Pattern Recognition 52 (2016) 206–217

Contents lists available at ScienceDirect

Pattern Recognition

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

Manifold-based constraints for operations in face space

Ankur Patel [a,1], William A.P. Smith [b,*]

[a] Conversant Inc., 101 North Wacker, 23rd Floor, Chicago, IL 60606, United States
[b] University of York, Department of Computer Science, Deramore Lane, York, YO105GH, UK

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form
25 August 2015
Accepted 3 October 2015
Available online 22 October 2015

Keywords:
Optimisation on manifolds
Face space
3D morphable models
Constrained optimisation
Statistical modelling

ABSTRACT

In this paper, we constrain faces to points on a manifold within the parameter space of a linear statistical model. The manifold is the subspace of faces which have maximally likely distinctiveness and different points correspond to unique identities. We provide a detailed empirical validation for the chosen manifold. We show how the Log and Exponential maps for a hyperspherical manifold can be used to replace linear operations such as warping and averaging with operations on this manifold. Finally, we use the manifold to develop a new method for fitting a statistical face shape model to data, which is both robust (avoids overfitting) and overcomes model dominance (is not susceptible to local minima close to the mean face). We provide experimental results for fitting a dense 3D morphable face model to data using two different objective functions (one underconstrained and one with many local minima). Our method outperforms generic nonlinear optimisers based on the BFGS Quasi-Newton method and the Levenberg–Marquardt algorithm when fitting using the Basel Face Model.

1. Introduction

Modelling "face space" (the manifold on which valid faces lie) is a longstanding goal in statistical shape analysis and computer vision and has been performed in various domains including 2D [1] and 3D [2] shape, appearance [3] and texture [4]. These approaches can be viewed as manifold learning where the faces are assumed to lie on an unknown manifold, the structure of which is learnt from data. Most commonly, the manifold is assumed to be a hyperplane (linear subspace) and the principal axes of the plane are estimated from training data using Principal Components Analysis (PCA). Applying these models to face analysis tasks requires a means to fit the model to observed data. Often this fitting process is underconstrained, prone to converge on local minima and computationally expensive. For these reasons, there is strong motivation for developing more constrained face space models in order to reduce the search space of the fitting process.

An alternative to manifold learning is to assume that the structure of the face space manifold is known. For example, the Grassmannian manifold of subspaces of a vector space has been used in face recognition [5] and the Kendall manifold of shapes has been used to model face shape [6].

The model we propose in this paper can be viewed as a hybrid of these two approaches in the sense that we assume the shape of the manifold is known (hyper-ellipsoidal) but we use manifold learning (PCA) to discover its principal axes from data. The motivation for this choice of model is as follows.

Psychological results [7,8] have shown that the parameter space of a PCA-based model has an interesting perceptually motivated interpretation: *identity* relates to direction in parameter space while *distinctiveness* is related to vector length (or equivalently distance from the mean). The reason for this is that increasing the length of a parameter vector simply exaggerates its differences from the average linearly, in other words its *features*, whereas rotating a parameter vector changes the *mix* of features present in the face. This is the justification for using angular difference in face space as a measure of dissimilarity for face recognition [4].

This decomposition also allows a useful probabilistic interpretation. Under the assumption that the original data forms a Gaussian cloud in a high dimensional space, each model parameter is independent and distributed according to a Gaussian distribution. This means that all faces lie on or near the surface of a hyperellipsoid in parameter space, with the probability density over the parameter vector lengths following a chi-square distribution. In other words, distinctiveness is subject to a statistical

* Corresponding author. Tel.: +44 (0)1904 325575.
E-mail addresses: ankurrp@gmail.com (A. Patel), william.smith@york.ac.uk (W.A.P. Smith).
[1] Tel.: +1 847 2179598.

http://dx.doi.org/10.1016/j.patcog.2015.10.003
0031-3203/

prior with the distinctiveness of most samples clustered around the expected length.

In this paper, we use these observations to motivate a representation for faces which decomposes face appearance into identity and distinctiveness subspaces. We focus on statistical models of 3D face shape. However, any class of objects amenable to linear statistical modelling using PCA could make the same identity/distinctiveness decomposition. We use ideas from differential geometry to develop tools which operate in the identity subspace, i.e. which retain constant distinctiveness. We provide empirical justification for constraining samples to have fixed distinctiveness, determined by the expected vector length.

We propose a new algorithm for fitting a statistical face model to data. Many such methods have been proposed previously, the details being dependent on the precise nature of the model and data. This inevitably involves a nonlinear optimisation over the model parameters. Our approach is more general and can be applied to any objective function. It operates via gradient descent on the manifold of equal distinctiveness. In other words, we solve for identity and assume distinctiveness takes its expected value. We show how the method naturally lends itself to a coarse-to-fine optimisation strategy and how the result avoids local minima or overfitting without having to select a regularisation weight parameter. We show that this offers improved performance over two generic nonlinear optimisation algorithms.

