such as the specify object might exist in the environment. With these priors, we can get more reliable understanding of the images.

In the latest two decades, the renaissance of the MRF model in computer vision has begun due to powerful energy minimization algorithms. A lot of inference algorithms have been developed to solve the MRF optimization problems, such as graph cut [4], belief propagation [5], tree-reweighted message passing [6], dual decomposition [7], fusion move [8], iterated conditional modes, and their extensions. In the literature [9], Szeliski et al. gave a set of energy minimization benchmarks and used them to evaluate the solution quality and runtime of several energy minimization algorithms. Felzenszwalb and Zabih [10] reviewed the dynamic programming and graph algorithms and then discussed their applications on computer vision. A review for the linear programming to solve max-sum problem was given in [11]. On the other hand, a framework of learning image priors for MRF was introduced by Roth and Black [12]. Schmidt et al. [13] revisited the generative aspects of MRF and analyzed the quality of common image priors in a fully application-neutral setting. New models based on MRF such as MPF were proposed. It was proved that the convex energy MPF can be used to encourage arbitrary marginal statistics [14]. Some excellent books about MRF models in image analysis such as [2] are also available.

The MRF has been successfully applied to image analysis such as restoration, matting [15], and segmentation, as well as two-dimensional (2D) fields such as stereo matching, super resolution [16], optical flow, image inpainting [17], motion estimation, and 2D-3D registration [18]. The MRF was also used to solve the high-level vision problems such as object classification [19, 20], face analysis [21], face recognition [22], and text recognition [23]. Many optimization problems can be formulated in the MRF, for example, color to gray transformation [24], feature detection scale-selection [25], and so forth. Additionally, Boykov and Funka-Lea [26] presented a survey of various energy-based techniques for binary object segmentation. S. Geman and D. Geman [1] firstly applied the MRF to image restoration. Sun et al. [27] used belief propagation algorithm and combinet it with occlusion to solve the stereo problem. Detry et al. [28] proposed an object representation framework that encodes probabilistic spatial relations between 3D features. Then the authors of [28] organize the proposed framework in the MRF.

In the remainder of the paper, Section 2 gives a sketchy of MRF and related concepts. Section 3 provides two most frequently used inference algorithms for MRF. Section 4 briefly introduces two labeling applications of MRF in low-level vision. Section 5 summarizes the contribution and offers the future works in the topic.

## 2. Problem Formulation with MRF

As a branch of probability theory [2], MRF is an undirected graphical model in which a set of random variables have a Markov property. To solve a special computer vision problem involving pixel interaction and partially observed information into an optimization problem using MRF model, we will go over the graphical models that visualize the structure of the

5	4	3	4	5
4	2	1	2	4
3	1	0	1	3
4	2	1	2	4
5	4	3	4	5

Figure 1: An example of the 5th-order neighbor system.

probabilistic models using diagrammatic representations. A graph consists of nodes and edges. Each node means an event, and each edge represents the relationship between the events. MRF is used to find the most optimal label configuration.

For a labeling problem, we need to specify a set of nodes, labels, and edges. Without loss of generality, let  $\mathbf{M}$  be a set of indexes  $\mathbf{M} = \{1, \dots, m\}$ , and let  $\mathbf{P} = \{p_1, \dots, p_m\}$  be a set of observed nodes. In vision problem, a node often represents the pixel intensity or some other image features. Let  $\mathbf{L}$  be a set of labels.  $\mathbf{L}$  can be continuous or discrete, but in most cases, all the labels we set are discrete:  $\mathbf{L} = \{l_1, \dots, l_m\}$ .

As stated above, a label means some quantity of the real scene. The simplest case is binary form where  $\mathbf{L} = \{0, 1\}$ . Such black and white model is often used to classify the foreground and background regions in the image. In general cases, the label value is more meaningful. For example, in stereo and image restoration problem, larger label value means depth information or lighter pixel intensity. Additionally,  $\mathbf{L}$  also can be unordered labels of which the value has no semantic meaning, such as for object classification.

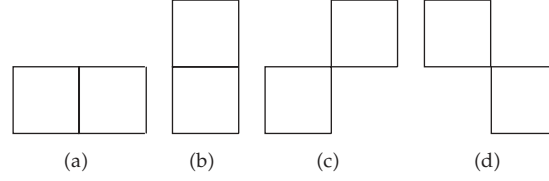
$\mathbf{N} = \{N_i \mid \forall i \in \mathbf{M}\}$  represents the neighbor system to indicate the interrelationship between nodes or the order of MRF. The edges are added between one node  $P_i$  and its neighbors  $N_i$ . Usually, the neighbor system should satisfy [2] the following:

- (1) a site does not neighbor with itself:  $i \notin N_i$ ,
- (2) the neighboring relationship is mutual:  $i \in N_j \Leftrightarrow j \in N_i$ .

The definition of the neighbor system is important because it reflects how far the contextual constraint is. For a regular array data, as in Figure 1, the neighbors of  $i$  are defined as the set of sites within a radius of  $\sqrt{r}$  from  $i$  where  $r$  is the order of the neighbor system.  $N_i = \{i, j \in \mathbf{M} \mid [f(\text{pixel}_i, \text{pixel}_j)]^2 \leq r, i \neq j\}$  where  $f(a, b)$  measures the Euclidean distance between  $a$  and  $b$ .

Another concept here is "clique." A clique is a subset of  $\mathbf{P}$  which plays the near role of the neighbor system. However, the nodes in a clique are ordered which means that  $(P_i, P_j)$  is a different form  $(P_j, P_i)$ . Figure 2 shows some examples of clique types.

