Computer Vision and Image Understanding 112 (2008) 30–38

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Measuring uncertainty in graph cut solutions

Pushmeet Kohli [a,*], Philip H.S. Torr [b]

[a] Microsoft Research, Cambridge, UK
[b] Oxford Brookes University, Oxford, UK

ARTICLE INFO

Article history:
Received 25 October 2007
Accepted 2 July 2008
Available online 15 July 2008

Keywords:
Parameter learning
Inference
Min-marginals
Graph cuts

ABSTRACT

In recent years graph cuts have become a popular tool for performing inference in Markov and conditional random fields. In this context the question arises as to whether it might be possible to compute a measure of uncertainty associated with the graph cut solutions. In this paper we answer this particular question by showing how the min-marginals associated with the label assignments of a random field can be efficiently computed using a new algorithm based on dynamic graph cuts. The min-marginal energies obtained by our proposed algorithm are exact, as opposed to the ones obtained from other inference algorithms like loopy belief propagation and generalized belief propagation. The paper also shows how min-marginals can be used for parameter learning in conditional random fields.

1. Introduction

Researchers in computer vision have extensively used graph cuts to compute the maximum a posteriori (MAP) solutions for various discrete pixel labelling problems such as image restoration, segmentation and stereo. One of the primary reasons for the growing popularity of graph cuts is their ability to find the globally optimal solutions for an important class of energy functions in polynomial time [20]. Even for problems where graph cuts do not guarantee optimal solutions they can be used to find solutions which are strong local minima of the energy [5]. These solutions for certain problems have been shown to be better than the ones obtained by other methods [4,24].

Graph cuts however do suffer from a big disadvantage. Unlike other inference algorithms, they do not provide any uncertainty measure associated with the solution they produce. This is a serious drawback since researchers do not obtain any information regarding the probability of a particular latent variable assignment in a graph cut solution. Inference algorithms such as loopy belief propagation (LBP), generalized belief propagation (GBP) and the recently introduced tree re-weighted message passing (TRW) [19,27] output approximate marginal or min-marginal energies associated with each latent variable. Note that for tree-structured graphs, the simple max-product belief propagation algorithm gives the exact max-marginal probabilities/min-marginal energies[1] for different label assignments in $O(nl^2)$ time where n is the number of latent variables, and l is the number of labels a latent variable can take.

This paper addresses the problem of efficiently computing the min-marginals associated with the label assignments of any latent variable in a Markov random field (MRF). Our method can work on all MRFs or CRFs that can be solved using graph cuts. First, we show how in the case of binary variables, the min-marginals associated with the labellings of a latent variable are related to the *flow-potentials* (defined in Section 3) of the node representing that latent variable in the graph constructed in the energy minimization procedure. The exact min-marginal energies can be found by computing these *flow-potentials*. We then show how flow-potential computation is equivalent to the problem of minimizing a *projection* of the original energy function.[2]

Minimizing a *projection* of an energy function is a computationally expensive operation and requires a graph cut to be computed. In order to obtain the min-marginals corresponding to all label assignments of all random variables, we need to solve $O(nl)$ number of st-mincut problems. In this paper, we present an algorithm based on dynamic graph cuts [16] which solves these $O(nl)$ problems extremely quickly. Our experiments show that the running time of this algorithm i.e. the time taken for it to compute the min-marginals corresponding to all latent variable label assignments is of the same order of magnitude as the time taken to compute a single graph cut. The first version of this paper appeared as [17]. This extended version shows how the min-marginals

* Corresponding author.
E-mail addresses: pkohli@microsoft.com (P. Kohli), philiptorr@brookes.ac.uk (P.H.S. Torr).

[1] We will explain the relation between max-marginal probabilities and min-marginal energies later in Section 2. To make our notation consistent with recent work in graph cuts, we formulate the problem in terms of min-marginal energies (subsequently referred to as simply min-marginals).

[2] A projection of the function $f(x_1, x_2, \ldots, x_n)$ can be obtained by fixing the values of some of the variables in the function $f(.)$. For instance $f'(x_2, \ldots, x_n) = f(0, x_2, \ldots, x_n)$ is a projection of the function $f(.)$.

1077-3142/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2008.07.002

obtained using our method can be used for parameter learning in CRFs.

1.1. Overview of dynamic graph cuts

Dynamic computation is a paradigm that prescribes solving a problem by dynamically updating the solution of the previous problem instance. Its hope is to be more efficient than a computation of the solution from scratch after every change in the problem. A considerable speedup in computation time can be achieved using this procedure, especially, when the problem is large and changes are few. Dynamic algorithms are not new to computer vision. They have been extensively used in computational geometry for problems such as range searching, intersections, point location, convex hull, proximity and many others [6].

Boykov and Jolly [3] were the first to use a *partially* dynamic st-mincut algorithm in a vision application. They proposed a technique with which they could update capacities of *certain* graph edges, and recompute the st-mincut dynamically. They used this method for performing interactive image segmentation where the user could improve segmentation results by giving additional segmentation cues (seeds) in an online fashion. However, their scheme was restrictive and did not allow for general changes in the graph. In one of our earlier papers, we proposed a new algorithm overcoming this restriction [16], which is faster and allows for arbitrary changes to be made in the graph. The running time of this new algorithm has been empirically shown to increase linearly with the number of edge weights changed in the graph. In this paper, we will use this algorithm to compute the exact min-marginals efficiently.

