Optik 124 (2013) 5732–5738

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Image segmentation using adaptive loopy belief propagation

Sheng-Jun Xu [a,b,*], Jiu-Qiang Han [a], Jun-Qi Yu [b], Liang Zhao [b]

[a] MoE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China
[b] School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China

ARTICLE INFO

Article history:
Received 10 November 2012
Accepted 4 April 2013

Keywords:
Adaptive loopy belief propagation
Markov random fields
Local interaction region MRF
Message self-convergence
Label pruning

ABSTRACT

Loopy belief propagation (LBP) algorithm over pairwise-connected Markov random fields (MRFs) has become widely used for low-level vision problems. However, Pairwise MRF is often insufficient to capture the statistics of natural images well, and LBP is still extremely slow for application on an MRF with large discrete label space. To solve these problems, the present study proposes a new segmentation algorithm based on adaptive LBP. The proposed algorithm utilizes local region information to construct a local region model, as well as a local interaction region MRF model for image segmentation. The adaptive LBP algorithm maximizes the global probability of the proposed MRF model, which employs two very important strategies, namely, "message self-convergence" and "adaptive label pruning". Message self-convergence can improve the reliability of a pixel in choosing a label in local region, and label pruning can dismiss impossible labels for every pixel. Thus, the most reliable information messages transfer through the LBP algorithm. The experimental results show that the proposed algorithm not only obtains more accurate segmentation results but also greater speed.

1. Introduction

The application of the Markov random fields (MRFs) model for solving early vision problems has recently resulted in exciting advances. The MRF model describes the probabilistic relationship between an observed image and an estimated image or scene pixels. However, the statistical inference in the MRF model remains NP-Hard [1]. Good approximation techniques based on loopy belief propagation (LBP) [2] and graph cuts [3,4] have been developed and demonstrated for a number of computer vision problems [3,5–7]. However, two problems need to be considered. The first problem is that conventional Pairwise MRF lacks the representational power required to capture the rich statistics of the natural image [8]. Although several higher-order learning algorithms have been proposed, including [9–11], the learning of such algorithms still presents hard problems [12]. Based on Pairwise MRF, Wang [13] proposed a super-pixel MRF for incorporating local data interaction. On the other hand, Jia [14] showed the over-segmentation of an image by a mean shift algorithm and represented it using a region adjacent graph (RAG). LBP was then run over the RAG for image segmentation. However, the non-overlapping region often produced blocky or staircase artifacts between neighboring regions. The present paper proposes an overlapping RAG for solving this problem. As such, the segmentation problem can be converted to a locally aggregated global optimization problem.

Another problem is the inefficiency of the LBP algorithm in terms of probability inference for the maximum a posteriori (MAP) estimation of the MRF model with a large discrete label space. The computational complexity of the LBP is quadratically proportional to the number of labels, which obscures the application of this algorithm in a number of computer vision problems. A simple strategy in maintaining the tractability of inference is the reduction of the number of possible states for each node in the MRF. A number of studies have recently developed several efficient strategies. Chan [15] proposed a local belief aggregation (LBA) algorithm that restricts the number of messages aggregated from a neighboring node. However, the LBA may sometimes incorrectly discard segment states, and this algorithm remains very slow, thus hindering its practical application on large-scale graphs. Scott et al. [16] used a hierarchical scheme, in which "rough" results computed at higher levels initialize the search space at lower levels and enable the message computations through the use of a smaller search space compared with traditional belief propagation. However, this hierarchical BP may also lose states in lower levels. Lan et al. [17] proposed an approximation BP algorithm that used an adaptive state space method to reduce the number of states in each pixel in higher-order MRF. The experimental results showed that the learned higher-order model outperforms the learned Pairwise MRF model both visually and quantitatively. However, the learning of the Fields of Experts models, which are exploited for learning

* Corresponding author at: MoE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China. Tel.: +86 189 92816228; fax: +86 029 82207827.
E-mail address: duplin@sina.com (S.-J. Xu).

0030-4026/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.ijleo.2013.04.013

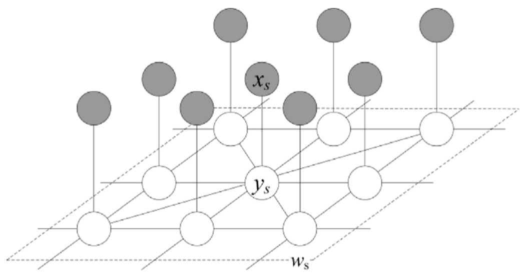

Fig. 1. Local region MRF model.

higher-order MRF, presents a hard problem. Similar to the aforementioned approaches, this method still has problems in terms of losing states.

The present study builds an overlapping local region MRF model based on Pairwise MRF, thereby incorporating more context information into the local region. Furthermore, the current study proposes an adaptive LBP as inference in the MAP probability based on the proposed local region MRF. The proposed algorithm utilizes two optimization schemes, namely, "message self-convergence" and "adaptive label pruning." Based on these two schemes, during the iteration processing, the probability of choosing the label assigned to more pixels in a local region rapidly increases while all others rapidly decrease. This step discards a few labels with minimal probability. The final step, according to the MAP criterion, obtains the segmentation results. The experimental results show the effectiveness of the proposed algorithm.

The remainder of the present paper is as follows: the next section presents the proposed MRF segmentation model. Section 3 describes the details of the adaptive LBP algorithm. Section 4 discusses the experiments and results. Finally, Section 5 concludes the current paper and states the final remarks.