Though we could get more static information of the problem domain with larger neighboring system, the computational complexity of the problem will also increase exponentially with the size of neighborhood. In most cases, a 4-neighborhood system is used for simplification and efficiency. MRF is an undirected graph where a set of random variables have a Markov property. In the random field, each random variable  $p_i$  in the set  $\mathbf{P}$  can take a label from  $\mathbf{L}$ . Usually, a mapping function  $\mathbf{F} = \{f_1, \dots, f_m\}$  in which  $\mathbf{F} : \mathbf{P} \rightarrow \mathbf{L}$



**Figure 2:** (a) Horizontal pair-site cliques, (b) vertical pair-site cliques, (c) and (d) diagonal pair-site cliques.

can represent for this processing.  $\mathbf{F}$  is also called configuration. Denote  $\Pr(f_i(p_i) = l_i)$  as the probability of a pixel  $p_i$  taking the label  $l_i$ . Then the configuration is a joint probability:  $\Pr(\mathbf{F}(\mathbf{P}) = \mathbf{L}) = \Pr(f_1(p_1) = l_1, \dots, f_m(p_m) = l_m)$ . Note that  $\Pr(f_i(p_i) = l_i) > 0$  for all  $p_i \in \mathbf{P}$ . The Markov property is a basic property if the conditional probability

$$\Pr(p_i | p_{M \setminus i}) = \Pr(p_i | p_{N_i}), \quad (2.1)$$

where  $M \setminus i$  means the entire element in  $\mathbf{M}$  other than  $i$  and  $N_i$  is the neighbor system of  $p_i$ . The Markov property means that the state probability of one node only depends on its neighbors rather than other remaining nodes. Gibbs random field is a random field in which the probability obeys the Gibbs distribution in the form of

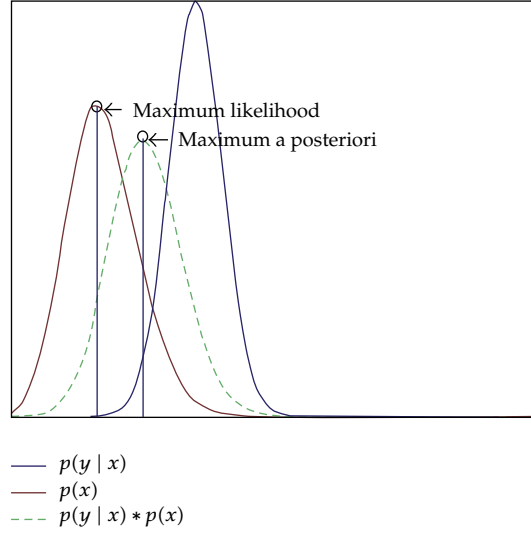
$$\Pr(f) = Z^{-1} \times e^{-(1/T)E(f)}, \quad (2.2)$$

where  $Z = \sum_{f \in F} e^{-(1/T)E(f)}$  is a normalizing constant called the partition function, and  $T$  is a constant which shall be assumed to be 1 unless otherwise stated.  $E(F) = \sum_{c \in C_i} V_c(f)$   $f \in c$ ,  $i \in M$  is the energy function.  $C$  is the clique defined on the graph, and  $V_c(f)$  is the clique potential function.

Hammersley-Clifford theorem states that if a probability distribution is a positive and satisfies the Markov properties in an undirected graph  $\mathbf{G}$ , the distribution is a Gibbs random field. That is, its probability can be factorized over the cliques of the graph. This theorem provides a simple way to calculate the joint probability using the clique potential. According to the Bayes' rule, the posterior distribution for a given set  $\mathbf{y}$  and their evidence  $p(\mathbf{y} | \mathbf{x})$ , combined with a prior  $p(\mathbf{x})$  over the unknowns  $\mathbf{x}$ , is given by

$$p(\mathbf{x} | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}. \quad (2.3)$$

If we do not know the prior information, the maximum likelihood (ML) criterion may be used where  $\mathbf{x}^* = \operatorname{argmax} p(\mathbf{y} | \mathbf{x})$ . However, sometimes we can still obtain the knowledge about the prior distribution of  $\mathbf{x}$ . Thus, the maximum of *a posteriori* (MAP) estimation is the best way to get the optimization where  $\mathbf{x}^* = \operatorname{argmax} p(\mathbf{y} | \mathbf{x})p(\mathbf{x})$ . Figure 3 illustrates the difference between ML criterion and MAP criterion. MAP probability is one of the most popular statistical criteria for optimization, and in fact, it is the first choice in MRF vision modeling.



**Figure 3:** Schematic comparison of ML and MAP methods.

Logarithmize both sides, and then we can obtain the negative posterior log-likelihood

$$-\log p(x | y) = -\log p(y | x) - \log p(x) + \log p(y), \quad (2.4)$$

where  $\log p(y)$  is a constant used to make integration of  $p(x | y)$  equal to 1. To find the MAP solution, we simply minimize (2.4). Rewrite the clique potential  $E(F) = \sum_{c \in C_i} V_c(f)$   $f \in c, i \in M$ , then (2.4) becomes an energy function

$$E(x, y) = E_d(x, y) + E_s(x), \quad (2.5)$$

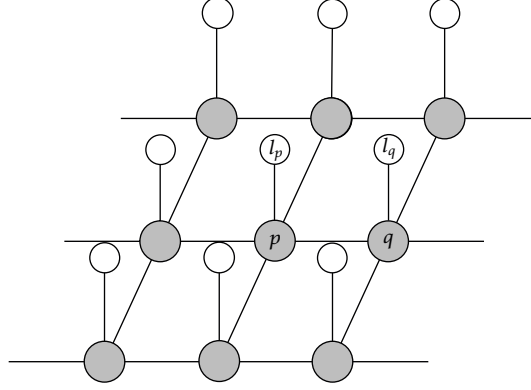
where the  $E_d(x, y)$  can be treated as the clique potential whose clique size is 1, and  $E_s(x)$  is the remaining clique potential or the observed image prior distribution. In most vision problems, the single-site clique potential is also called unary potential or data energy. Similarly,  $E_s(x)$  is called smooth potential or smooth energy. With (2.1), (2.5) can be rewritten as

$$E(f_p, N_p) = E_d(f_p, N_p) + E_s(f_p), \quad (2.6)$$

where  $E_s(f_p) = \sum_{q \in N_p \setminus q} E_s(f_p, f_q)$ . Therefore,

$$E(F) = \sum_{p \in P} \sum_{q \in N_p \setminus q} E_s(f_p, f_q) + \sum_{p \in P} E_d(f_p, N_p). \quad (2.7)$$

Most vision problems map the minimization of an energy function over an MRF. In some degree, the energy function can be seen as a mathematical representation of the scene and should precisely measure the global quantity of the solution as well as can be easy to find



**Figure 4:** The model for a 4-neighbor MRF (i.e., the pairwise MRF). The dash circles are the observed nodes, and the white circles are the unobserved labels.

the global minimization. When the energy function (2.7) is minimized, the corresponding posteriori  $\Pr(x | y)$  gets the maximum.

