Image and Vision Computing 28 (2010) 1031–1038

Contents lists available at ScienceDirect

Image and Vision Computing

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

Generalized sparse MRF appearance models ☆

René Donner [a,*], Georg Langs [a], Branislav Mičušik [b], Horst Bischof [c]

[a] Computational Image Analysis and Radiology Lab, Department of Radiology, Medical University of Vienna, Währinger Gurtel 18-22, 1090 Vienna, Austria
[b] George Mason University, 4400 University Drive MS 4A5, Fairfax, VA 22030, USA
[c] Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, 8010 Graz, Austria

ARTICLE INFO

Article history:
Received 18 June 2008
Received in revised form 4 December 2008
Accepted 6 July 2009

Keywords:
Markov Random Fields
Appearance models
Object localization
Medical imaging

ABSTRACT

Image segmentation or registration approaches that rely on a local search paradigm (e.g, Active Appearance Models, Active Contours) require an initialization that provides for considerable overlap or a coarse localization of the object to be segmented or localized. In this paper we propose an approach that does not need such an initialization, but localizes anatomical structures in a global manner by formulating the localization task as the solution of a *Markov Random Field (MRF).*

During search *Sparse MRF Appearance Models (SAMs)* relate a priori information about the geometric configuration of landmarks and local appearance features to a set of candidate points in the target image. They encode the correspondence probabilities as an MRF, and the search in the target image is equivalent to solving the MRF. The resulting node labels define a mapping of the modeled object (e.g. a sequence of vertebrae) to the target image interest points. The local appearance information is captured by novel symmetry-based interest points and local descriptors derived from *Gradient Vector Flow (GVF).* Alternatively, arbitrary interest points can be used. Experimental results are reported for two data-sets showing the applicability to complex medical data. The approach does not require initialization and finds the most plausible match of the query structure in the entire image. It provides for precise, reliable and fast localization of the structure.

1. Introduction

The reliable and fast segmentation of anatomical structures is a central issue in medical image analysis. It has been tackled by a number of powerful approaches. Among them are Active Shape Models [3], Active Appearance Models [4,6], Active Feature Models [17], Graph-Cuts [2], Active Contours [11], or Level-Set approaches [27]. These methods have been successfully applied to segment structures in cardiac MRIs [25] for the registration in functional heart imaging [33], rheumatoid arthritis monitoring [16], the delineation of vertebrae in the spine [28] or as preprocessing step in bone densiometry [34].

An important question and limitation of these methods is their capture range, i. e. how accurate an initialization needs to be to avoid local minima and to ensure the convergence at meaningful segmentation results. ASMs and AAMs need to be placed with considerable overlap with the object of interest. Graph-cuts need a set of seed points placed within and outside of the object, while Active Contours need spatial constraints to ensure the delineation of the correct object. Usually the initialization is either done manually or by application-specific approaches.

A related line of work deals with object recognition and the parts based representation of objects and object categories. The approach presented in [9] learns a probabilistic representation of the objects using interest points and solves the recognition task using a bayesian approach, while [7] build tree graphs of objects parts automatically from image data. While both show good results for recognizing objects in scenes the small number of objects parts limits their use for localization. [26] showed how a part-based model can be learnt from segments of the objects' boundaries and used for object detection. The reliance on object boundaries, though, makes it not well suited for radiological data where the reliable detection of edges is obstructed by high levels of noise and weak gradients. The relation of these methods to the approach proposed in this paper highlights the potential of going beyond the localization of a priori known structures. It indicates that learning more detailed models from large data-sets might be feasible by using strategies like the one presented in this paper.

The recent advances in solving the inference problem in multi-label Markov Random Fields have opened the door for new approaches to computer vision challenges like stereo reconstruction [35], 3D from single images [30], image registration [32], segmentation [23,10], early vision [8] and shape matching [37]. Seminal works like tree reweighted message passing [12], second order

☆ Expanded version of a paper presented at the 18th British Machine Vision Conference (Warwick, September 2007).

* Corresponding author. Tel.: +436508214307.
E-mail addresses: donner@icg.tu-graz.ac.at (R. Donner), georg.langs@meduniwien.ac.at (G. Langs), bmicusik@gmu.edu (B. Mičušik), bischof@icg.tu-graz.ac.at (H. Bischof).

0262-8856/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2009.07.010

cone programming [15] and linear programming [13]/the MAX-SUM approach [36] have provided the community with the insight on how to solve MRFs in practice with near-optimal solutions. While the optimal solution of a Graph-Cut can be found very efficiently, multi-label MRFs are NP-hard to solve and thus only approximate solutions are possible in reasonable time for non-trivial cases.

The main contribution of this paper is a method that formulates the localization of an object model in an image as the solution to the optimal labeling task of a *Markov Random Field (MRF)* that encodes the relation between the model and the *entire* search image. The method integrates a Sparse Appearance Model acquired during a training phase with a global search method. On the one hand this enables us to take advantage of the rich descriptive power of local appearance features, and on the other hand we can use an efficient discrete optimization framework to perform the localization of such a clique of points. The detection is performed in a fast manner by solving the MRF with the MAX-SUM algorithm [36]. The approach finds a solution that minimizes the combined costs of non-rigid deformations and local descriptor feature differences between the target and the model.

