Medical Image Analysis 32 (2016) 101–114

Contents lists available at ScienceDirect

Medical Image Analysis

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

Active appearance pyramids for object parametrisation and fitting

Qiang Zhang[a,*], Abhir Bhalerao[a], Edward Dickenson[b], Charles Hutchinson[b]

[a] Department of Computer Science, University of Warwick, Coventry, CV4 7AL, UK
[b] University Hospitals Coventry and Warwickshire, Coventry, CV2 2DX, UK

ARTICLE INFO

Article history:
Received 25 June 2015
Revised 8 March 2016
Accepted 10 March 2016
Available online 1 April 2016

Keywords:
Lumbar spinal stenosis
Active appearance model
Part-based model
Active appearance pyramid

ABSTRACT

Object class representation is one of the key problems in various medical image analysis tasks. We propose a part-based parametric appearance model we refer to as an Active Appearance Pyramid (AAP). The parts are delineated by multi-scale Local Feature Pyramids (LFPs) for superior spatial specificity and distinctiveness. An AAP models the variability within a population with local translations of multi-scale parts and linear appearance variations of the assembly of the parts. It can fit and represent new instances by adjusting the shape and appearance parameters. The fitting process uses a two-step iterative strategy: local landmark searching followed by shape regularisation. We present a simultaneous local feature searching and appearance fitting algorithm based on the weighted Lucas and Kanade method. A shape regulariser is derived to calculate the maximum likelihood shape with respect to the prior and multiple landmark candidates from multi-scale LFPs, with a compact closed-form solution. We apply the 2D AAP on the modelling of variability in patients with lumbar spinal stenosis (LSS) and validate its performance on 200 studies consisting of routine axial and sagittal MRI scans. Intervertebral sagittal and parasagittal cross-sections are typically used for the diagnosis of LSS, we therefore build three AAPs on L3/4, L4/5 and L5/S1 axial cross-sections and three on parasagittal slices. Experiments show significant improvement in convergence range, robustness to local minima and segmentation precision compared with Constrained Local Models (CLMs), Active Shape Models (ASMs) and Active Appearance Models (AAMs), as well as superior performance in appearance reconstruction compared with AAMs. We also validate the performance on 3D CT volumes of hip joints from 38 studies. Compared to AAMs, AAPs achieve a higher segmentation and reconstruction precision. Moreover, AAPs have a significant improvement in efficiency, consuming about half the memory and less than 10% of the training time and 15% of the testing time.

1. Introduction

Representation and segmentation of anatomical objects is of vital importance in the understanding of medical images. A standard approach which has proven robust and efficient, is to learn and leverage prior knowledge of the object garnered from statistics of its parametric form. To achieve this, the following steps are implemented: delineating the object class with a coherent parametric form; learning a prior model of the object class by formulating the statistics of the parameters; and fitting the parametric model to new, unseen instances while regularising the solution with the learned prior model.

The most commonly used strategy is to describe the objects with deformable appearances such as morphable models (Jones and Poggio, 1998), statistical deformable models (Rueckert et al., 2003) and AAMs (Cootes et al., 2001; Matthews and Baker, 2004). The correspondence in the training data are established by annotating the landmarks at consistent features of interest from subjects. The prior knowledge is then usually learned through a linear model by applying eigen analysis, e.g., PCA. As a generative method, AAMs can not only achieve a robust segmentation, but also synthesise new instances and code the appearance with compact parameters for higher-level interpretation, such as for the diagnosis and grading of pathologies. AAMs are widely adopted and have proven successful, but in clinical applications face challenges such as their sensitivity to local minima during fitting, and computational costs when built on 3D data.

In addition to the holistic methods, part-based models have shown superior performance in computer vision tasks including object detection and tracking. Notable examples are sub-model AAMs (Roberts et al., 2006; Roberts, 2008), Deformable Part Models ([16]; Felzenszwalb et al., 2010a; 2010b), CLMs (Cristinacce and Cootes, 2006; 2008; Saragih et al., 2011) and mixture-of-trees models (Zhu and Ramanan, 2012), in which an object is decomposed into locally rigid parts with a geometric model capturing

* Corresponding author. Tel.: +4407442336388.
E-mail address: qiang@dcs.warwick.ac.uk (Q. Zhang).

http://dx.doi.org/10.1016/j.media.2016.03.005
1361-8415/

spatial relationships among parts. Among these the models reported applied for clinical applications are sub-model AAMs and CLMs. For example in Cristinacce and Cootes (2008) the CLMs show superior performance over AAMs on brain and dental images. In Lindner et al. (2013), combined with random forests regression CLMs are reported to have the best performance in segmenting femur radiographs. The fitting process is implemented by local feature searching followed by a regularisation imposed through a prior model of the global shape. CLMs decompose the complex appearance into parts with simpler structures therefore suffer less from the high dimension low sample space (HDLSS) problem when compared to AAMs. Moreover they are able to utilise advanced feature detection algorithms such as boosted regression (Cristinacce and Cootes, 2007), random forests (Lindner et al., 2013), regularised mean-sift (Saragih et al., 2009), and shape optimisation methods such as pictorial structures (Antonakos et al., 2015) and non-parametric model (Xiong and De la Torre, 2013). Due to the small local support of the feature patches however, the local feature detectors in CLMs are plagued by the problem of ambiguity, which results in errors in landmark location as the detection becomes trapped in local minima (Saragih et al., 2011). In addition, the existing part-based models coarsely delineate the objects focussing on capturing the key features which is sufficient in computer vision tasks, but in clinical applications a more delicate appearance model is needed to preserve the structural details and parametrise the entire anatomical appearance.