To solve a specific problem, we need to determine the energy form and the parameters involved. Though there are many types of clique potential functions, there exists a unique normalized potential, called the canonical potential. In literature, the energy function can be expressed as either a parametric form or a nonparametric form [2]. Here, we take the second-order clique potential, for example, which is also called the pairwise model. The pairwise MRF is the most commonly used model in which each node interacts with its adjacent nodes. It is the lowest-order constraint to convey contextual information and is widely used due to its simple form and low computational cost [2]. The pairwise MRF models the statistics of the first derivative in the image structure (Figure 4). The corresponding energy function is

$$E(F) = \sum_{p \in P} \sum_{q \in N_p \setminus q} E_s(f_p, f_q) + \sum_{p \in P} E_d(f_p, N_p) \cdots N = 4. \quad (2.8)$$

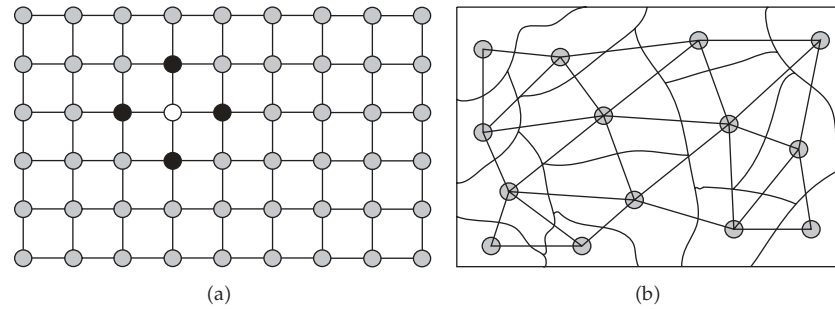
Usually,  $E_d$  is the local evidence of  $p$  taking the label  $f_p$  such as the intensity or the color value. Equation (2.8) can be rewritten as

$$E(F) = \sum_{p \in P} \sum_{q \in N_p \setminus q} E_s(f_p, f_q) + \sum_{p \in P} E_d(f_p) \cdots N = 4. \quad (2.9)$$

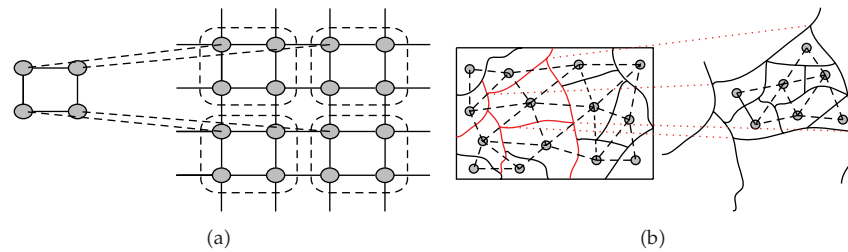
In the binary MRF case,  $E_s(f_p, f_q) = f_p \times f_q$ , where  $f_p \in \{0, 1\}$ . In the multilabel case, Potts model is the most widely used one which can prevent the edges of objects from oversmoothing. Usually, Potts model takes the form

$$V(f_p, f_q) = \begin{cases} 0 \cdots f_p = f_q, \\ \alpha \cdots f_p \neq f_q, \end{cases} \quad (2.10)$$

where  $\alpha$  may be a constant or  $\alpha = |f_p - f_q|$ .



**Figure 5:** (a) Standard pairwise MRF model with image-grid data, where the circle represents a node or a pixel. The black circles are neighbors of the white one; (b) an example of MRF used in segmentation, where the nodes of neighboring segments are connected by applying Delaunay triangulation method.

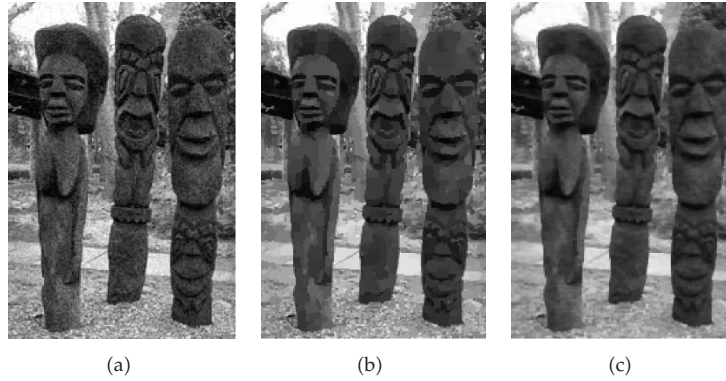


**Figure 6:** Two examples of hierarchical MRFs. (a) is the most common quadtree where four-grid nodes on the low level are decomposed into one node on the high level. (b) is an irregular pyramid with a large-scale MRF and a corresponding small-scale MRF.

As is illustrated in Figure 5, in the pairwise MRF model, a node is attached to a pixel in the image, while edges are constructed between the node and its four neighborhoods. With such model, the corresponding energy function can be efficiently minimized using many inference algorithms. Other graph structures are also used. For example, in image segmentation, an image is partitioned into several regions. Each region can be regarded as a node, and edges may be constructed between adjacent segmented regions. To make the optimization more efficient, a hierarchical MRF model is used. It mainly uses the pyramid structure and performs in a coarse-to-fine scheme which uses a coarser solution to initialize a finer solution. It is well known that hierarchical methods can significantly improve the convergence rate and reduce the execution time. In [29–31], a regular pyramid downsampling method was applied, while Zitnick and Kang [32] used an irregular pyramid. Figure 6 illustrates an example.

Although most MRFs use the pairwise model due to its simplicity, a scheme of more complex interaction, for example, 8-neighborhood or more numbers of pairwise terms, is also used sometimes. People usually use 26-neighborhood in 3D volumetric images or video analysis. Higher-order clique potentials can capture more complex interactions of random variables. For example, calculating the curvature of an object requires interaction of at least three nodes. Computational time for the clique potential increases exponentially with the size of the clique and poses a difficult energy minimization scenario, which poses a tough question. Recently, there have been many attempts to go beyond pairwise MRF. One approach is to transform the higher-order problem into pairwise problem by adding auxiliary variables.





