Cocluster analysis of thalamo-cortical fibre tracts extracted from diffusion tensor MRI

Cui Lin*, Shiyong Lu, Xuwei Liang and Jing Hua

Department of Computer Science,
Wayne State University,
5143 Cass Avenue, Detroit, MI, 48202, USA
E-mail: cuilin@wayne.edu
E-mail: shiyong@wayne.edu
E-mail: george.liang@wayne.edu
E-mail: jinghua@wayne.edu
*Corresponding author

Otto Muzik

Department of Pediatrics and Radiology,
Wayne State University Medical School,
Children's Hospital of Michigan PET Center,
3901 Beaubien Blvd Detroit, MI, 48201, USA
E-mail: otto@pet.wayne.edu

Abstract: As the central relay station of the human brain, the thalamus modulates sensory signals to and from the cerebral cortex. The reciprocal connectivity between the cerebral cortex and the thalamus is believed to play an essential role in consciousness and various neurological disorders. Thus, in-vivo analysis of thalamo-cortical connectivity is important for our understanding of normal and pathological brain processes. In this paper

- We propose a new partitioning paradigm, called *coclustering*, in order to segment the thalamus into thalamic nuclei based on their cortical projections. In contrast to the traditional clustering paradigm, a coclustering procedure not only simultaneously partitions cortical voxels and thalamic voxels into groups, but also identifies the corresponding strong connectivities between the two classes of groups
- We develop the first coclustering algorithm, Genetic Coclustering Algorithm (GCA), to solve the coclustering problem
- We apply GCA to segment the thalamus into thalamic nuclei and visualise main thalamo-cortical fibre tracts.

Keywords: coclustering; cortex; thalamus; fibre tracking; thalamo-cortical connectivity; data mining; bioinformatics.

Reference to this paper should be made as follows: Lin, C., Lu, S., Liang, X., Hua, J. and Muzik, O. (2008) 'Cocluster analysis of thalamo-cortical fibre tracts extracted from diffusion tensor MRI', *Int. J. Data Mining and Bioinformatics*, Vol. 2, No. 4, pp.342–361.

Biographical notes: Cui Lin is currently a PhD candidate in the Department of Computer Science at Wayne State University. She received her BE from Beijing Information Technology Institute, after which, she spent a few years working for IBM as a system analyst and project manager, and certified as an IBM DB2 expert. She is currently a member of the Scientific Workflow Research Laboratory (the SWR Lab). Her research interests include bioinformatics, scientific workflows, and data mining. She is a member of IEEE.

Shiyong Lu is currently an Assistant Professor in the Department of Computer Science at Wayne State University. He received his PhD from Stony Brook University in 2002 and BE from the University of Science and Technology of China at Hefei in 1993. His research interests include scientific workflows and medical informatics. He has published over 60 refereed international journal and conference papers in the above areas. He is the Founding Chair of the *IEEE International Workshop on Scientific Workflows* and an editorial board member for *International Journal of Medical Information Systems and Informatics*. He is a member of IEEE.

Xuwei Liang is a Graduate student of Computer Science in the Department of Computer Science, Wayne State University. He obtained his BS Degree at the Beijing Institute of Technology in 1986. He is now conducting research in the Graphics and Imaging Lab. His research interests include computer graphics, visualisation, and biomedical imaging especially in diffusion tensor imaging.

Jing Hua is an Assistant Professor of Computer Science at Wayne State University and the Director of Graphics and Imaging Lab. He received his PhD Degree (2004) in Computer Science from Stony Brook University. His research interests include computer graphics, visualisation, biomedical imaging and informatics. He has published over 80 journal and conference papers in the above areas. His research is currently funded by the NSF, the NIH, and the Michigan State Foundations. He serves as an editorial board member for *Scientific Journals International* and *International Journal of Technology Enhanced Learning* and a program committee member for many international conferences.

Otto Muzik received his PhD Degree in Physics in 1998 from the University of Vienna, Austria. He spent his postdoctoral years first at the Nuclear Research Center in Julich, Germany and subsequently at the University of Michigan in Ann Arbor working on quantitative methods in positron emission tomography (PET). In 1993, he joined the PET Research group at Children's Hospital of Michigan, Detroit. His research includes multimodality integration of advanced imaging techniques such as PET, MRI and DTI. He is Professor of Pediatrics and Radiology and PI or Co-Investigator on several NIH grants funded in the PET Imaging group.

1 Introduction

The thalamus is believed to relay information selectively to various parts of the cortex, as functionally specific neuronal clusters in the thalamus (thalamic nuclei) project to different cortical areas. Thalamic nuclei have strong reciprocal connections with the cerebral cortex, forming thalamo-cortical circuits that are believed to be involved in consciousness and play an important role in many neurological disorders (Bagary et al., 2002; Juhasz et al., 2001; Henderson et al., 2000; Banati et al., 2000; Creutzfeldt, 1995). Since a better understanding of the connectivity pattern between the cortex and thalamus is crucial for neuroscience in general and the treatment of neurological diseases in particular, a computational framework for the assessment of thalamic nuclei and their cortical projections is greatly needed.

Unfortunately, currently available methods for the in-vivo analysis of thalamo-cortical connectivity are still suboptimal. Existing brain analysis work either focuses on fibre tracking (which groups fibres into bundles (Corouge et al., 2004; Anders Brun et al., 2004; Ding et al., 2003; Brun et al., 2003; Shimony et al., 2003; Zhang and Laidlaw, 2002), or focuses on thalamic nuclei segmentation (Deoni et al., 2005; Jonasson et al., 2005; Wiegell et al., 2003; Imig and Morel, 1985). In this paper, we propose a method that not only supports the functionalities of both these methods but also supports the assessment of thalamo-cortical connectivity. As shown in Figure 1, each thalamic nuclei is mainly connected to a particular region of the cortex. For example, most voxels in thalamic nuclei T_2 are connected to voxels of cortical region C_2, with only few connections to other cortical regions, and vice versa. Therefore, in order to perform an accurate in-vivo analysis of the thalamo-cortical connections, what is needed is a partitioning procedure that not only simultaneously partitions cortical voxels and thalamic voxels into groups, but also identifies the corresponding strong connectivities between the two classes of groups. Traditional clustering algorithms, such as K-means (Han and Kamber, 2001), CHAMELEON (Karypis et al., 1999), and DBSCAN (Ester et al., 1996), are not applicable since they can only perform partitioning on one class of objects, either on cortical voxels, on fibres, or on thalamic voxels, without the consideration of the above anatomical connectivity constraint. As a result, they will fail to identify accurately the corresponding thalamo-cortical connectivity patterns.

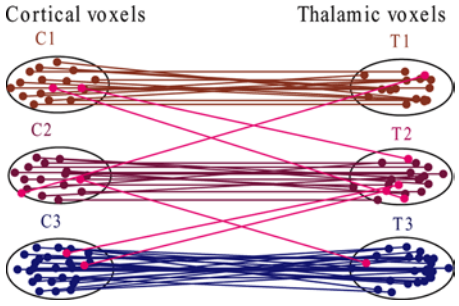
The main contributions of this paper are:

1 We propose a new partitioning paradigm, called *coclustering*, to model the thalamo-cortical connectivity analysis problem. In contrast to the traditional clustering paradigm, a coclustering procedure not only simultaneously partitions cortical voxels and thalamic voxels into groups, but also identifies the corresponding strong connectivities between the two classes of groups.

