Articulated Shape Matching Using Locally Linear Embedding and Orthogonal Alignment

Diana Mateus, Fabio Cuzzolin, Radu Horaud, and Edmond Boyer
INRIA Rhône-Alpes
655, avenue de l'Europe, Montbonnot Saint-Martin, FRANCE
Diana.Mateus@inrialpes.fr

Abstract

In this paper we propose a method for matching articulated shapes represented as large sets of 3D points by aligning the corresponding embedded clouds generated by locally linear embedding. In particular we show that the problem is equivalent to aligning two sets of points under an orthogonal *transformation acting onto the d-dimensional embeddings. The method may well be viewed as belonging to the* model-based clustering *framework and is implemented as an EM algorithm that alternates between the estimation of correspondences between data-points and the estimation of an optimal alignment transformation. Correspondences are initialized by embedding one set of data-points onto the other one through out-of-sample extension. Results for pairs of voxelsets representing moving persons are presented. Empirical evidence on the influence of the dimension of the embedding space is provided, suggesting that working with higher-dimensional spaces helps matching in challenging real-world scenarios, without collateral effects on the convergence.*

1. Introduction

Shape matching is a central problem in computer vision as it allows to find shape classes for object recognition, to track objects over time, to build spatio-temporal representations useful for shape modeling, for action and/or gesture recognition, etc. Although methods are available both for rigid objects and deformable surfaces, articulated shape matching remains a challenging problem. Rigidity (and hence isometry) is only locally preserved and knowledge about how an articulated shape is split into rigid pieces is often not available. Whenever an object is represented by a cloud of 2D or 3D points, matching two different poses of the same object reduces to establishing assignments between points of the two clouds. Situations involving occlusions, missing data, outliers, and noise have already been addressed within the framework of matching *rigid* objects. A number of methods were suggested including hypothesize-and-test implemented as tree search, iterative closest point (ICP), and probabilistic assignment.

In addition to the difficulties associated with rigid alignment mentioned above, the problem of articulated alignment is more complex for at least two reasons: (i) Shape isometry is preserved only locally and not globally and (ii) object sub-parts may collaps together, as it is the case when one arm lies along the torso. The lack of a global transformation that, in theory, maps points from one pose onto points of another pose, leads to consider the more general problem of *maximum subgraph matching*.

As we claim here, those obstacles can be overcome by *aligning the two clouds of points in an embedding space*. The first method known by us to solve the weighted graph matching problem through an eigen (or spectral) decomposition was proposed by Umeyama [**?**]. More recently, Several spectral methods were proposed [**?**, **?**] which compute non-linear embeddings of the input dataset by means of the SVD decomposition of an affinity matrix which depends on the structure of the data. ISOMAP, for instance, is based on the matrix of geodesic distances between points which are substantially preserved under articulated motion (factoring out changes in the topology of the moving body), as pointed out by other researchers [**?**]. By intuition, if the affinity matrix captures only local isometric properties of the input cloud, the shape of the resulting embedded cloud is only weakly affected by articulated deformations. In consequence, two clouds associated with different *articulated* poses of the same object can be *rigidly aligned in the embedding space*, such that each point of the first cloud is (in principle) associated with its nearest neighbor in the second (aligned) cloud.

In practice, as discussed here as well as in a companion paper [**?**], embedded shapes can only be aligned up to a $d \times d$ *orthogonal transformation*, where d is the dimensionality of the embedded space. Unlike its Euclidean sub-group, the orthogonal group does not have a Lie struc-