**Figure 7:** An example of image denoising. (a) is the original noisy image. (c) is the denoising result by BP using pairwise interaction [5]. (b) is the result by BP with the learned 2-by-2 model. The results no longer exhibit any piecewise constancy. Not only edges are preserved, but also smoothly varying regions are preserved better than those with higher-order clique [39].

For instance, Kohli et al. [33] proposed an efficient graph cut method based on special class of higher-order potential, that is, robust  $P^n$  Potts model. Rother et al. [34] transformed the minimizing sparse higher-order energy function into an equivalent quadratic minimization problem. Potetz and Lee [35] introduced an efficient belief propagation algorithm where the computational complexity increases linearly with the clique size. Kwon et al. [36] decomposed high-order cliques as hierarchical auxiliary nodes and used hierarchical gradient nodes to reduce the computational complexity. Another way is to perform direct computing using factor graph representation [37]. Kwon et al. [38] proposed a nonrigid registration method using the MRF with a higher-order spatial prior. Experiments show that using high-order potential the performances of image denoising are significantly improved, as is shown in Figure 7.

### 3. Inference Methods

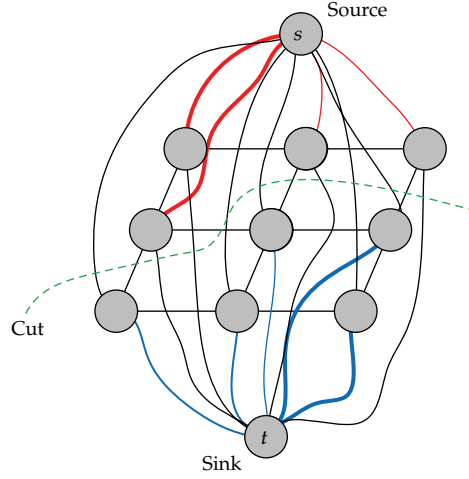
Over the years, a large number of inference algorithms have been developed, which can be mainly classified into two categories, that is, message passing algorithms such as loopy belief propagation and move making algorithms such as graph cuts. In this section, we briefly introduce two classic inference methods for approximating energy minimums, that is, belief propagation and graph cut.

#### 3.1. Graph Cut

Graph cut (GC) was first applied in computer vision by Greig et al. [40], which describes a large family of MRF inference algorithms based on solving min-cut/max-flow problems. Given a type of computer vision problems which can be formulated in terms of an energy function, GC can get the minimum energy configuration corresponds to the MAP theory.

Suppose that  $G = (V, E)$  is a directed graph in which the edge weight is nonnegative,  $V$  represents vertices, and  $E$  denotes edges. The graph has two special terminals (vertices), that is, the source  $s$  and the sink  $t$ . A cut  $C = (S, T)$  is a partition of  $V$ . An  $s$ - $t$  cut is a cut that splits





**Figure 8:** An example of min-cut/max-flow graph cut. The gray circles represent the nodes, and the solid lines are the edges between the nodes. The curve of each “flow” is connected to the source terminal or sink terminal. The potential of flow is measured by the width of line. The dotted line indicates a cut of graph partition.

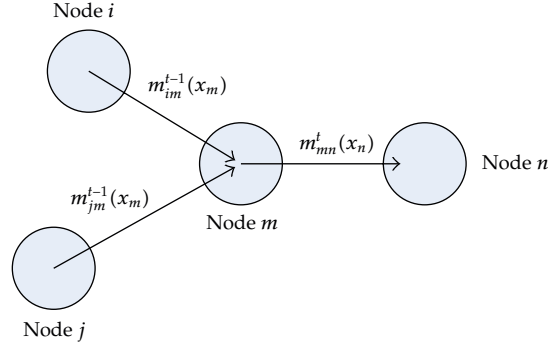
the source and the sink to be in different subsets where  $s \in S$  and  $t \in T$ . Besides, according to graph theory, the potential of a cut can be measured by the sum of the weights of the edge crossing the cut. To find a cut which can minimize  $s$ - $t$  cut problem is equivalent to compute the maximum flow from the source  $s$  to the sink  $t$ . Maximum flow is the maximum “amount of water” that can be sent from the source to the sink by interpreting graph edges as directed “pipes” with capacities equal to edge weights. As illustrated in Figure 8, the GC algorithm is ideally designed to solve the max-flow problem.

It was reported that GC can obtain the exact solution in the binary label case. In multilabel case, GC requires solving a series of related binary inferences and then obtains the approximated global optimal solutions. Two of the most popular GC algorithms are  $\alpha$ - $\beta$  swap and  $\alpha$ -expansion. In the  $\alpha$ - $\beta$  algorithm, a swap move takes some subset of nodes that currently label with  $\alpha$  and assign their label with  $\beta$ , and vice versa. The  $\alpha$ -expansion algorithm increases the set of nodes taking  $\alpha$  by moving it to other nodes. When there is no more swap or expansion move, a local minimum is found. Comparing the two algorithms,  $\alpha$ -expansion is more accurate and efficient. Also  $\alpha$ -expansion can produce a result with lower energy. However, the condition of  $\alpha$ -expansion is more strict. When using the  $\alpha$ -expansion, the interaction potential must be metric, that is,

$$V_{ij}(\alpha, \alpha) + V_{ij}(\beta, \beta) \leq V_{ij}(\alpha, \beta) + V_{ij}(\beta, \alpha). \quad (3.1)$$

For  $\alpha$ - $\beta$  swap, it must be semimetric, that is,

$$V_{ij}(\alpha, \alpha) + V_{ij}(\beta, \beta) \leq V_{ij}(\alpha, \beta) + V_{ij}(\beta, \alpha). \quad (3.2)$$



**Figure 9:** Message passing in BP.  $m_{ij}(x_i)^t$  is a message from node  $i$  to its neighbor  $j$  and indicates what the state should be in node  $j$ .

More details about  $\alpha$ - $\beta$  swap and  $\alpha$ -expansion can be found in [4]. In addition, Kolmogorov and Roth [41] wrote a survey about graph cut and pointed out that GC can be applied to both submodular and nonsubmodular functions. Other more recent developments in GC include order-preserving GC [42] and combination GC [43].