2 We develop the first coclustering algorithm, Genetic Coclustering Algorithm (GCA), to solve the coclustering problem.

3 We apply GCA to perform in-vivo analysis of the thalamo-cortical connections and produce a 3-D visualisation of seven thalamic nuclei groups together with their cortical projections.

Figure 1 Strong connectivity between a thalamic nuclei and its corresponding cortical region (see online version for colours)



This paper extends (Lin et al., 2006) with a comprehensive study of the effects of various operators and parameters on the convergence performance of the GCA algorithm (Section 5). First, the image acquisition and data preprocessing procedure is described in Section 5.1. Second, the evaluation of the effects of operators viz., selection, mutation and K-means, as well as parameters of population size and mutation probability on the convergence performance of the GCA algorithm is reported in Section 5.2. Finally, a 3-D visualisation of the thalamic nuclei as well as their connectivities to the corresponding cortical regions is presented in Section 5.3; this visualisation adds more detailed annotations to the one presented in Lin et al. (2006). We also extend (Lin et al., 2006) with a description of the background and related work in Section 2.

Organisation. The rest of the paper is organised as follows: Section 2 presents the background of this research and related work. Section 3 formalises the coclustering model with respect to thalamo-cortical connectivity analysis. Section 4 introduces our coclustering algorithm, the Genetic Coclustering Algorithm (GCA), to solve the coclustering problem. Section 5 presents experimental results obtained from analysis of real brain data. Finally, Section 6 concludes the paper and comments on future directions of this research.

2 Background and related work

As the central relay for a human brain, the thalamus modulates sensory signals to and from the cerebral cortex. Multiple functional pathways relay through the thalamus and form the thalamic cytoarchitecture, which is divided into functionally specific clusters that are referred to as thalamic nuclei (Wiegell et al., 2003). Thalamic nuclei have strong reciprocal connections with the cerebral cortex, forming the thalamo-cortical network. The pathological changes of such connections have been implicated in a large number of diseases, such as neocortical epilepsy (Juhasz et al., 2001), Parkinson's disease (Henderson et al., 2000), Schizophrenia (Bagary et al., 2002), and multiple sclerosis (Banati et al., 2000). These diseases can be treated by surgical ablation or electric stimulation, which requires both the study of the global connectivity and localisation of the place of individual structures within the overall scheme (Scannell et al., 1999).

Although an area of intense research, computational frameworks that efficiently model connectivity patterns in human brains are still suboptimal. For example,

Shimony et al. (2003) apply a fuzzy c-means algorithm in which each fibre is associated with a cluster by a membership function that indicates the degree that a fibre belongs to a cluster. Zhang and Laidlaw (2002) used a hierarchical clustering algorithm for the fibre clustering, Brun et al. (2003) consider fibres are similar to each other if they start from and end at the same area and define a measure that uses the distance between the end points. Other work focused on thalamus nuclei segmentation (Imig and Morel, 1985; Deoni et al., 2005; Jonasson et al., 2005; Wiegell et al., 2003) or cortex segmentation (Goldenberg et al., 2002). However, none of the above work provided an efficient computational framework to assess thalamo-cortical connectivity patterns.

Diffusion Tensor Imaging (DTI) is a Magnetic Resonance (MR) acquisition technique that measures the directional dependence of motion of water molecules in tissue (Moberts et al., 2005). DTI is a non-invasive technique which allows in vivo assessment of the internal structure of organised issue, such as white matter. In fibrous tissue of human brain white matter, water tends to diffuse less in the directions perpendicular to the fibre structure. This makes it possible to study the local fibre orientations indirectly by interpreting the water diffusion within the voxel (Anders Brun et al., 2004). Therefore, performing fibre tracking may provide valuable insights into fibre tract connectivity, by creating fibre traces from virtual particles, which travel along the direction of maximum diffusion, starting from a set of seed points (Corouge et al., 2004; Brun et al., 2003; Ding et al., 2001). DTI has been applied to a wide variety of fields, including neocortical epilepsy, brain development, brain tumor, and multiple sclerosis (Hesseltine et al., 2006; Tummala et al., 2003; Taber et al., 2002).

3 The coclustering model

In this section, we present our coclustering model, which models the thalamo-cortical analysis problem. In this model, the structure of the cortex and thalamus is represented as a bipartite graph $G = (C, T, F)$ as illustrated in Figure 1, where C is the set of cortical voxels, each of which is represented by a vector $\overrightarrow{X_n} = (x, y, x)$, encoding its three-dimensional coordinates; T is the set of thalamic voxels, represented by $\overrightarrow{Y_n} = (x, y, x)$; and F is the set of fibres connecting C and T. Although not required by our model, the working hypothesis is that most thalamic voxels within one thalamic nuclei usually connect to a specific cortical region, and the connectivities to other cortical regions are relatively weaker, and vice versa. The goal of a coclustering procedure is to group the two classes of objects simultaneously while minimising the cross-connectivity cost between them. More specifically, a coclustering procedure will partition both classes of objects into K groups so that

1 for each class, similar objects are within the same group, while dissimilar objects are in different groups
2 there is a one-to-one correspondence between the clusters in the first class (Class of cortical voxels) and the second class (Class of thalamic voxels); the corresponding cluster to a cluster is called its *spouse cluster*
3 the total cross-connectivity cost which results from those edges between a cluster in one class and a non-spouse cluster in the other class is minimised.

To achieve the above goals, we define several notions. First, we define the centroid of a cluster and its Within-Cluster Variation (WCV) to quantify the similarity of objects within one cluster.

The centroid of a cortical cluster C_K is defined as:

$$\overrightarrow{\mu_k} = \frac{\sum_{\overrightarrow{X_n} \in C_k} \overrightarrow{X_n}}{|C_k|}$$

where $|C_k|$ represents the number of cortical voxels in cluster C_k. Similarly, the centroid of a thalamic cluster T_k is defined as:

$$\overrightarrow{\nu_k} = \frac{\sum_{\overrightarrow{Y_n} \in T_k} \overrightarrow{Y_n}}{|T_k|}$$

where $|T_k|$ represents the number of thalamic voxels in cluster T_k.

The WCV of cortical cluster C_k is defined as:

$$\mathrm{WCV}(C_k) = \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} (X_{nd} - \mu_{kd})^2.$$

Similarly, the WCV of thalamic cluster T_k is defined as:

$$\mathrm{WCV}(T_k) = \sum_{\overrightarrow{Y_n} \in T_k} \sum_{d=1}^{D} (Y_{nd} - \nu_{kd})^2.$$

