

# Probabilistic Clustering and Shape Modelling of White Matter Fibre Bundles Using Regression Mixtures

Nagulan Ratnarajah<sup>1</sup>, Andy Simmons<sup>2</sup>, and Ali Hojjatoleslami<sup>1</sup>

<sup>1</sup> Medical Image Computing, School of Biosciences, University of Kent, U.K.

<sup>2</sup> Neuroimaging Department, Institute of Psychiatry, King's College London, U.K.

**Abstract.** We present a novel approach for probabilistic clustering of white matter fibre pathways using curve-based regression mixture modelling techniques in 3D curve space. The clustering algorithm is based on a principled method for probabilistic modelling of a set of fibre trajectories as individual sequences of points generated from a finite mixture model consisting of multivariate polynomial regression model components. Unsupervised learning is carried out using maximum likelihood principles. Specifically, conditional mixture is used together with an EM algorithm to estimate cluster membership. The result of clustering is a probabilistic assignment of fibre trajectories to each cluster and an estimate of cluster parameters. A statistical shape model is calculated for each clustered fibre bundle using fitted parameters of the probabilistic clustering. We illustrate the potential of our clustering approach on synthetic and real data.

**Keywords:** Probabilistic Clustering, Regression Mixture, Fibre Tractography, Shape Model.

## 1 Introduction

White matter (WM) fibre clustering is becoming an important field of clinical neuroscience research since it facilitates insights about anatomical structures in health and disease, allows clear visualizations of fibre tracts and enables the calculation of relevant statistics across subjects. A number of algorithms have been developed for clustering and labelling WM fibre bundles in DTI. Deterministic clustering algorithms [1-3] assign each trajectory to only one cluster, which may lead to biased estimators of cluster parameters if the clusters overlap. Probabilistic clustering algorithms [4], on the contrary, deal with the inherent uncertainty in assigning the trajectories to clusters. Quantitative parameters can be estimated by a weighted average over cluster members and thus more robust results may be obtained, which are less sensitive to the presence of outliers. Maddah et al. [4] proposed a probabilistic approach using a gamma mixture model and a distance map. This method assumes that the number of clusters is known and the approach requires manual user initialisation of the cluster centres. A problem for this approach was establishing correspondence between points.