### 3.2. Belief Propagation

Belief propagation is a power inference tool originally developed for tree-Bayesian networks [45]. It is recently extended to those “cycle” graphs such as MRF. Although BP can only guarantee convergence with the Bethe free energy in MRF [46], it can obtain reasonable results in practice. In standard BP with pairwise MRF, a variable  $m_{ij}(x_j)$  can be treated as a “message” from a node  $i$  to its neighbor  $j$  which contains the information about what the state of node  $j$  should be in. The message is a vector of the same dimension as the number of possible labels. The value of each dimension manifests how this label might be corresponding to the node.

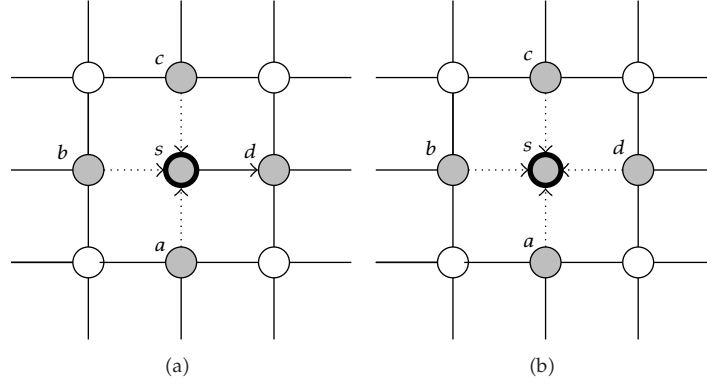
Let  $\phi_{ij}(x_i, x_j)$  be the pairwise interaction potential of  $p_i$  with  $p_j$ , and  $\phi_i(x_i, x_j)$  is the “local evidence” of  $p_i$ . Usually, the message must be nonnegative. A large value of the message means that the node “believes” the posterior probability of  $X_j$  is high. The message updating rule is

$$m_{ij}(x_i)^t = \sum_{x_i} \varphi_i(x_i) \phi_{ij}(x_i, x_j) \prod_{k \in N_i \setminus j} m_{ki}(x_i)^{t-1}, \quad (3.3)$$

where  $t$  represents the number of interaction  $T$  as showed in Figure 9.

The belief is the product of “local evidence” of the node and all messages send to it

$$b_i(x_i) = k \phi_i(x_i) \sum_{j \in N_i} m_{ij}(x_i). \quad (3.4)$$



**Figure 10:** (a) Messages passing from node  $s$  to its neighbor  $u$ . (b) The belief of node  $s$  is calculated according to its neighbors' messages.

The standard BP described above is also called *sum-product* BP. There is another variant BP which is more simple to use, that is, *max-product* (or *max-sum* in log domain). In *max-product* BP, (3.3) and (3.4) are represented as

$$m_{ij}(x_i)^t = \max_{x_i} \left( \varphi_i(x_i) + \phi_{ij}(x_i, x_j) + \sum_{k \in N_i \setminus j} m_{ki}(x_i)^{t-1} \right), \quad (3.5)$$

$$b_i(x_i) = k \left( \sum_{j \in N_i} m_{ij}(x_i) + \varphi_i(x_i) \right). \quad (3.6)$$

The sketch map of this process is illustrated in Figure 10. Several speed-up techniques are attempted, for example, distance transformation, checkerboard updating, and multiscale BP [5], so that the belief propagation can converge efficiently. In another way, Yu et al. [47] used the predictive coding, linear transform coding, and envelope point transform to improve the BP efficiency.

Although BP is an implicitly efficient inference algorithm for MRF with loops, it can only converge to the stationary point of the Bethe approximation of the free energy. Recently, a generalized belief propagation (GBP) algorithm proposed by [48] has received more attention due to its better convergence property against BP. It can converge to a more accurate stationary point of Kikuchi free energy [46]. More details about the GBP algorithm can be found in [48].

BP and graph cut are both good optimal techniques which can find "global" minima over cliques and produce plausible results in practice. A comparison between the two different approaches for stereo vision was described in [49]. GC can get lower energy, but the performance of BP is comparative to GC relative to the ground truth.

In addition to the two typical methods, many other inference algorithms have been proposed in latest few years. Fusion move [8] is proposed for multilabel MRF. By employing QPBO graph cut, the fusion move can efficiently combine two proposal labels in a theoretically sound way, which is in practice often globally optimal. Alahari et al. [50] improved the computational and memory efficiency of algorithms for solving multilabel

energy functions arising from discrete MRF by recycling, reducing, and reusing. Kumar et al. [51] provided an analysis of linear programming relaxation, the quadratic programming relaxation, and the second-order cone programming relaxation to obtain the maximum a posteriori estimate of a general discrete MRF. Komodakis and Tziritas [52] proposed an exemplar-based framework and used priority BP to find MRF solutions. Ishikawa [53] introduced a method to exactly solve a first-order MRF optimization problem in more generality than previous ones. Cho et al. [54] used patch transform representation to manipulate images in the patch domain. The patch transform is posed as a patch assignment problem on an MRF, where each patch should be used only once, and neighboring patches should fit to form a plausible image.

## 4. Applications

Here, we provide MRF solutions for two typical problems in computer vision, that is, stereo matching and image segmentation. These problems require labeling each pixel with a value to represent the disparity and foreground or background. They can be easily modeled using MRF and solved by energy minimization.

### 4.1. Stereo Matching

Stereo matching has always been one of the most challenging and fundamental problems in computer vision. Comprehensive research has been done in the last decade [32, 55–58]. A latest evaluation of these various methods can be found in [59]. In the last few years, as is shown in [44], the global methods based on MRF have reached the top performance.