Second, we define the Total Within-Cluster Variation to quantify a particular partitioning. The Total Within-Cluster Variation (TWCV) of a cortical partition $(C_1, \ldots, C_K)$ is defined as:

$$\begin{aligned}
&\mathrm{TWCV}(C_1, \ldots, C_K) \\
&= \sum_{k=1}^{K} \mathrm{WCV}(C_k) \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} (X_{n_d} - \mu_{k_d})^2 \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} (X_{n_d})^2 + \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} (\mu_{k_d})^2 - \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} (2 X_{n_d} \mu_{k_d}) \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} {X_{n_d}}^2 + \sum_{k=1}^{K} |C_k| \sum_{d=1}^{D} (\mu_{k_d})^2 - \sum_{k=1}^{K} \sum_{d=1}^{D} (2 \times \mathrm{SCF}_{k_d} \times \mu_{k_d}) \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} {X_{n_d}}^2 + \sum_{k=1}^{K} \sum_{d=1}^{D} \frac{(\mathrm{SCF}_{k_d})^2}{|C_k|} - \sum_{k=1}^{K} \sum_{d=1}^{D} \frac{2 \times {\mathrm{SCF}_{k_d}}^2}{|C_k|} \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{X_n} \in C_k} \sum_{d=1}^{D} {X_{n_d}}^2 - \sum_{k=1}^{K} \frac{1}{|C_k|} \sum_{d=1}^{D} (\mathrm{SCF}_{k_d})^2
\end{aligned}$$

where SCF_{k_d} is the sum of the dth features of all voxels in C_k. Similarly, the TWCV of a thalamic partition $(T_1, \ldots, T_K)$ is defined as:

$$\begin{aligned}
&\mathrm{TWCV}(T_1, \ldots, T_K) \\
&= \sum_{k=1}^{K} \mathrm{WCV}(T_k) \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{Y_n} \in T_k} \sum_{d=1}^{D} (Y_{n_d} - \nu_{k_d})^2 \\
&= \sum_{k=1}^{K} \sum_{\overrightarrow{Y_n} \in T_k} \sum_{d=1}^{D} {Y_{n_d}}^2 - \sum_{k=1}^{K} \frac{1}{|T_k|} \sum_{d=1}^{D} (\mathrm{STF}_{k_d})^2,
\end{aligned}$$

where STF_{k_d} is the sum of the dth features of all voxels in T_k.

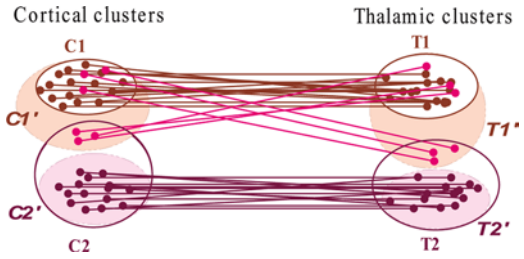
Third, in order to minimise the cross-connectivity cost, for each thalamic cluster, we define the set of cortical voxels that are connected to it as its *shaded thalamic cluster*. More formally, given a thalamic partition $(T_1, \ldots, T_K)$, the shaded cortical cluster C'_K $(k = 1, \ldots, K)$ is defined as:

$$C'_k = \{c \,|\, c \in C, \exists t \in T_K, (c, t) \in F\};$$

A shaded cortical cluster T'_k can be defined similarly.

For example, in Figure 2, all the cortical voxels that are connected to voxels in thalamic cluster T_1 forms the shaded cortical cluster C'_1, while all thalamic voxels that are connected to the voxles in cortical cluster C_1 forms the shaded thalamic cluster T'_1. In an ideal coclustering, as C'_2 and T'_2, a shaded cluster should *coincide* with the corresponding spouse cluster. However, this is not always the case in general. The cross-connectivity cost can be characterised by the disagreement between shaded clusters and spouse clusters and quantified by the total within cluster variance of shaded clusters with respect to their corresponding spouse clusters, called STWCV, which is defined as follows.

Figure 2 Shaded cortical and thalamic clusters (see online version for colours)



The Shaded Within-Cluster Variation (SWCV) of cortical cluster C'_K is defined as:

$$\mathrm{SWCV}(C'_k) = \sum_{\overrightarrow{X'_n} \in C'_k} \sum_{d=1}^{D} (X'_{n_d} - \mu_{k_d})^2.$$

Note that, instead of using the centroid of C'_k, the centroid of the C_k is used to calculate SWCV(C'_k). The intuition is that, in an ideal partitioning, the shaded partition $C'_1 \ldots C'_K$ should mostly coincide with $C_1 \ldots C_K$.

Similarly to the shaded cortical partition, the centroid of the T_k is used to calculate SWCV(T'_k) for finding the optimal solution. The SWCV of thalamic cluster T'_k is defined as:

$$\text{SWCV}(T'_k) = \sum_{\overrightarrow{Y'_n} \in T'_k} \sum_{d=1}^{D} (Y'_{n_d} - \nu_{k_d})^2.$$

The Shaded Total Within-Cluster Variation (STWCV) of cortical partition $(C'_1, \ldots, C'_K)$ is defined as:

$$\text{STWCV}(C'_1, \ldots, C'_K) = \sum_{k=1}^{K} \text{SWCV}(C'_k) = \sum_{k=1}^{K} \sum_{\overrightarrow{X'_n} \in C'_k} \sum_{d=1}^{D} (X'_{n_d} - \mu_{k_d})^2.$$

Similarly, the STWCV of thalamic partition $(T'_1, \ldots, T'_K)$ is defined as:

$$\text{STWCV}(T'_1, \ldots, T'_K) = \sum_{k=1}^{K} \text{SWCV}(T'_k) = \sum_{k=1}^{K} \sum_{\overrightarrow{Y'_n} \in T'_k} \sum_{d=1}^{D} (Y'_{n_d} - \nu_{k_d})^2.$$

Statement of the problem: Finally, the coclustering problem can be formally stated as follows: given a cluster number K, a bipartite $G = (C, T, F)$, and a distance metric d for nodes in C and T, partition (C, T) into K cluster pairs (C_1, T_1), $(C_2, T_2), \ldots,$ (C_K, T_K), such that the following objective function OTWCV is minimised:

$$\begin{aligned}\text{OTWCV} = {} & \text{TWCV}(C_1, \ldots, C_K) + \text{TWCV}(T_1, \ldots, T_K) \\ & + \text{STWCV}(C'_1, \ldots, C'_K) + \text{STWCV}(T'_1, \ldots, T'_K).\end{aligned}$$

4 Our proposed GCA algorithm

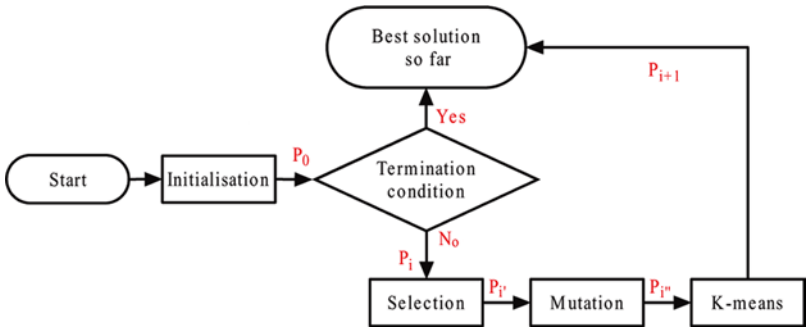
In this section, we propose the first coclustering algorithm, Genetic Coclustering Algorithm (GCA), to solve the coclustering problem. It is based on the genetic algorithm approach (Maulik and Bandyopadhyay, 1994) by working on a coding of the solution space over which the search has to be performed. These encoded solutions are called *chromosomes* and the objective function value OTWCV can be calculated for each solution according to its definition in Section 3. Each solution is encoded by a string of symbols and GCA evolves solutions over generations. During each generation, GCA produces a new population from the current population by applying genetic operators viz., selection, mutation, and K-means operator. Each solution in the population is associated with a figure of merit (fitness value) depending on OTWCV. The selection operator selects a solution from the current population for the next population with a probability proportional to its fitness value. The mutation operator toggles each position in a string with a probability, called the

Mutation Probability (MP). Finally the K-means operator is introduced to speed up GCA's convergence to the global optimum (Krishna and Murty, 1999). The rational that we use a genetic algorithm approach is that it has been shown that genetic algorithms that maintain the best discovered solution either before or after the selection operator asymptotically can converge to the global optimisation (Rudolph, 1994).

