Coclustering for Cross-Subject Fiber Tract Analysis Through Diffusion Tensor Imaging

Cui Lin, *Member, IEEE*, Darshan Pai, Shiyong Lu, *Senior Member, IEEE*, Otto Muzik, and Jing Hua, *Member, IEEE*

***Abstract*—One of the fundamental goals of computational neuroscience is the study of anatomical features that reflect the functional organization of the brain. The study of physical associations between neuronal structures and the examination of brain activity *in vivo* have given rise to the concept of anatomical and functional connectivity, which has been invaluable for our understanding of brain mechanisms and their plasticity during development. However, at present, there is no robust and accurate computational framework for the quantitative assessment of cortical connectivity patterns. In this paper, we present a quantitative analysis and modeling tool that is able to characterize anatomical connectivity patterns based on a newly developed coclustering algorithm, termed the business model-based coclustering algorithm (BCA). We apply BCA to diffusion tensor imaging (DTI) data in order to provide an automated and reproducible assessment of the connectivity patterns between different cortical areas in human brains. BCA not only partitions the cortical mantel into well-defined clusters, but also maximizes the connectivity strength between these clusters. Moreover, BCA is computationally robust and allows both outlier detection as well as parameter-independent determination of the number of clusters. Our coclustering results have showed good performance of BCA in identifying major white matter fiber bundles in human brains and facilitate the detection of abnormal connectivity patterns in patients suffering from various neurological diseases.**

***Index Terms*—Coclustering, corticocortical connectivity, DTI, fiber tracking.**

I. Introduction

With ever-improving imaging technologies, the complexity and scale of brain imaging data have continued to grow at an explosive pace. Recent advances in imaging technologies, particularly in diffusion tensor imaging (DTI), have allowed us to have an increased understanding of both normal and abnormal brain structures and functions [1]. It is well understood that normal brain functions are dependent on the interactions between those that are linked through a complex neural network and specialized functional areas of the brain that process information within local and global networks.

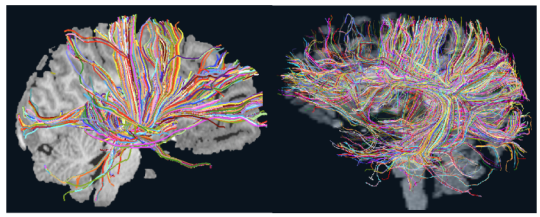
Manuscript received July 22, 2008; revised November 20, 2009 and December 18, 2009. First published January 29, 2010; current version published March 17, 2010. This work was supported in part by the Michigan Technology Tri-Corridor Basic Research Grant MTTC05-135/GR686 and the National Science Foundation Grant IIS-0713315, Grant IIS-0915933, and Grant IIS-0937586.

C. Lin, D. Pai, S. Lu, and J. Hua are with the Department of Computer Science, Wayne State University, Detroit, MI 48202 USA (e-mail: cuilin@wayne.edu; darshan@wayne.edu; shiyong@wayne.edu; jinghua@wayne.edu).

O. Muzik is with the Positron Emission Tomography (PET) Center, Children's Hospital of Michigan, Detroit, MI 48201 USA (e-mail: otto@pet.wayne.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITB.2010.2040286



(a) (b)

Fig. 1. Cerebral cortex sends and receives connections from various subcortical structures. (a) Reciprocal connections between cerebral cortex and thalamus, called *thalamocortical connectivity*; (b) reciprocal connections between different cortical areas, called *corticocortical connectivity*.

Although the cerebral cortex sends and receives connections from various subcortical structures, such as thalamus [see Fig. 1(a)], the largest part of the connections arriving at the cerebral cortex comes from the cerebral cortex itself [see Fig. 1(b)]. One of the most promising approaches to identify distinct functional areas within the cerebral cortex originates from the notion that functionally discrete areas of the cortical mantel can be delineated based on their corticocortical connectivity patterns. Therefore, assessing such connectivity patterns of corticocortical fiber tracts is crucial for our understanding of the mechanisms involved in brain function and may provide clues toward the identification and characterization of many neurological diseases.

DTI is a noninvasive magnetic-resonance-imaging-based technique which allows the *in vivo* measurement of local water diffusion in brain tissues. Most of the intra- and extracellular fluid present in the human body diffuses in the direction along fibrous tissue structures, such as brain white matter. For each voxel, a diffusion tensor can be obtained, characterizing the preferred water diffusion direction. Therefore, DTI gives us the local diffusion profile of the brain and allows the assessment of fiber connectivity patterns within the brain.