For MAP estimation, let  $\mathbf{P}$  be the set of the image pixels in image pair, and let  $\mathbf{L}$  be the set of disparity. The initial data cost, which is calculated by the truncated linear transform which is robust to noise or outlier, is defined as

$$D(f_p) = \lambda \cdot \min \left( \sqrt{\sum_{c \in \{L, a, b\}} (\mathbf{I}_c^L(p) - \mathbf{I}_c^R(p - f_p))^2}, T \right), \quad (4.1)$$

where  $\lambda$  is the cost weight which determines the portion of energy that data cost possesses in the whole energy, and  $T$  represents the truncating value. The parameters can be set with empirical values from experiments.  $\mathbf{I}_c^L(p)$  represents  $p$ 's intensity in the left image of channel  $c$ .  $\mathbf{I}_c^R(p)$  is similarly defined. Birchfield and Tomasi's pixel dissimilarity is used to improve the robustness against the image sampling noise. The smooth cost which expresses the compatibility between neighboring variables embedded in the truncated linear model, is defined as:

$$V(f_p, f_q) = \min(|f_p - f_q|, K), \quad (4.2)$$

where  $K$  is the truncating value. The smooth cost based on the truncated linear model is also referred to as discontinuity preserving cost since it can prevent the edges of objects from



**Figure 11:** Comparison of different optimization algorithms. (a) image is the original left image. (b) is the result BP. (c) is the result by  $\alpha$ -expansion. (d) is the result TRW [44].

oversmoothing. The corresponding energy function used here is the most conspicuous one and is defined as

$$E(f) = \sum_{p \in \mathbf{P}} D(f_p) + \sum_{(p,q) \in \mathbf{N}} V(f_p, f_q), \quad (4.3)$$

where  $\mathbf{N}$  contains the edges in the four-connected neighborhood set.

The objective is to find a solution which minimizes (4.3). The solution means the correct depth information in the scene. Figure 11 shows the results of “Tsukuba” data set using different energy minimization methods available in [44]. In the past decades, segment-based stereos [32] have been boomed as they perform well in reducing the ambiguity associated with textureless regions and enhancing noise tolerance by aggregating over pixels with homogenous properties. Usually, those algorithms firstly segment the source image. Then the matching cost is computed over the entire segment. A plane fit method is applied to refine the result.

## 4.2. Binary Image Segmentation

Binary image segmentation is widely used in medical image analysis and object recognition. Here, each pixel is assigned with a label  $l$  with  $0 \leq l \leq 1$ . In the simplest case, we have  $l \in \{0, 1\}$ , where 0 represents the pixel belonging to the background and 1 to the foreground. The segmentation result should be accurate and fine enough for successful applications such as object category, photo editing, and image retrieval. Although segmentation is regarded as one of the most difficult problems due to the complexity of real scene and noise corruption, MRF model can often successfully deal with this challenging problem.

The corresponding energy function is represented the same as (2.9). The data cost  $E_d$  represents whether the pixel property is consistent with the statistic distribute of possible region. It may be simple to take such an absolute difference of pixel intensity and the mean of region gray level. Alternatively, the complex data term often leads to better results. For example, in [60], the data cost uses the color data model which is the log-likelihood of a pixel and is modeled as two separate Gaussian mixture models. The smoothness term is a simple Potts model

$$E_s(f_q, f_p) = K \times \text{Dis}(m, n)^{-1} [f_q = f_p]. \quad (4.4)$$



**Figure 12:** Comparison of different optimization algorithms for binary segmentation of the image “sponge.” (a) image is original. (b) is the result of BP. (c) is the result of  $\alpha$ -expansion. (d) is the result of TRW. There is a slight difference among the three results. Besides, since this is a binary labeling problem, the  $\alpha$ -expansion finds the global optimum in a single iteration [44].

In (4.4),  $\text{Dis}(m, n)$  is the Euclidean distance of pixel  $m$  and pixel  $n$ , and  $[f_p = f_q]$  denotes the indicator function taking 0 and 1.  $K$  is a constant. If  $K = 1$ , the smoothness term recovers the Ising model which encourages smoothness everywhere.  $K$  determines how coherent the similar grey level in a region is. Recently, user interaction was proposed to refine the results in [60–62]. Usually, the user first marks some pixels to indicate the background and foreground. With those labeled pixels, we can get the corresponding region statistics.

GC is the most common optimal tool for binary MRF combined with both color (texture) information and edge information. Further, the marked pixels can be used as the seeds in the cut-based algorithm. A graph cut extension, that is, grabcut [60], was proposed for iterative minimization of the energy. Figure 12 shows the results of binary segmentation with different methods using identity parameters [44].

Considering multilabel segmentation, Micusik and Pajdla [63] formulated single-image multi-label segmentation into coherent regions in texture and color as a max-sum problem. As a region merging method, Mignotte [64] used MRF fusion model combining several segmentation results to achieve a more reliable and accurate result.

More recently, Panda and Nanda [65] proposed an unsupervised color image segmentation scheme using the homotopy continuation method and compound MRF model. Chen et al. [66] proposed image segmentation method based on MAP or ML estimation. Li [67] introduced a multiresolution MRF approach to texture segmentation problems. Rivera et al. [68] presented a new MRF model for parametric image segmentation. Some other works [69–71] carried out for learning of the prior distribution. MRF is also widely used in medical image segmentation. Zhang et al. [72] proposed segmentation of brain MR images through a hidden MRF. Scherrer et al. [73] used expectation maximization to segment the images in an MRF model. Anguelov et al. [74] segment 3D scanned data into objects using GC. Hower et al. [75] investigated in the context of neuroimaging segmentation. As a low level vision problem, the segment is often applied for object classification. Honghui et al. [76] proposed a robust supervised label transfer method for semantic segmentation of street scenes. Feng et al. [77] recently proposed a method to optimize the MRF, which can automatically determine the number of labels in balance of accuracy and efficiency.

## 5. Conclusion

It is now acknowledged that MRF is one of the most successful approaches for solving labeling problems in computer vision and image analysis. The most challenge of MRF models is to develop its efficient inference algorithm in order to find the low-energy configuration.

As in computer vision, there are too many nodes. For example, consider two frame images with the size of  $a \times b$ . If each node takes  $N$  possible labels, the computation space is  $N^{a \times b}$ . Clearly, the inference algorithm should be efficient enough to overcome this dilemma. Secondly, constructing reasonable MRF also plays key roles, especially for some new vision applications. For instance, there are many different grid topologies and nonlocal topologies. Thirdly, the parameters of MRF model should be efficiently learned from image instead of manually or empirically chosen. Furthermore, further studies can focus on the energy functions which can not be efficiently solved by using state-of-the-art methods.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China and Microsoft Research Asia (NSFC-60870002, 60802087), NCET, Zhejiang Provincial S&T Department (2010R10006, 2010C33095), and Zhejiang Provincial Natural Science Foundation (R1110679).