1.2. Our contributions

To summarize, the key contributions of this paper include:

- The discovery of a novel relationship between min-marginal energies and node flow-potentials in the residual graph obtained after the graph cut computation.
- An efficient algorithm based on dynamic graph cuts to compute min-marginals by minimizing energy function projections.
- The use of min-marginals for learning parameters of CRFs used for modelling labelling problems.

1.3. Organization of the paper

The paper starts by describing the basics of random fields and graph cuts and proceeds to discuss the relationship between min-marginals and node-flow-potentials. We then show how max-marginal probabilities can be found by minimizing projections of the energy function defining a MRF or CRF, and how dynamic graph cuts can be used to efficiently compute the minimum values of these projections. Our algorithm can handle all energy functions that can be solved using graph cuts [9,11,14,20].

We discuss random fields and min-marginal energies in Section 2. In Section 3, we formulate the st-mincut problem, define terms that would be used in the paper, and describe how certain energy functions can be minimized using graph cuts. In Section 4, we show how min-marginals can be found by minimizing projections of the original energy function. In the same section we describe a novel algorithm based on dynamic graph cuts to efficiently compute the minima of these energy projections. In Section 5, we discuss some applications of our algorithm.

2. Notation and preliminaries

We will now describe the notation used in the paper. We will formulate our problem in terms of a *pairwise* MRF.[3] Note that the pairwise assumption does not affect the generality of our formulation since any MRF involving higher order interaction terms can be converted to a *pairwise* MRF by addition of auxiliary variables in the MRF [28].

Consider a discrete random field $\mathbf{X}$ defined over a lattice $\mathscr{V} = \{1, 2, \ldots, n\}$ with a neighbourhood system $\mathscr{N}$. Each random variable $X_i \in \mathbf{X}$ is associated with a lattice point $i \in \mathscr{V}$ and takes a value from the label set $\mathscr{X}_v$. The neighbourhood system $\mathscr{N}$ of the random field is defined by the sets $\mathscr{N}_i, \forall i \in \mathscr{V}$, where $\mathscr{N}_i$ denotes the set of all neighbours of the variable X_i. Any possible assignment of labels to the random variables is called a *labelling* or *configuration*. It is denoted by the vector $\mathbf{x}$, and takes values from the set $\mathscr{X}$ defined as $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2 \times \cdots \times \mathscr{X}_n$. Unless noted otherwise, we use symbols u and v to denote values in $\mathscr{V}$, and i and j to denote particular values in $\mathscr{X}_u$ and $\mathscr{X}_v$, respectively.

A random field is said to be a Markov random field (MRF) with respect to a neighbourhood system $\mathscr{N} = \{\mathscr{N}_v | v \in \mathscr{V}\}$ if and only if it satisfies the positivity property: $\Pr(\mathbf{x}) > 0 \forall \mathbf{x} \in \mathscr{X}$, and the Markovian property:

$$\Pr(x_v | \{x_u : u \in \mathscr{V} - \{v\}\}) = \Pr(x_v | \{x_u : u \in \mathscr{N}_v\}), \quad \forall v \in \mathscr{V}. \tag{1}$$

Here we refer to $\Pr(X = \mathbf{x})$ as $\Pr(\mathbf{x})$ and $\Pr(X_i = x_i)$ as $\Pr(x_i)$. A conditional random field (CRF) may be viewed as an MRF globally conditioned on the data.

The MAP–MRF estimation problem aims to find the configuration $\mathbf{x}$ which has the highest probability. It can be formulated as an energy minimization problem where the energy corresponding to a MRF configuration $\mathbf{x}$ is defined as

$$(\mathbf{x}|\theta) = -\log \Pr(\mathbf{x}|\mathbf{D}) - \text{const.} \tag{2}$$

Here θ is the energy parameter vector defining the MRF [19]. The energy functions characterizing MRFs used in computer vision can usually be written as a sum of unary and pairwise terms:

$$E(\mathbf{x}|\theta) = \sum_{v \in \mathscr{V}} \left(\phi(x_v) + \sum_{u \in \mathscr{N}_v} \phi(x_u, x_v) \right) + \text{const.} \tag{3}$$

In the paper, $\psi(\theta)$ is used to denote the value of the energy of the MAP configuration of the MRF and is defined as

$$\psi(\theta) = \min_{\mathbf{x} \in \mathscr{X}} E(\mathbf{x}|\theta). \tag{4}$$

In what follows, the term *optimal solution* will be used to refer to the MAP solution of the random field.

2.1. Min-marginal energies

A min-marginal is a function that provides information about the minimum values of the energy E under different constraints. Following the notation of [19], we define the min-marginal energies $\psi_{v;j}, \psi_{uv;ij}$ as

$$\psi_{v;j}(\theta) = \min_{\mathbf{x} \in \mathscr{X}, x_v = j} E(\mathbf{x}|\theta), \tag{5}$$

$$\psi_{uv;ij}(\theta) = \min_{\mathbf{x} \in \mathscr{X}, x_u = i, x_v = j} E(\mathbf{x}|\theta). \tag{6}$$

In words, given an energy function E whose value depends on the variables $(X_1, X_2, \ldots, X_n)$, $\psi_{v;j}(\theta)$ represents the minimum energy value obtained if we fix the value of variable X_v to j $(x_v = j)$ and minimize over all remaining variables. Similarly, $\psi_{uv;ij}(\theta)$ represents the value of the minimum energy obtained by assigning labels i and j to

[3] Pairwise MRFs have cliques of size at most two.