Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 2163–2172
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Incorporating shape prior into geodesic active contours for detecting partially occluded object

Wen Fang, Kap Luk Chan*

School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798, Singapore

Abstract

A new method to incorporate shape prior knowledge into geodesic active contours for detecting partially occluded object is proposed in this paper. The level set functions of the collected shapes are used as training data. They are projected onto a low dimensional subspace using PCA and their distribution is approximated by a Gaussian function. A shape prior model is constructed and is incorporated into the geodesic active contour formulation to constrain the contour evolution process. To balance the strength between the image gradient force and the shape prior force, a weighting factor is introduced to adaptively guide the evolving curve to move under both forces. The curve converges with due consideration of both local shape variations and global shape consistency. Experimental results demonstrate that the proposed method makes object detection robust against partial occlusions.

Keywords: Geodesic active contours; Shape prior; Object detection

1. Introduction

Object detection and tracking is necessary in a wide variety of applications of computer vision such as video surveillance, object-based coding, and robotics, etc. [1–3]. Variational methods have taken an important role in realizing this task. A number of variational methods [4–7] have been developed to date, they can be roughly divided into two categories: boundary-based approach and region-based approach. The boundary-based approach depends on information of edges to detect objects. Due to their relatively reliable performance and fast computation, they have been widely adopted [8]. However, as only edge information is utilized, their detection accuracy is determined by the strength of the image gradient. The presence of object occlusion and other scene clutter can lead to erroneous detection. Region-based approach depends on information of the entire region such as color or texture to detect objects [9–11]. They are more robust than boundary-based methods since the information of the whole region is explored. However, they need more computations and the region is restricted to be uniform and occlusion free.

This paper focuses on the boundary-based variational approach to detect objects that may be partially occluded, with nonuniform inner region and also cater for the need for fast computation, such as for the surveillance application. There are two well-known variational methods: the active contour method [4] and the level set method [12]. The active contour (or snake) method is a parametric deformable model which pulls the curve towards the objects by minimizing an energy function composed of internal and external energies. The main problem of this method is that it cannot handle the topological changes of objects. Hence, in recent years, the level set method has been used more favorably [13–15]. The level set method is a geometric deformable model which solves its partial differential equation by minimizing the geodesic distance in a Riemannian space [7]. The level set method has the ability to extract contour in concave regions and cope with topological changes of objects. However, like most of the boundary-based methods, it is unable to find the correct objects in the presence of object occlusion or background clutter. As only the edge information is explored, curves will be "deceived" by spurious contours resulted from large local gradient nearby and converge to the wrong location.

* Corresponding author. Tel.: +65 67904514.
E-mail address: eklchan@ntu.edu.sg (K.L. Chan).

0031-3203/$30.00 © 2007 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2006.12.014

Various attempts have been reported to solve the aforementioned problem. A common approach is to learn the object shape from a set of training samples, and incorporate this prior knowledge as a global shape constraint into the variational methods to confine the detection of current shape within a cluster of similar shapes [16,17]. Chen [18,19] proposed to use the average of a training set of aligned curves as their shape model, while Zhang [20] defined their shape model to be the level set function of a preselected reference shape. In Refs. [21–23], only a single shape was considered as the shape prior information also. The above methods focus more on the numerical computation of the shape priors and their incorporations into a variational framework, they do not elaborate on the problem of representing the shape of objects. The shape priors they used are too simple to capture the rich characteristics of objects.

There are reported research work that use statistical distribution to construct the shape priors. In Ref. [11], the author used the level set functions of training shapes as training data, and estimated their distribution with a Gaussian function. A shape model was constructed in the original data space which is a high dimensional space. A similar shape prior model was developed in Ref. [24]. The distribution of the level set representations of a cluster of object shapes were directly described by a Gaussian function in the original data space. The problems of these methods are that they suffer from the curse of dimensionality. The training data they used are created from certain representations of shapes (e.g. level set functions), hence their dimensionality depends on the size of the shape template which can often be very large. The training data distribute sparsely in a high dimensional space, making it difficult to describe the real data structure. Hence, dimensionality reduction techniques are employed in constructing the prior model process. Wang [25] computed the corresponding points across a set of training shapes and constructed a statistical model in the subspace determined by the principal component analysis (PCA) on these points. But using marker points to describe boundaries suffers from the numerical instability and inability to accurately capture high curvature locations. Leventon [26] extended Wang's technique into a level set-based framework and modeled the level set representations of object shapes as a Gaussian distribution. PCA was used to reduce the dimensionality of training data and a Gaussian function was estimated in the subspace. To incorporate this shape model into the curve evolution equation, a maximum a posteriori (MAP) approach was used to find a best matching shape, and the level set curve was forced to move towards it. In Refs. [27,28], Tsai integrated the shape model of Leventon's into the reduced version of the Mumford-Shah functional [5]. The same shape prior model was also used in Refs. [29,30]. However, Leventon's method suffers from slow computation. At each step of its curve, evolution, the maximum a posterior position must be found to adjust the pose of the prior model. Moreover, many parameters need to be set some of which are determined empirically.

In order to employ the shape prior knowledge to constrain the curve evolution within a cluster a similar shapes, a collection of training shapes are needed from which we can learn the general characteristics of object shapes. Furthermore, a more sophisticated method to represent these shape characteristics should be considered. Hence, in this paper, we aim at using a statistical method to describe the structure of a set of preselected shapes and construct the shape prior model based on these shapes. Besides, we found that the following points should be considered:

- Since the level set representations of object shapes are commonly used as training data and they suffer from the curse of dimensionality, dimension reduction technique is needed. The data distribution should be estimated in the lower dimensional subspace.
- If the shape prior model is created to improve the detection performance in surveillance videos, then it should be incorporated into the variational methods without causing significant increase of computational complexity.

In this paper, a new method to incorporate a shape prior model into geodesic active contours is proposed. The level set function of object shapes are collected as training data and are projected onto the first k principal directions using PCA. In the reduced subspace, the data distribution is approximated by a Gaussian function. A shape prior model is constructed and is used to constrain the curve evolution process. To speed up the convergence under the influence of shape prior knowledge, we propose to explicitly align the shape model with the shape of current curve under evolution. Thus less iterations are needed in adjusting the pose compared to Leventon's MAP approach to seek the pose parameter. Moreover, we develop a weighting scheme to balance the influence between the image gradient and shape prior information by quantifying their differences, unlike Lenvenon's work in which some of the parameters have to be set empirically.

The rest of the paper is organized as follows: in Section 2, the basic rationale of geodesic active contour method is reviewed. Section 3 describes the construction process of our shape prior model. The incorporation of the shape model into the geodesic active contour equation is presented in Section 4. Finally, experimental results and concluding remarks are given in Sections 5 and 6.

2. Geodesic active contours

Let $C(q)$: $[0, 1] \to \mathbb{R}^2$ be a parametrized curve and let I: $[0, a]\times[0, b] \to \mathbb{R}^+$ denote a given image. The geodesic active contour method [7] evolves its curve by finding a geodesic curve in a Riemannian space with a metric derived from the image

$$\min \int_0^1 g(|\nabla I(C(q))|)|C'(q)|\,\mathrm{d}q, \tag{1}$$

where g is a stopping function which contains information regarding the boundary of the object. g is usually of the form $g = 1/(1 + |\nabla \hat{I}|^p)$, where $\hat{I}$ is a smoothed version of I and p can be 1 or 2.

Computing the Euler–Lagrange of Eq. (1) and searching the gradient descent direction to find its minimum value, Eq. (2)