## References

- [1] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
- [2] S. Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer, Berlin, Germany, 2009.
- [3] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, Berlin, Germany, 1st edition, 2011.
- [4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [5] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," *International Journal of Computer Vision*, vol. 70, no. 1, pp. 41–54, 2006.
- [6] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1568–1583, 2006.
- [7] N. Komodakis, N. Paragios, and G. Tziritas, "MRF energy minimization and beyond via dual decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 531–552, 2011.
- [8] V. Lempitsky, C. Rother, S. Roth, and A. Blake, "Fusion moves for markov random field optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1392–1405, 2010.
- [9] R. Szeliski, R. Zabih, D. Scharstein et al., "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068–1080, 2008.
- [10] P. F. Felzenszwalb and R. Zabih, "Dynamic programming and graph algorithms in computer vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 721–740, 2011.
- [11] T. Werner, "A linear programming approach to max-sum problem: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1165–1179, 2007.
- [12] S. Roth and M. J. Black, "Fields of experts: a framework for learning image priors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '05)*, vol. 2, pp. 860–867, San Diego, Calif, USA, June 2005.
- [13] U. Schmidt, Q. Gao, and S. Roth, "A generative perspective on MRFs in low-level vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '10)*, pp. 1751–1758, San Francisco, Calif, USA, June 2010.
- [14] O. Woodford, C. Rother, and V. Kolmogorov, "A global perspective on MAP inference for low-level vision," in *Proceedings of the 12th International Conference on Computer Vision, (ICCV '09)*, pp. 2319–2326, October 2009.
- [15] S. Y. Lin and J. Y. Shi, "A Markov random field model-based approach to natural image matting," *Journal of Computer Science and Technology*, vol. 22, no. 1, pp. 161–167, 2007.
- [16] T. Kasetkasem, M. K. Arora, and P. K. Varshney, "Super-resolution land cover mapping using a Markov random field based approach," *Remote Sensing of Environment*, vol. 96, no. 3-4, pp. 302–314, 2005.



- [17] H. Heng-Feng, L. Jin-Jang, L. Cheng-Shian, and C. Hsuan-Ying, "Image inpainting using structure-guided priority belief propagation and label transformations," in *Proceeding of the 20th International Conference on Pattern Recognition, (ICPR '10)*, pp. 4492–4495, August 2010.
- [18] D. Zikic, B. Glocker, O. Kutter et al., "Markov random field optimization for intensity-based 2D-3D registration," in *Medical Imaging 2010: Image Processing*, vol. 7623 of *Proceedings of SPIE*, 2010.
- [19] J. Verbeek and B. Triggs, "Region classification with Markov field aspect models," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '07)*, pp. 1059–1066, June 2007.
- [20] D. Larlus and F. Jurie, "Combining appearance models and Markov random fields for category level object segmentation," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, (CVPR '08)*, vol. 1–20, pp. 864–870, June 2008.
- [21] M. Lievin and F. Luthon, "Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of face features in video," *IEEE Transactions on Image Processing*, vol. 13, no. 1, pp. 63–71, 2004.
- [22] X. G. Wang and X. O. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, 2009.
- [23] S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, "Improving offline handwritten text recognition with hybrid HMM/ANN models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 767–779, 2011.
- [24] M. L. Song, D. C. Tao, C. Chen, X. L. Li, and C. W. Chen, "Color to gray: visual cue preservation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1537–1552, 2010.
- [25] H. Mirzaalian and G. Hamarneh, "Vessel scale-selection using MRF optimization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '10)*, pp. 3273–3279, June 2010.
- [26] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
- [27] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo matching using belief propagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787–800, 2003.
- [28] R. Detry, N. Pugeault, and J. H. Piater, "A probabilistic framework for 3D visual object representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 10, pp. 1790–1803, 2009.
- [29] A. S. Willsky, "Multiresolution Markov models for signal and image processing," *Proceedings of the IEEE*, vol. 90, no. 8, pp. 1396–1458, 2002.
- [30] Z. Kato, M. Berthod, and J. Zerubia, "Multiscale Markov random field models for parallel image classification," in *Proceedings of the IEEE 4th International Conference on Computer Vision*, pp. 253–257, May 1993.
- [31] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical MRF model," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216–1231, 2000.
- [32] C. L. Zitnick and S. B. Kang, "Stereo for image-based rendering using image over-segmentation," *International Journal of Computer Vision*, vol. 75, no. 1, pp. 49–65, 2007.
- [33] P. Kohli, L. Ladicky, and P. H. S. Torr, "Robust higher order potentials for enforcing label consistency," *International Journal of Computer Vision*, vol. 82, no. 3, pp. 302–324, 2009.
- [34] C. Rother, P. Kohli, W. Feng, and J. Y. Jia, "Minimizing sparse higher order energy functions of discrete variables," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, (CVPR Workshops '09)*, vol. 1–4, pp. 1382–1389, June 2009.
- [35] B. Potetz and T. S. Lee, "Efficient belief propagation for higher-order cliques using linear constraint nodes," *Computer Vision and Image Understanding*, vol. 112, no. 1, pp. 39–54, 2008.
- [36] D. Kwon, K. J. Lee, I. D. Yun, and S. U. Lee, "Solving MRFs with higher-order smoothness priors using hierarchical gradient nodes," in *Proceedings of the 10th Asian Conference on Computer Vision*, pp. 121–134, November 2010.
- [37] D. Tarlow, I. Givoni, and R. S. Zemel, "HOP-MAP: efficient message passing with high order potentials," *Journal of Machine Learning Research*, vol. 9, pp. 812–819, 2010.
- [38] D. Kwon, K. J. Lee, I. D. Yun, and S. U. Lee, "Nonrigid image registration using dynamic higher-order MRF model," in *Proceedings of the 10th European Conference on Computer Vision, (ECCV '08)*, vol. 5302, part 1 of *Lecture Notes in Computer Science*, pp. 373–386, Marseille, France, October 2008.
- [39] X. Y. Lan, S. Roth, D. Huttenlocher, and M. J. Black, "Efficient belief propagation with learned higher-order Markov random fields," in *Proceedings of the 9th European Conference on Computer Vision (ECCV '06)*, vol. 3952, part 2, pp. 269–282, Graz, Austria, May 2006.