More specifically, GCA maintains a population (set) of Z coded solutions, where Z is a parameter specified by the user. Each solution is coded by a string $(c_1 \ldots c_N, t_1 \ldots t_N)$ for N cortical voxels and N thalamic voxels, where each c_i or t_i, called an *allele*, denotes a cortical voxel or a corresponding thalamic voxel. Each allele takes a value from $1, \ldots, K$, representing the cluster number to which the voxel belongs. For example, $(c_1c_2c_3c_4c_5, t_1t_2t_3t_4t_5) = (21332, 21345)$ encodes a bicluster partitioning of five cortical voxels and five thalamic voxels, in which c_1 and c_5 belong to cortical cluster 2, while their connected thalamic voxels t_1 and t_5 belong to thalamic clusters 2 and 5, respectively.

Figure 3 illustrates the flowchart of GCA. It starts with the initialisation phase, which generates the initial population P_0, and then run the selection operator, the mutation operator, and the K-means operator sequentially on the current population P_i to obtain the next generation P_{i+1}. This sequence of operators are run iteratively to produce one generation after another until a termination condition is reached.

Figure 3 Flowchart of the GCA algorithm (see online version for colours)



During evolution, some solutions in which some cortical clusters or thalamic clusters are empty might be produced. These solutions are called *illegal* solutions. As will be shown later, although illegal solutions are not needed eventually, they are helpful for GCA's convergence to the global optimum. To deal with illegal solutions, we define the notion of *legality ratio*. Given a solution S_z that is encoded by $(c_1 \ldots c_N, t_1 \ldots t_N)$, let n_1 be the number of nonempty cortical clusters and n_2 be the number of nonempty thalamic clusters in S_z, the *legality ratio* of S_z is defined as:

$$e(S_z) = (n_1 + n_2)/(2 \times K).$$

A solution S_z is *legal* if $e(S_z) = 1$ and *illegal* otherwise. For example, given $K = 4$, the solution $c_1c_2c_3c_4c_5, t_1t_2t_3t_4t_5 =$ '43223, 43223' is illegal since cortical cluster 1 is empty. We describe our design of the operators in the sequel.

4.1 Phase 1: initialisation operator

GCA starts with the initialisation phase, which randomly generates the initial population P_0 of Z solutions, where Z is a parameter specified by the user. Each allele c_i in a solution $(c_1 \ldots c_N, t_1 \ldots t_N)$ is initialised to a cluster number randomly selected from the uniform distribution over the set $\{1, 2, \ldots, K\}$; t_i is initialised to a value equal to its corresponding c_i.

Illegal solutions are permitted but are considered as the most undesirable solutions by defining their OTWCVs as $+\infty$ and assigning them with low fitness values (will be defined in the next subsection). With a low fitness value, an illegal solution will have a lower probability for survival. This flexibility of allowing illegal solutions in the evolution process not only avoids the overhead of illegal solution elimination and thus improves the time performance of the algorithm, but also provides the chances for illegal solutions to mutate to legal solutions during the mutation phase.

4.2 Phase 2: the selection operator

We use proportional selection for the selection operator in which, the population of the next generation is determined by Z independent random experiments. Each experiment randomly selects a solution from the current population $(S_1, S_2, \ldots, S_z)$ according to the probability distribution $(p_1, p_2, \ldots, p_Z)$ defined by

$$p_z = \frac{F(S_z)}{\sum_{z=1}^{Z} F(S_z)}$$

where $F(S_z)$ denotes the fitness value of solution S_z with respect to the current population and is defined as follows.

$$F(S_z) = (\mathrm{OTWCV}_{\max} - \mathrm{OTWCV}_{S_z}) \times e(S_z)$$

where $\mathrm{OTWCV}_{\max}$ is the maximal value of OTWCV that has been encountered till the present generation.

The intuition behind this fitness function is that each solution will have a probability to survive by being assigned with a positive fitness value, but a solution with a smaller OTWCV has a greater fitness value and hence has a higher probability to survive. Illegal solutions are also allowed to survive but with lower fitness values than all legal solutions in the current population. Finally, illegal strings that have more empty clusters are assigned with smaller fitness values and hence have lower probabilities for survival.

4.3 Phase 3: the mutation operator

The mutation operator is very useful for GCA to reach better solutions based on the evolutional theory that offsprings produced by mutations might be superior to their parents. More importantly, the mutation operator performs the functionality of shaking the algorithm out of a local optimum and of moving it towards the global optimum (Rudolph, 1994).

Given a solution $S_z = (c_1 \ldots c_N, t_1 \ldots t_N)$, the mutation operator mutates each allele c_i or t_i to new values k_1 and k_2 simultaneously (might be equal to $c_i t_i$), where

k_1 and k_2 are numbers randomly selected from $(1, 2, \ldots, K)$ with probability MP respectively and independently where $0 < \mathrm{MP} < 1$ is a parameter called the *Mutation Probability* which is specified by the user. To define the probability mass function of mutation, we firstly define the closeness measurement between cortical voxels $\overrightarrow{X_i}$ and cortical cluster C_{k_1} as

$$CS(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}}) = \max_{k=1}^{K}\{d(\overrightarrow{X_i}, \overrightarrow{\mu_k})\} - d(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}}),$$

where $d(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}})$ is the Euclidean distance between cortical voxel $\overrightarrow{X_i}$ and the centroid $\overrightarrow{\mu_{k_1}}$ of the k_1th cortical cluster.

The closeness measurement between thalamic voxel Y_i and thalamic cluster C_{k_2} is defined by

$$TS(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}}) = \max_{k=1}^{K}\{d(\overrightarrow{Y_i}, \overrightarrow{\nu_k})\} - d(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}})$$

where $d(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}})$ is the Euclidean distance between thalamic voxel $\overrightarrow{Y_i}$ and the centroid $\overrightarrow{\nu_{k_2}}$ of the k_2th cluster on thalamus.

During mutation, we replace each allele $c_i t_i$ by $k_1 k_2$ for $i = (1, \ldots, N)$ simultaneously where k_1 and k_2 are selected from $(1, \ldots, K)$ with the probability distribution $(p_{11}, p_{12}, \ldots, p_{k_1 k_2}, \ldots, p_{KK})$ in which

$$p_{k_1 k_2} = \frac{CS(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}}) + CS(\overrightarrow{X_i}', \overrightarrow{\mu_{k_2}}) + TS(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}}) + TS(\overrightarrow{Y_i}', \overrightarrow{\nu_{k_1}})}{\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}(CS(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}}) + CS(\overrightarrow{X_i}', \overrightarrow{\mu_{k_2}}) + TS(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}}) + TS(\overrightarrow{Y_i}', \overrightarrow{\nu_{k_1}}))}.$$

The distance between a voxel and an empty cluster is defined to be 0 to increase the chance of converting an illegal solution to a legal one. The above mutation operator is defined such that

1 $\overrightarrow{X_i}$ and $\overrightarrow{Y_i}$ might be reassigned randomly to each cluster with a positive probability

2 the probability of changing allele value $c_i t_i$ to a cluster number $k_1 k_2$ is greater if $\overrightarrow{X_i}$ and $\overrightarrow{Y_i}$ are closer to the centroid of the k_1th cortical cluster and the k_2th thalamic cluster.