We present a generative part-based appearance model we refer to as an AAP. An AAP utilises the power of local feature searching and shape regularisation algorithms like a part-based model. Meanwhile it enhances the robustness of part searching with multi-scale local feature descriptors. Compared to CLMs, AAPs are more robust to initialisation having a wider capture range, plus individual landmarks on the shape are less prone to becoming trapped in local minima. Moreover an AAP is able to model the anatomical variations among the population and reconstruct delicate appearance as well if not better than AAMs, and have superior performance in computational efficiency and precision. Our work differ from the previous part-based models in that instead of fitting the shape, we focus on a parametric representation which can model and visualise the whole appearance variations within an object class, and fit the model to new instance to obtain the parametric representation of both the shape and appearance. The main contributions integrated in the proposed method are threefold: (1) A multi-scale LFP as the part delineation which offers a comprehensive description of the local feature and shows resistance to local minima; (2) An efficient AAP fitting algorithm derived from the weighted Lucas and Kanade (LK) methods (Lucas et al., 1981); (3) A shape regulariser integrating multiple landmark candidates from the LFPs, with a closed-form solution of the maximum likelihood (ML) shape.

In this paper, we detail how AAPs are constructed, trained and fitted and demonstrate that the appearance of an object can be delineated with multi-scale parts and that an associated deformation can be approximated by a set of locally rigid transformations of the parts. We set out the context of the problem in Section 2 and detail the AAP in Section 3. In Section 4 we derive an efficient fitting algorithm based on the weighted LK method and a regulariser utilising multi-scale landmark candidates. In Section 5 we apply 2D AAPs for modelling and fitting of lumbar vertebrae in axial and parasagittal MRI slices, which exhibit varied LSS. We demonstrate their performance against AAMs and CLMs by measuring the convergence range, segmentation accuracy and reconstruction precision. We also present experiments of 3D AAPs validated on CT data of the pelvis focussing on the hip joint. We compare the storage, computational saving as well as the segmentation and reconstruction quality against AAMs. We conclude with a discussion of the relative merits of AAPs and give proposals for further improvement .[1]

2. Background

The range of object representation and active fitting methods proposed in the literature strive to improve performance and precision. The methods have thus been adapted in various ways: to allow the prior models to compactly capture variation yet be able fit to unseen instances containing pathology; and prevent the fitting becoming trapped in local minima whilst maintaining a simplicity in object parametrisation and efficiency in fitting. We consider the challenge of local minima during fitting and how the choice of delineation (parametrisation) of objects can resolve this problem, but also result in a more flexible parts model which is efficient.

2.1. Local minima

Local minima are a problem facing all shape and appearance based methods. They not only reduce the convergence range, which affects the initialisation, but also introduce large errors to the fitting results. In both holistic and part-based methods, a coarse-to-fine strategy is often employed, which naturally increases the 'capture range' of the initialisation. However, even if at the finest level the model is close to the desired solution, the occurrence of local minima is still likely to divert the model from it (Brunet et al., 2009).

Part-based models such as CLMs are plagued by the local minima problem due to their small local support and the large appearance variation. The most effective strategy is to manipulate the scale. For instance, an efficient constrained mean shift method is proposed by Saragih et al. (2009); 2011), in which a varying kernel density estimate (KDE) is applied to perform coarse-to-fine fitting. The method starts with a smooth unimodal Gaussian model, and refines the fidelity by reducing the smoothness and increasing the number of modes. Roberts et al. (2009), searches for the local patches with coarse-to-fine resolution and use the results as an initialisation for the AAM fitting. Tresadern et al. (2009), use a hierarchy of shape models to extend the CLM where the relationships between landmarks at each level is modelled by a MRF: the local models 'select' the best candidate points and the global model acts as a regulariser. They demonstrate an improvement in performance over CLMs. Despite the optimisation in feature searching algorithms, the choice of the feature scale (size of the image patches) itself is a trade-off between the location specificity and textural properties. Also the features at different landmarks themselves can have salient edges at varying scales (see for example Fig. 2(b)), therefore an unitary scale for the descriptors across all landmarks will not capture faithfully all the salient features. We confirm that a LFP combining multi-scale local features at each landmark gives a more comprehensive description, and the shape fitting with multiple landmark estimations shows an ability to resist local minima.

2.2. Object class representation

In medical images, structural degeneration is often seen with the local appearance changes. For example in MRIs of patients with LSS, vertebral degeneration is often seen as an abnormal shape along with local intensity changes which could indicate facet joint thickening (Fig. 5(b)) and/or disc herniation (Fig. 5(c)) and occasionally inflammation or fractures. In this instance, because the intensity and structural variations are related and coupled, a

[1] Videos as well as other supplementary materials are available online at http://sites.google.com/site/activeappearancepyramids/.