2. Local region MRF

An MRF image is defined on a $W \times H$ rectangular lattice. Let $S=\{s=(i,j)|1 \le i \le W,\ 1 \le j \le H\}$ be the set of image lattice sites. Let $X=\{x_s|s \in S,\ x_s \in \{0, 1, 2, ..., 255\}\}$ denote the observed image, and $Y=\{y_s|s \in S,\ y_s \in \Lambda\}$ express the label field, where $\Lambda=\{1, ..., L\}$ is a common label space, and L is the number of classes.

In the Pairwise MRF framework, the present study defines a pixel-labeling problem as assigning to every pixel x_s a label y_s. According to the Bayesian theorem, the label process can be described as the MAP problem for an appropriately defined MRF [18]:

$$Y^* = arg\max_x P(Y|X, \Theta) = arg\max_x \frac{1}{Z}\prod_s \phi_s(x_s, y_s)\prod_{sr} \psi_{sr}(y_s, y_r) \quad (1)$$

where $P(Y|X,\Theta)$ is a global probability of MRF, Θ is vector of hyperparameters of MRF, $\phi_s(x_s,y_s)$ is the conditional probability density function of x_s given a label y_s, $\psi_{sr}(y_s,y_r)$ is the prior probability function of the label field Y, and Z is a normalization constant.

However, Pairwise connected models are often insufficient for capturing the full complexity of the joint distribution of the problem [8]. The current section describes an overlapping local region MRF to solve this problem. As shown in Fig. 1, for the observed image $\boldsymbol{X}$, let a local region w_s with size $w \times w$ be centered at pixel $\boldsymbol{s}$, such that the observed image $\boldsymbol{X}$ is partitioned into $W \times H$ subimages, where W and H are the width and height of the image, respectively. Each subimage is assumed to be an MRF. In any subimage, the intensities of any two pixels are independent of each other. The relationship between the observed subimage and its label field is given as a Gaussian mixture model (GMM), and the Gaussian conditional probability function is defined as follows:

$$P_{w_s}(x_{w_s}|y_{w_s},\theta) = \prod_{s \in w_s} p(x_s|y_s,\theta) = \prod_{s \in w_s} \frac{1}{\sqrt{2\pi\sigma_l^2}} \cdot \exp\left\{-\frac{(x_s-\mu_l)^2}{2\sigma_l^2}\right\}$$
$$= \frac{1}{(2\pi\sigma_l^2)^{|w_s|/2}} \cdot \exp\left\{-\sum_{s \in w_s}\frac{(x_s-\mu_l)^2}{2\sigma_l^2}\right\} \quad (2)$$

where $|w_s|$ is the number of pixels belonging to local region w_s; $\theta=(\mu_1, \sigma_1, \mu_2, \sigma_2, \cdots \mu_L, \sigma_L)$; and μ_l and σ_l are the mean and variance of class l, respectively.

The label filed in each subimage is also assumed to be an MRF. To reduce the computational complexity, the assumption is that the label of a pixel depends only on the labels at its local region. As the Hammersley–Elifford theorem suggests [19], the joint distribution of all labels in the local region follows a Gibbs distribution and follows the form:

$$P_{w_s}(y_s|y_{w_s},\beta) = \frac{\exp\left\{-\beta H_{w_s}(y)\right\}}{\sum_{y \in L}\exp\left\{-\beta H_{w_s}(y)\right\}} \quad (3)$$

where β is the prior hyperparameter of the Gibbs distribution. In the local region, the prior function $H_{w_s}(y)$ is defined as:

$$H_{w_s}(y) = \frac{1}{\sum_{(sr) \in w_s}\left\{1-\delta(y_s,y_r)\right\}} \quad (4)$$

where y_s is the label of the center site s, and $y_r \in N(y_s)$ is the neighborhood label of y_s in the local region. $H_{w_s}(y)$ is a adaptive local region prior function since the labels y_s and y_r are changed in each iteration. The present study considers pairs of symmetrical cliques within the 3×3 neighborhood of each pixel. $\delta(y_s,y_r)$ is the kronecker delta:

$$\delta(y_s,y_r) = \begin{cases} 1 & \text{if} \quad y_s \ne y_r \\ 0 & \text{else} \end{cases} \quad (5)$$

where $\sum_{(sr) \in w_s}\left\{1-\delta(y_s,y_r)\right\}$ is the number of neighbors equal to the central label. In a local region, the local Gibbs prior information of each pixel is measured by counting the number of neighbors label equal to the central label. A greater number denotes a larger number of pixels with higher probability chose this label in the local region, and vice versa, in which pixels tend to take the same label value.

According to the Bayesian rule and Eqs. (2) and (3), the local region posterior probability can be represented as:

$$P_{\text{Local}}(y_s|x_s,x_{w_s},y_{w_s}) \propto P(x_s,x_{w_s \backslash s}|y_s,y_{w_s \backslash s})P(y_s|y_{w_s \backslash s})$$
$$= \prod_{r \in w_s \backslash s} P(x_s|y_s)P(x_r|y_r)P(y_s|y_{w_s \backslash s}) \quad (6)$$

where $w_s \backslash s$ is the set of sites in the local region w_s without the center site s.

The estimation for the optimal label y^* is the process of maximizing the following posterior:

$$y^* = arg\max_x P_{Local}(y_s|x_s,x_{w_s},y_{w_s}) \propto arg\max_x P(x_s|y_s)P(y_s|y_{w_s \backslash s})$$
$$\times \prod_{r \in w_s \backslash s} P(x_r|y_r) \quad (7)$$

From the local posterior probability Eq. (6), the estimate of label x_s not only considers the intensity field of the neighborhood y_{w_s} but