The first property ensures that an arbitrary solution, including the global optimum, might be generated by the mutation from the current solution with a positive probability; the second property encourages that each $\overrightarrow{X_i}$ and $\overrightarrow{Y_i}$ are moving towards a closer cluster with a higher probability.

4.4 Phase 4: the K-means operator

In order to speed up the convergence process, we introduce the following K-means operator based on the idea of the classical K-means algorithm (Han and Kamber, 2001). Even though there have been many genetic algorithms (Bhuyan et al., 1991; Jones and Beltramo, 1991) that may converge to the global optimum because of mutation operator, they employed either an expensive crossover operator to generate valid child chromosomes from parent chromosomes or a costly fitness function or both (Krishna and Murty, 1999). To address this issue, GCA hybridises the genetic

algorithm with k-means operator and replaces crossover operator, resulting in an algorithm that combines the effectiveness and simplicity of k-means and robustness of GA. In addition, We treat illegal solutions and legal solutions separately.

For an illegal solution $S_z = (c_1 \ldots c_N, t_1 \ldots t_N)$, we replace each $c_i t_i$ by new values $k_1 k_2$ for $i = 1, \ldots, N$ simultaneously, where k_1 and k_2 are numbers selected from $(1, 2, \ldots, K)$ such that the value returned by the following τ function is minimised.

$$\tau(k_1, k_2) = d(\overrightarrow{X_i}, \overrightarrow{\mu_{k_1}}) + d(\overrightarrow{X_i'}, \overrightarrow{\mu_{k_2}}) + d(\overrightarrow{Y_i}, \overrightarrow{\nu_{k_2}}) + d(\overrightarrow{Y_i'}, \overrightarrow{\nu_{k_1}}).$$

The distance between a voxel and an empty cluster centroid is defined to be 0 with the effort to convert an illegal solution to a legal one. For a legal solution $S_z = (c_1 \ldots c_N, t_1 \ldots t_N)$, we replace each $c_i t_i$ by new values kk for $i = 1, \ldots, N$ simultaneously, where k is a number selected from $(1, 2, \ldots, K)$ such that the value returned by the following λ function is minimised:

$$\lambda(k) = d(\overrightarrow{X_i}, \overrightarrow{\mu_k}) + d(\overrightarrow{X_i'}, \overrightarrow{\mu_k}) + d(\overrightarrow{Y_i}, \overrightarrow{\nu_k}) + d(\overrightarrow{Y_i'}, \overrightarrow{\nu_k}).$$

Here, we replace $c_i t_i$ to the same new cluster number to reduce cross-connectivity cost.

5 Experimental results

In this section, first, we present how we acquire the datasets based on medical imaging techniques; second, we provide our study of the convergence performance of GCA; and finally, we present our 3-D visualisation of the analysis results.

5.1 Image acquisition and data preprocessing

MRI data was acquired on a GE 1.5 Tesla Signa unit (GE Medical Systems, Milwaukee, Wisconsin). Initially the DTI sequence consists of an image volume with no diffusion weighting ($b = 0\,\mathrm{s/mm^2}$) followed by the acquisition of image volumes in six gradient directions ($[1, 0, 1], [-1, 0, 1], [0, 1, 1], [0, -1, 1], [1, 1, 0], [1, -1, 0]$) with a b-value of $1000\,\mathrm{s/mm^2}$. For each b-value and gradient direction, six images were acquired and magnitude averaging was used to avoid artifacts from subject motion.

In order to test the performance of an automated source/target definition procedure for the creation of fibre tracts connecting different parts of the brain, we have designed a prototype software environment, which allows the coregistration of each of the seven gradient volumes obtained from the DTI acquisition sequence independently to a high-resolution SPGR image volume. Independent coregistration of each gradient volume is often necessary due to patient motion. Moreover, the gradient image volumes are resampled in order to match the resolution of the SPGR image volume. Subsequently, fibre tracts were calculated based on the direction of the principal eigenvector in each image voxel. The calculation of fibre tracts requires the following constraints:

1 fibre tracts originate only from those voxels with an FA value of > 0.2

2 tracking is terminated when it reaches a voxel with FA < 0.2

3 the fibre tract is terminated when the deviation between adjacent eigenvectors is larger than 75 degrees.

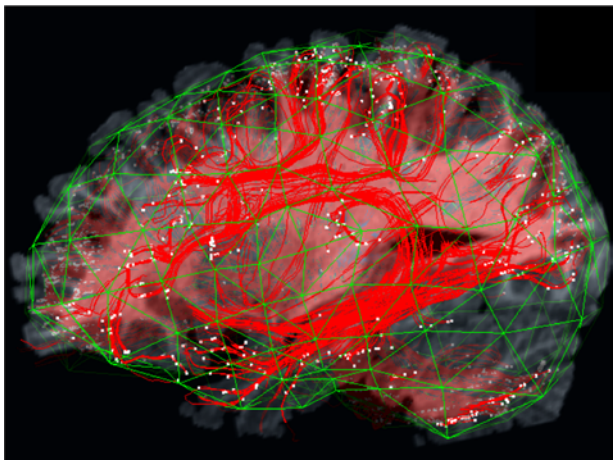
Moreover, the Fractional Anisotropy (FA) and Apparent Diffusion Coefficient (ADC) of each fibre tract are calculated as follows:

$$\mathrm{FA} = \sqrt{\left(3\sum_{i=1}^{3}(\lambda_i - \lambda_{\mathrm{mean}})^2\right) \Big/ \left(2\sum_{i=1}^{3}\lambda_i^2\right)}$$

$$\mathrm{ADC} = \left(\sum_{i=1}^{3}\lambda_i\right)/3$$

where λ_i is the eigenvalue calculated on a voxel-by-voxel basis from the Stejskal-Tanner equation (Stejskal and Tanner, 1965). The FA is based on the normalised variance of the eigenvalues and shows the differences between an isotropic diffusion (where the diffusion tensor is represented by a sphere) and a linear diffusion ('cigar-shape' ellipsoid). Its range is between 0 and 1 with 0 representing an isotropic diffusion and 1 representing a highly directional diffusion. Finally, ADC equals the mean eigenvalue and characterises the diffusivity of an image voxel. All fibre tracts that originate in a source volume can be displayed and inspected in the 3D space. Figure 4 shows a representative rendering of a portion of fibre tracts obtained using tracking of fibres within a brain.

Figure 4 Fibre tracts detected with diffusion tensor imaging data. Fibres (red) are shown relative to a sagittal cut through a T1-weighted MR image volume. Fibre end points on the cortical surface are rendered as white dots (see online version for colours)



5.2 GCA convergence performance evaluation

Our experiments were conducted on a Dell Dimension 8200 PC machine with 2.4 GHz CPU and 512M RAM. GCA was implemented in C with Microsoft Visual Studio. As GCA has very fast convergence speed (within 10 minutes for running 80 generations), no large generation number G is required for reaching a convergence value. Therefore, we only need to focus on the effect of each operator and other two parameters on our coclustering results. In each case, GCA was run for ten times and the average values were recorded in Figures 5 and 6.