The second contribution of the paper is the introduction of symmetry-based point detectors and descriptors based on Gradient Vector Flow [38] that is particularly well suited for medical image data. State of the art local symmetry detectors [14,21] are either computationally expensive or use radial symmetry detection with predefined radii. Recently [20] proposed an approach to detect symmetry in the constellation of interest points detected by existing point detection methods. The detector and local descriptor explained in this work uses Gradient Vector Flow to derive a stable and repeatable description of local image structure. Based on the GVF formulation it integrates local details as well as more global structure that allows for a stable detection and identification of individual points. Experiments indicate that these are particularly favorable properties for medical data-sets which exhibit a level of fine structure variation that can deteriorate the performance of purely local descriptors. Alternatively, we show how arbitrary interest points can be used. We report results for interest points based on local symmetry and a complementary local descriptor derived from Gradient Vector Flow as well as for a combined approach employing Harris corners. The approach outlined in this paper thus represents the generalization of the work first published in [5].

The remainder of the paper is structured as follows: In Section 2 we explain the interest point detector and local descriptor. Section 3 details Markov Random Fields. In Section 4 the construction of SAMs is outlined and the mapping of a source model to target points by the MRF is explained in detail. In Section 5 we outline how to use an arbitrary interest point detector in the SAM framework. In Section 6 we present the experimental evaluation of our approach, followed by a conclusion and an outlook in Section 7.

2. Symmetry based interest points and descriptors

Many structures of interest to medical experts, like bones, veins and many anatomical structures or their parts exhibit a shape with a high degree of symmetry w.r.t. one or more axes. This property of (local) symmetry is well preserved even when dealing with projections like radiographs, or 2D slices of 3D data-sets like MRIs, as the cross sections of these body parts will appear as round or elongated structures. Even regions of interest that do not exhibit this property can be localized by observing their neighborhood, e.g. an initialization for meniscoids can be provided by correctly localizing the discs and vertebrae of the spine.

2.1. Interest points from local symmetry

To detect points of high local symmetry we use the Gradient Vector Flow field, which was originally proposed in [38] to increase the capture range of Active Contours. A GVF field $GVF(\mathbf{J})$ of a binary edge map $\mathbf{J}$ is a non-conservative vector field $\vec{v}(x,y) = [u(x,y), v(x,y)]$ that minimizes the energy functional

$$\int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\Delta\mathbf{J}|^2 |\vec{v} - \Delta\mathbf{J}|^2 \, dx\, dy. \quad (1)$$

It is nearly equal to the gradient of the edge map $\vec{v} \approx \Delta\mathbf{J}(\mathbf{I})$ if $\Delta\mathbf{J}(\mathbf{I})$ is large, and varies slowly in homogeneous regions of $\mathbf{J}$. The parameter μ governs the smoothness of the resulting fields.

Its strengths include the ability to detect even weak structures while being robust to high amounts of noise in the image. This allows to derive sensible interest points for structures of varying size within an image without the need to set a scale parameter. Instead of computing the GVF of $\mathbf{J}$ as used in the Active Contour context we compute the GVF directly from the gray level image $\mathbf{I}$ as $\mathbf{G} = GVF(\mathbf{I})$, yielding the complex matrix $\mathbf{G}$ used for all subsequent computations. The resulting field $\mathbf{G}$ is depicted in Fig. 1 for a section of a hand radiograph, overlaid over the image $\mathbf{I}$.

The field magnitude $|\mathbf{G}|$ is largest in areas of high image gradient, and the start- and endpoints of the field lines of $\mathbf{G}$ are located at symmetry maxima. E.g. in the case of a symmetrical structure formed by a homogeneous region surrounded by a different gray level value the field will point away from or towards the local symmetry center of the structure, as shown in Fig. 1(a) and (c). The symmetry interest points, which we will call *gvf-Points*, are thus defined as the local minima of $|\mathbf{G}|$.

In comparison, popular interest point detectors like the Harris corner detector or the Difference of Gaussians (DoG) approach do not possess a preference for local symmetry. An illustration of the interest points detected by DoG and interest points derived from local symmetry is shown in Fig. 2(a) and (b).

After detecting the interest points the orientation $\alpha_i \in [0, 2, \pi[$ of each interest point i can be estimated by the mode $\mathbf{e}^1$ of the largest variance of the field vectors **Gi** in a local neighborhood around interest point i (10 pixel radius throughout this paper):

$$\alpha_i = \arctan(\mathbf{e}_1) \quad (2)$$

The scale *si* of the region around the interest point can be estimated by the mean distance of the interest point i to the two closest local maxima of $|\mathbf{G}|$ in the directions of α_i and $\alpha_i + \pi$. If the scale varies only within a limited range as for the medical images examined in this paper the scale can remain fixed.

2.2. Local descriptors from gradient vector flowfields

In addition to indicating interest points the GVF field can serve as an appearance descriptor. It integrates structures of different scale (local structure if present, global properties if no strong local structure is available), and thus provides an adaptive and stable description in the case of fine and highly variable structure. We use it to define a descriptor of the local regions at the positions of the symmetry interest points. Several highly discriminating local descriptors have been proposed in recent years, including Shape Context [1], SIFT [19] and GLOH [24]. While most of these approaches yield descriptors suitable for building the MRF, they would require additional computations. In contrast, we can directly use $\mathbf{G}$ to describe local context. Furthermore our experiments show that highly discriminant descriptors are unfavorable for the match confidence estimation employed in this work, potentially eliminating valid candidate points.

Each interest point is assigned a local descriptor $\mathbf{D}_i \in \mathbb{R}^{2 \cdot nd \cdot nd}$ which is derived from $\mathbf{G}$ by extracting patches around the interest