- [40] D. M. Greig, B. T. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society Series B Methodological*, vol. 51, no. 2, pp. 271–279, 1989.
- [41] V. Kolmogorov and C. Rother, "Minimizing nonsubmodular functions with graph cuts—a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1274–1279, 2007.
- [42] X. Liu, O. Veksler, and J. Samarabandu, "Order-preserving moves for graph-cut-based optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, pp. 1182–1196, 2010.
- [43] N. Komodakis and G. Tziritas, "Approximate labeling via graph cuts based on linear programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1436–1453, 2007.
- [44] D. Scharstein and R. Szeliski, "Middlebury Stereo Vision Research Page," 2011, <http://vision.middlebury.edu/stereo/eval>.
- [45] J. Pearl, "Reverend bayes on inference engines: a distributed hierarchical approach," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 133–136, Pittsburgh, Pa, USA, 1982.
- [46] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.
- [47] T. L. Yu, R. S. Lin, B. Super, and B. Tang, "Efficient message representations for belief propagation," in *Proceedings of the 11th IEEE International Conference on Computer Vision, (ICCV '07)*, vol. 1–6, pp. 604–611, October 2007.
- [48] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Generalized belief propagation," *Advances in Neural Information Processing Systems*, vol. 13, pp. 689–695, 2001.
- [49] M. F. Tappen and W. T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters," in *Proceedings of the 9th IEEE International Conference on Computer Vision*, vol. 2, pp. 900–907, October 2003.
- [50] K. Alahari, P. Kohli, and P. H. S. Torr, "Dynamic hybrid algorithms for MAP inference in discrete MRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1846–1857, 2010.
- [51] M. P. Kumar, V. Kolmogorov, and P. H. S. Torr, "An analysis of convex relaxations for MAP estimation of discrete MRFs," *Journal of Machine Learning Research*, vol. 10, pp. 71–106, 2009.
- [52] N. Komodakis and G. Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2649–2661, 2007.
- [53] H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1333–1336, 2003.
- [54] T. S. Cho, S. Avidan, and W. T. Freeman, "The patch transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1489–1501, 2010.
- [55] Q. X. Yang, L. Wang, R. G. Yang, H. Stewenius, and D. Nister, "Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 3, pp. 492–504, 2009.
- [56] A. Barbu and S. C. Zhu, "Incorporating visual knowledge representation in stereo reconstruction," in *Proceedings of the 10th IEEE International Conference on Computer Vision, (ICCV '05)*, vol. 1-2, pp. 572–579, Beijing, China, October 2005.
- [57] N. D. F. Campbell, G. Vogiatzis, C. Hernandez, and R. Cipolla, "Using multiple hypotheses to improve depth-maps for multi-view stereo," in *Proceedings of the 10th European Conference on Computer Vision, (ECCV '08)*, vol. 5302, part 1, pp. 766–779, 2008.
- [58] J. Starck and A. Hilton, "Correspondence labelling for wide-timeframe free-form surface matching," in *Proceedings of the 11th IEEE International Conference on Computer Vision, (ICCV '07)*, vol. 1–6, pp. 2189–2196, October 2007.
- [59] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp. 7–42, 2002.
- [60] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut' - interactive foreground extraction using iterated graph cuts," *Acm Transactions on Graphics*, vol. 23, no. 3, pp. 309–314, 2004.
- [61] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of the 8th International Conference on Computer Vision*, vol. 1, pp. 105–112, July 2001.
- [62] I. Grinias, N. Komodakis, and G. Tziritas, "Flooding and MRF-based algorithms for interactive segmentation," in *Proceedings of the 20th International Conference on Pattern Recognition, (ICPR '10)*, pp. 3943–3946, August 2010.

- [63] B. Micusik and T. Pajdla, "Multi-label image segmentation via max-sum solver," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1–8, pp. 1994–1999, 2007.
- [64] M. Mignotte, "A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1610–1624, 2010.
- [65] S. Panda and P. K. Nanda, "Unsupervised color image segmentation using compound Markov random field model," in *Proceedings of the 3rd International Conference on Pattern Recognition and Machine Intelligence*, vol. 5909, pp. 291–296, December 2009.
- [66] S. F. Chen, L. L. Cao, J. Z. Liu, and X. O. Tang, "Iterative MAP and ML estimations for image segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '07)*, vol. 1–8, pp. 289–294, June 2007.
- [67] C. T. Li, "A conditional random field approach to unsupervised texture image segmentation," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, 2010.
- [68] M. Rivera, O. Ocegueda, and J. L. Marroquin, "Entropy-controlled quadratic Markov measure field models for efficient image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 3047–3057, 2007.
- [69] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4085–4098, 2010.
- [70] X. Feng, C. K. I. Williams, and S. N. Felderhof, "Combining belief networks and neural networks for scene segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 467–483, 2002.
- [71] X. Jianxiong and Q. Long, "Multiple view semantic segmentation for street view images," in *Proceedings of the 12th International Conference on Computer Vision, (ICCV '09)*, pp. 686–693, October 2009.
- [72] Y. Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.
- [73] B. Scherrer, M. Dojat, F. Forbes, and C. Garbay, "Agentification of Markov model-based segmentation: application to magnetic resonance brain scans," *Artificial Intelligence in Medicine*, vol. 46, no. 1, pp. 81–95, 2009.
- [74] D. Anguelov, B. Taskar, V. Chatalbashev et al., "Discriminative learning of Markov random fields for segmentation of 3D scan data," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR '05)*, vol. 2, pp. 169–176, June 2005.
- [75] D. Hoyer, V. Singh, and S. Johnson, "Label set perturbation for MRF based neuroimaging segmentation," in *Proceedings of the 12th International Conference on Computer Vision, (ICCV '09)*, pp. 849–856, October 2009.
- [76] Z. Honghui, X. Jianxiong, and Q. Long, "Supervised label transfer for semantic segmentation of street scenes," in *Proceedings of the 11th European Conference on Computer Vision, (ECCV '10)*, vol. 6315, pp. 561–574, September 2010.
- [77] W. Feng, J. Y. Jia, and Z. Q. Liu, "Self-validated labeling of Markov random fields for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1871–1887, 2010.





# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