5.2.1 Effects of K-means on GCA's convergence performance

Figure 5(a) and (b) shows the effects of K-means operator on GCA's convergence performance. Figure 5(b) is a zoom-in version of Figure 5(a) without case 3 that is described below. In this study, we choose population size $Z = 200$, mutation probability $\mathrm{MP} = 0.1$, and generation number $G = 80$. We compare four cases:

1 *case 1*: GCA with illegal solutions elimination during the selection phase

2 *case 2*: GCA without the mutation phase

3 *case 3*: GCA without the K-means phase

4 *case 4*: GCA with all phases.

Figure 5 (a) Effects of K-means on GCA's convergence performance and (b) effects of selection and mutation on GCA's convergence performance (see online version for colours)

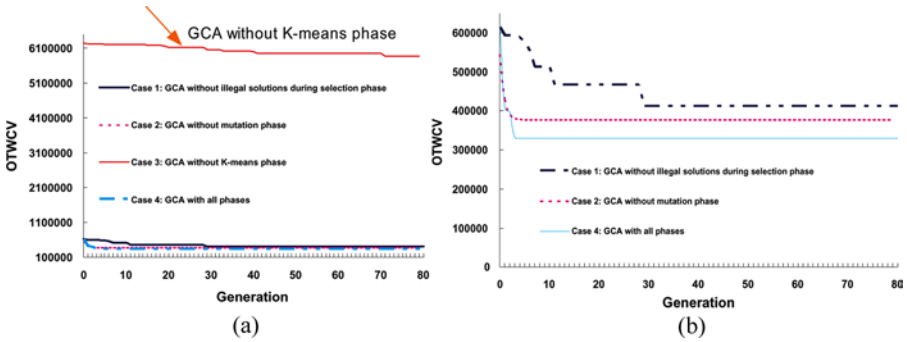
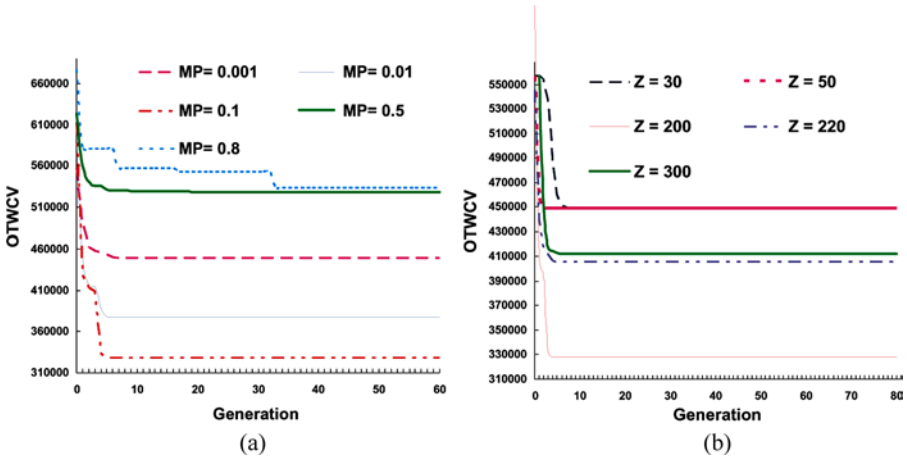


Figure 6 (a) Effects of population size on GCA's convergence performance and (b) effects of mutation probability on GCA's convergence performance (see online version for colours)



As shown in Figure 5, GCA with all the phases has the best convergence performance after 80 generations with OTWCV of 328520.6199. The K-means operator is critically

important for the convergence of GCA since without K-means, GCA hardly converges and ends up with the largest OTWCV value of 5868383.83 among the four cases, which is 17 times greater than the convergence of the best case – GCA with all phases.

5.2.2 *Effects of selection on GCA's convergence performance*

Figure 5(b) shows the effects of the selection phase on GCA's convergence performance. It demonstrates the benefits of allowing illegal solutions to survive with low probabilities. First, this strategy avoids additional computational overhead for eliminating illegal solutions. Second, it increases the convergence speed and improves the time performance of the algorithm. Finally, such a strategy promotes the chance of converting illegal solutions to legal ones, which might lead to the best solution after mutation and K-means phases. Figure 5(b) shows if illegal solutions elimination is not part of the selection phase, then after 80 generations, GCA converges to the OTWCV value of 411885, which is worse than case 2 and case 4.

5.2.3 *Effects of mutation on GCA's convergence performance*

Figure 5(b) shows that the mutation phase improves the convergence performance of GCA. In particular, after 80 generations, the OTWCV value derived by the GCA without the mutation phase is around 14% greater than the one derived by the GCA with all phases. The reason is that the mutation phase performs the functionality of shaking the algorithm out of a local optimum and of moving it towards the global optimum. In our experiment, the curve for the GCA without the mutation phase remains a fixed value after reaching a turning point (between 388537.2691 and 395949.7706), while the curve for the GCA with all phases continues to converge to a much smaller OTWCV value of 328520.6199.

5.2.4 *Effects of population size on GCA's convergence performance*

Figure 6(a) illustrates how the population size affects the convergence performance of GCA. Specifically, there are five curves with $\mathrm{MP} = 0.1$ and $G = 80$, each curve varies on the value of Z. Our experiments showed that tuning the value of Z has no distinct effects on the convergence speed and all of the five curves remain similar shapes. However, it significantly changes the convergence values, from 449419.6154 ($Z = 30$) to 448844.382 ($Z = 50$), and then OTWCV reduces to 328520.6199 when Z is equal to 200. After that, the convergence value rises up from 406181.9956 ($Z = 220$) to 412310.9311 ($Z = 300$) within 80 generations. Afterwards, the computation time become gradually intolerable as Z increases.

According to Grefenstette (1986), Goldberg et al. (1992a, 1992b) and Smith (1993), for a fixed number of generations of genetic algorithm, there exists an optimal population size under which the algorithm has the best convergence performance; therefore, larger population size does not necessarily imply better convergence performance. This explains our observation that the convergence performance is better for a population of 200 than a population of 300.

5.2.5 *Effects of mutation probability on GCA's convergence performance*

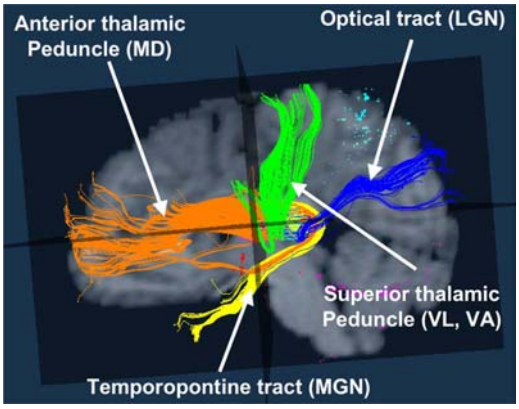
Figure 6(b) illustrates the effects of mutation probability on the convergence performance of GCA. Five curves are drawn with $Z = 100$ and $G = 80$, each curve

varies on the value of MP. Obviously, the shape of the curve with $\mathrm{MP} = 0.8$ is not as smooth as any other curves. It is because when mutation probability is relatively high, OTWCV may drop down sharply in one generation, although the convergence value does not fall into the lowest within 60 generations. As MP increases from 0.01 to 0.1, the convergence value tends to decrease from 448677.5381 to 377731.778663 and then to 328520.619907, which is the best convergence value we observed so far. After that, as MP continues to grow to 0.5, OTWCV begins to increase to 527926.9647, being apart from the best convergence value.