1.1. Related work

Perhaps the best known statistical face model is the Active Appearance Model (AAM) [3] which combines a linear model of 2D shape and 2D appearance. Rather than model appearance, the 3D Morphable Model of Blanz and Vetter [4] models the shape and texture which give rise to appearance via a model of image formation. Xiao et al. [9] have used a 3D model in conjunction with a 2D appearance model to enforce geometric constraints on the 2D shape generated.

Construction or training of a statistical face model involves a number of steps: (1) data collection, (2) registration (e.g. transforming the face data to a vector space) and (3) statistical analysis. When represented in a vector space, face-like samples can be synthesised by taking convex combinations of training faces. However, it is the statistical analysis which allows us to study how the face samples distribute themselves in high dimensional space and which regions of this space correspond to plausible faces, i.e. face space.

Although statistical face models have useful applications when used in a purely generative manner (e.g. for the synthesis of faces), the most compelling applications necessitate face analysis through fitting the model to observed data. This data may take many forms, such as the appearance of a face in one [4,3,9] or more [10,11] images, a noisy and incomplete 3D scan [12] or the location of a sparse set of feature points in an image [2].

When the objective function is underconstrained or ill-posed, the classical approach is to use Tikhonov regularisation (for a linear objective) or more generally to augment the objective function with a regularising term using a Lagrange multiplier. Typically, the regularisation term encourages smaller norms or equivalently, solutions closer to the mean face. With a suitable choice of the regularisation weight, this prevents overfitting and ensures that the resulting face is plausible. However, the optimal choice of regularisation weight may be different for different data samples. By choosing a conservative value, fitting results are likely to be too close to the mean face to capture features of the input face.

Much prior work uses such regularised optimisation approaches for face model fitting. For example linear regression [3], the inverse compositional algorithm [13], global optimisation [4], hybrid objective functions to encourage convexity [14] and alternating least squares for solving a multilinear system [15,16]. All of these approaches trade off satisfaction of a model-based prior against quality of fit. To ensure robust performance, these approaches must favour the prior, resulting in model dominance.

Recently, Brunton et al. [17] proposed a method to fit a statistical shape model to 3D data. They used a hard hyper box constraint, whereby each shape parameter was constrained to lie within $\pm k$ standard deviations of the mean. In other words, they assumed a uniform distribution over the hyper box as their prior. This has the advantage of being expressed as a linear inequality constraint on the parameters, enabling it to be incorporated into standard optimisation methods. Their hyper box is more conservative than the hyper-ellipsoid constraint that we propose here, with the two only intersecting at the corners of the hyper box. This is done so as to prevent extreme values of a single parameter being allowed by the constraint. We have not found this to be a problem in our experimental results and our manifold is motivated directly by the properties of assumed distribution over the parameters. Moreover, by assuming a uniform prior they do not discourage solutions close to the mean when the objective is over constrained.

There has been a recent interest in shape modelling on manifolds. Berkels et al. [18] show how to perform discrete geodesic regression on shape manifolds. This allows them to perform nonlinear regression in shape space according to a specified discrete path energy. For the specific case of the space of thin shells (including faces), Heeren et al. [19] provide a computational framework for calculating geodesics, allowing for plausible interpolations, averaging, and even shape extrapolation applications. In an altogether difference approach, Boscaini et al. [20] formulate shape interpolation and averaging in the space of Laplacians, from which shapes are subsequently reconstructed. Shapira and Ben-Chen [21] show how to align two face spaces (each corresponding to a different identity) by a non-rigid ICP between the corresponding manifold samples. This allows for shape analogies to be computed, providing a kind of expression transfer.

In this paper, we propose to solve the model fitting problem within the subspace of maximally likely faces. This requires the solution of an optimisation problem on a manifold. This problem has been considered previously in the medical imaging [22], signal processing [23], computer vision [24], robotics [25] and projective geometry [26] communities. Generic methods for optimisation on arbitrary manifolds have also been proposed [27]. In particular, the recently released Manopt toolbox [28] allows local optimisation on a number of manifolds through the expression of an objective and its gradient in the Euclidean embedding space. We focus on the case of a hyperspherical manifold and develop a hypherspherical gradient descent algorithm. In contrast to Manopt, our method operates in a coarse-to-fine manner in order to reduce susceptibility to local minima and exploits the closed nature of the manifold to reduce line searches to interval searches. We extend our previous presentation of this work [29] by demonstrating results on expression interpolation (Section 3.1) and underconstrained optimisation (Section 5.2), more thorough empirical evaluation of the manifold assumption and describing the theoretical ideas more thoroughly.

1.2. Outline

In Section 2 we begin by describing our statistical model and manifold. We first introduce tools from differential geometry which are necessary for developing our methodology and then provide empirical validation to justify our choice of manifold. In Section 3 we describe how warps and averages between two or more faces can be constrained to the manifold and compare the