5.3 3-D visualisation of the GCA results

The number of thalamic nuclei varies with the histological methods employed. Here we partition the thalamus into seven clusters, and five of them can be automatically identified, which correspond to five functionally defined nuclear groups: Medial Dorsal nucleus (MD), Lateral Geniculate Nucleus (LGN), Medial Geniculate Nucleus (MGN), Ventral Lateral nucleus (VL), Ventral Anterior nucleus (VA). Figure 7 shows segmentation of the thalamus based on the connectivity of thalamic voxels to different cortical projection areas. Using the proposed coclustering method we were able to identify thalamic voxels belonging to the MD, which project towards areas of the frontal lobe (e.g., cingulate gyrus) via the anterior thalamic peduncle. Thalamic voxels comprising both the VL and VA could be clustered based on their projection (via the superior thalamic peduncle) to the primary motor cortex (M1) as well as premotor cortex (PMC). Moreover, thalamic voxels representing the LGN were clustered according to their posterior projections to the occipital cortex via the optical tract. Finally, voxels in the thalamus representing the MGN could be identified through their projections to the middle and superior gyri of the temporal lobe via the temporopontine tract.

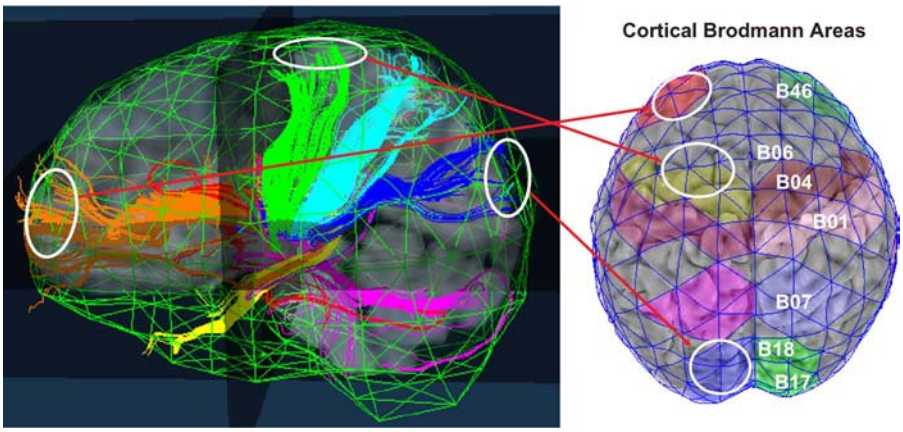
Figure 7 Segmentation of the thalamus based on projection areas of thalamic nuclei to the cortex. Strong connectivity between thalamic voxels and the cortex could be determined for the Medial Dorsal nucleus (MD), Lateral Geniculate Nucleus (LGN), Medial Geniculate Nucleus (MGN) as well as for both the Ventral Lateral nucleus (VL) and Ventral Anterior nucleus (VA) (see online version for colours)



The confirmation of our results was obtained through visual comparison of the obtained thalamic cortical projections to well-known Brodmann areas of the cortex.

Brodmann areas represent region of the cortex defined based on in vivo tissue staining which have been shown to roughly delineate functionally distinct areas of the cortex. As the various thalamic nuclei connect to different functional areas of the cortex, results of the GCA can be then judged on how well the projection areas of thalamo-cortical fibre tracts overlap with the corresponding Brodmann areas which have been independently identified on the cortical surface based on their spatial relationship to cortical landmarks. As an example, the following Brodmann areas in Figure 8 were considered: B46 – dorsolateral prefrontal cortex (DLPFC), B04 – primary motor cortex, B06 – supplementary motor cortex, B01 – primary somatosensory cortex, B07 – somatosensory association cortex, B17 – primary visual cortex and B18 – visual association cortex. Figure 8 shows the superior thalamic peduncle (green) projecting to B04 and B06 which are colour-coded on the cortical mesh. Similarly, the optical tract (dark blue) projects to B17 and B18. Finally, the anterior thalamic peduncle (orange) represents fibres originating from the MD and projecting partly to B46. Although short of an exact mathematical validation, these results indicate that our coclustering algorithm is consistent with anatomical data and warrants the application of this method in subsequent more clinically oriented studies.

Figure 8 The left panel shows the spatial relationship of thalamo-cortical fibre tracts with respect to finite cortical surface elements derived using landmark-constrained conformal mapping (Zou et al., 2006). Sets of finite elements were assigned to multiple Brodmann areas based on their spatial proximity to easily recognised cortical landmarks such as the central sulcus, sylvian fissure and the parieto-occipital sulcus. The right panel shows a superior view of the cortex with Brodmann areas colour-coded bilaterally. The validation of our clustering results was achieved by comparing the overlap between fibre projection areas (white circles) and independently defined Brodmann areas (see online version for colours)



6 Conclusions and future work

In this paper, we defined the coclustering problem and developed a coclustering algorithm, the Generic Coclustering Algorithm (GCA), which we then applied to the in vivo analysis of thalamo-cortical connectivity patterns. Our results showed that our

algorithm is able to segment the thalamus into seven thalamic nuclei groups based on their connection to functionally distinct cortical areas.

Although the coclustering problem is motivated by the need to assess thalamo-cortical connectivity patterns, we expect that it will have a wide range of applications. For example, in social science, our coclustering technique might be used in identifying social communities, in which husbands and wives can belong to the same or different social communities. Moreover, we anticipate that GCA might be useful in bioinformatics, where it can be used to identify families of transcription factors and their corresponding binding sites in the upstream promoter regions of genes.

References

Anders Brun, J.S., Knutsson, H., Park, H-J., Shenton, M.E. and Westin, C-F. (2004) 'Clustering fiber traces using normalised cuts', *Medical Image Computing and Computer-Assisted Intervention*, Vol. 1, pp.368–375.

Bagary, M.S., Foong, J., Maier, M., duBoulay, G., Barker, G.J., Miller, D.H. and Ron, M.A. (2002) 'A magnetisation transfer analysis of the thalamus in schizophrenia', *The Journal of Neuropsychiatry and Clinical Neurosciences*, Vol. 14, pp.443–448.

Banati, R.B., Newcombe, J., Gunn, A., Turkheimer, F., Heppner, F., Price, G., Wegner, F., Giovannoni, G., Miller, D.H., Perkin, G.D., Smith, T., Hewson, A.K., Bydder, G., Kreutzberg, G.W., Jones, T., Cuzner, M.L. and Myers, R. (2000) 'The peripheral benzodiazepine binding site in the brain in multiple sclerosis', *Brain*, Vol. 123, pp.2321–2337.

Bhuyan, J.N., Raghavan, V.V. and Elayavalli, V.K. (1991) 'Genetic algorithm for clustering with an ordered representation', *Proc. 4th Int. Conf. Genetic Algorithms*, Morgan Kaufman, San Mateo, CA.

Brun, A., Park, H-J., Knutsson, H. and Westin, C-F. (2003) 'Coloring of DT-MRI fiber traces using laplacian eigenmaps', *Lecture Notes in Computer Science*, Vol. 2809, pp.716–721.

Corouge, I., Gouttard, S. and Gerig, G. (2004) 'Towards a shape model of white matter fiber bundles using diffusion tensor MRI', *International Symposium on Biomedical Imaging, Conf. Proc.*, pp.344–347.

Creutzfeldt, O.D. (1995) *Cortex Cerebri: Performance, Structural and Functional Organisation of the Cortex*, Oxford University Press, Oxford.

Deoni, S.C.L., Josseau, M.J.C., Rutt, B.K. and Peters, T.M. (2005) 'Visualisation of thalamic nuclei on high resolution, multi-averaged T1 and T2 maps acquired at 1.5 T', *Human Brain Mapping*, Vol. 25, pp.353–359.

Ding, Z., Gore, J.C. and Anderson, A.W. (2001) 'Reconstruction, visualisation and quantification of neuronal fiber pathways', *Proceedings of the Conference on Visualisation 2001*, pp.453–456.

Ding, Z., Gore, J.C. and Anderson, A.W. (2003) 'Classification and quantification of neuronal fiber pathways using diffusion tensor MRI', *Magn. Res. Med.*, Vol. 49, pp.716–721.

Ester, M., Kriegel, H-P., Sander, J. and Xu, X. (1996) 'A density-based algorithm for discovering clusters in large spatial databases with noise', *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining*, pp.226–231.

Goldenberg, R., Kimmel, R., Rivlin, E. and Rudzsky, M. (2002) 'Cortex segmentation: a fast variational geometric approach', *IEEE Trans on Medical Imaging*, Vol. 21, No. 12, pp.1544–1551.

Goldberg, D.E., Deb, K. and Clark, J.H. (1992a) 'Accounting for noise in the sizing of populations', in Whitley, L.D. (Ed.): *Foundations of Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, pp.127–140.

Goldberg, D.E., Deb, K. and Clark, J.H. (1992b) 'Genetic algorithms, noise, and the sizing of populations', *Complex Systems*, Vol. 6, pp.333–362.

Grefenstette, J. (1986) 'Optimisation of control parameters for genetic algorithms', *IEEE Trans. Syst. Man Cybern.*, Vol. 16, No. 1, pp.122–128.

Han, J. and Kamber, M. (2001) *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, pp.349–353.

Henderson, J.M., Carpenter, K., Cartwright, H. and Halliday, G.M. (2000) 'Loss of thalamic intralaminar nuclei in progressive supranuclear palsy and parkinson's disease: clinical and therapeutic implications', *Brain*, Vol. 123, No. 7, pp.1410–1421.

Hesseltine, S.M., Lawb, M., Babba, J., Rada, M., Lopeza, S., Gea, Y., Johnsona, G. and Grossmana, R.I. (2006) 'Diffusion tensor imaging in multiple sclerosis: assessment of regional differences in the axial plane within normal-appearing cervical spinal cord', *American Journal of Neuroradiology*, Vol. 27, pp.1189–1193.

Imig, T.J. and Morel, A. (1985) 'Tonotopic organisation in lateral part of posterior group of thalamic nuclei in the cat', *Journal of Neurophysiology*, Vol. 53, pp.836–851.

Jonasson, L., Hagmann, P., Pollo, C., Bresson, X., Wilson, C.R., Meuli, R. and Thiran, J-P. (2005) *A Level Set Method for Segmentation of the Thalamus and its Nuclei in DT-MRI*, Elsevier Science, Vol. 53, pp.836–851.

Jones, D.R. and Beltramo, M.A. (1991) 'Solving partitioning problems with genetic algorithms', *Proceedings of the 4th Int. Conf. Genetic Algorithms*, Morgan Kaufman, San Mateo, CA.

Juhasz, C., Chugani, D.C., Muzik, O., Shah, A., Shah, J., Watson, C., Canady, A. and Chugani, H.T. (2001) 'Relationship of flumazenil and glucose pet abnormalities to neocortical epilepsy surgery outcome', *American Academy of Neurology*, Vol. 56, pp.1650–1658.

Karypis, G., Han, E-H. and Kumar, V. (1999) 'CHAMELEON: a hierarchical clustering algorithm using dynamic modeling', *Computer*, Vol. 32, No. 8, pp.68–75.

Krishna, K. and Murty, M.N. (1999) 'Genetic K-Means algorithm', *IEEE Transactions on Systems, Man and Cybernetics (Part B)*, Vol. 29, No. 3, pp.433–439.

Lin, C., Lu, S., Liang, X. and Hua, J. (2006) 'GCA: a coclustering algorithm for thalamo-cortico-thalamic connectivity analysis', *Proceedings of the IEEE International Workshop on Data Mining in Bioinformatics*, pp.163–168.

Maulik, U. and Bandyopadhyay, S. (1994) 'Genetic algorithm based clustering technique', *IEEE Trans on Evolutionary Computation*, Vol. 3, pp.96–101.

Moberts, B., Vilanova, A. and van Wijk, J.J. (2005) 'Evaluation of fiber clustering methods for diffusion tensor imaging', *Visualisation, 2005. VIS 05. IEEE*, pp.65–72.

Rudolph, G. (1994) 'Convergence analysis of canonical genetic algorithms', *Neural Networks, IEEE Transactions*, Vol. 5, pp.96–101.

Scannell, J., Burns, G., Hilgetag, C., O'Neil, M. and Young, M. (1999) 'The connectional organisation of the cortico-thalamic system of the cat', *Cerebral Cortex*, Vol. 9, pp.277–299.

Shimony, J.S., Snyder, A.Z., Lori, N. and Conturo, T.E. (2003) 'Automated fuzzy clustering of neuronal pathways in diffusion tensor tracking', *Soc. Mag. Reson. Med. 10, Conf. Proc.*, Vol. 49, pp.453–456.

Smith, R. (1993) 'Adaptively resizing populations: an algorithm and analysis', in Forrest, S. (Ed.): *Proceedings of the 5th Int. Conf. Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, p.653.

Stejskal, E. and Tanner, J. (1965) 'Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient', *Chem. Phys.*, Vol. 42, pp.288–292.

Taber, K.H., Pierpaoli, C., Rose, S.E., Rugg-Gunn, F.J., Chalk, J.B., Jones, D.K. and Hurley, R.A. (2002) 'The future for diffusion tensor imaging in neuropsychiatry', *J. Neuropsychiatry Clin. Neurosci.*, Vol. 14, pp.1–5.

Tummala, R.P., Chu, R.M., Liu, H., Truwit, C.L. and Halla, W.A. (2003) 'Application of diffusion tensor imaging to magnetic-resonance-guided brain tumor resection', *Pediatric Neurosurgery*, Vol. 39, pp.39–43.

Wiegell, M., Tuch, D., Larsson, H. and Wedeen, V. (2003) 'Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging', *NeuroImage*, Vol. 19, pp.391–401.

Zhang, S. and Laidlaw, D.H. (2002) *Hierarchical Clustering of Streamtubes*, Technical Report, CS-02-18.

Zou, G., Hua, J., Gu, X. and Muzik, O. (2006) 'An approach for intersubject analysis of 3D brain images based on conformal geometry', *Proceedings of International Conference on Image Processing*, pp.1193–1196.