Although DTI has been extensively applied to display both major and minor fiber tracts in human brains, a systematic framework allowing reproducible delineation of even major fiber tracts is still elusive. The objective of our research is to devise such a scheme that can facilitate the segmentation of distinct functional territories within the cortex based on quantitative assessment of fiber tract connectivity. Such a basic scheme is illustrated in Fig. 2: most cortical voxels in one region of the cortex are strongly connected to a specific remote cortical region with the connection to any other anatomical territory being much weaker. For example, most of the voxels in cortical region

1089-7771/$26.00 © 2010 IEEE

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

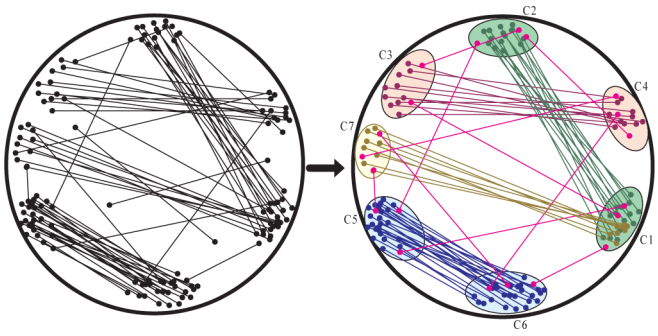


Fig. 2. Our proposed coclustering model.

C_2 are connected to voxels in cortical region C_1, with only few connections to other cortical regions.

Traditional clustering algorithms [2]–[8] are not applicable since they do not account for connectivity and anatomical constraints; as a result, they will fail to identify accurately the corresponding corticocortical connectivity patterns. Consequently, in order to perform an automated connectivity analysis, there is a pressing need to further develop an advanced partitioning scheme that simultaneously partitions voxels into groups, and at the same time, identifies the strongest pairwise connectivity among various groups.

The main contributions of this paper are:

1) we are the first to propose a new coclustering model, which formulates corticocortical connectivity analysis as a computational problem;
2) based on the coclustering model, we propose a business-model-based coclustering algorithm (BCA) to identify functional cortical areas based on the locations of terminal points of corticocortical fiber bundles and anatomical constraints. In contrast to traditional clustering paradigms, BCA is not only able to partition image voxels within the cortical mantel into well-defined clusters, but also is able to maximize the connectivity strength between such clusters. Moreover, BCA is able to automatically identify outliers as well as the number of cortical clusters with high efficiency;
3) The application of BCA to human DTI datasets enables automated, reproducible, and cross-subject assessment of the connectivity patterns of major fiber tracts in human brains.

This paper extends our previous method [9] with a more comprehensive study and improvement: first, while our previous work only proposed the coclustering algorithm for individual brain analysis, this paper extends it to cross-subject analysis of brain fiber tracts, resulting in a totally new section, Section V; second, the original coclustering algorithm requires the user to manually specify input parameters for the analysis of each individual brain; in this paper, we have proposed a strategy to automatically decide these parameters, thus automating the parameter-setting step. The transfer operator is also improved to enhance its performance and effectiveness by introducing additional transfer conditions; third, details about image acquisition, data preprocessing procedures, cross-subject experiments, and statistical evaluation for control groups, patient groups, and their intergroups are reported in order to further evaluate our coclustering algorithm and its application to the clinical research.

The rest of the paper is organized as follows: Section III formalizes the coclustering model for the corticocortical fiber tract analysis. Section IV proposes our coclustering algorithm, BCA, in order to assess fiber tract connectivity between remote cortical areas. Section V presents image data acquisition, DTI fiber tracking for visualizing coclustering results, experiments and evaluation of results on control groups, as well as the application of our framework to clinical research. Finally, Section VI concludes the paper and comments on future work.

II. Background and Related Work

Although the cerebral cortex sends and receives connections from various subcortical structures, the largest part of the connections arriving at the cerebral cortex comes from the cerebral cortex itself. Therefore, assessing connectivity patterns of corticocortical fiber tracts is crucial for our understanding of the mechanisms involved in brain function and may provide clues toward the identification and characterization of many neurological diseases.

Recently, tractography based on DTI has been shown to produce results that are consistent with known pathways formed by major white matter fiber tracts in human brains [10], [11]. As the characteristics of *in vivo* water diffusion depend upon the microscopic tissue architecture [12], changes in the diffusion characteristics serve as markers of change in the tissue microarchitecture. Assuming that water molecules diffuse in a homogeneous medium, diffusion measurements in each image voxel are fitted with a 3×3 second-order tensor model to characterize the self-diffusion of water molecules. This diffusion tensor represents an ellipsoid in space which is mathematically described by the direction and length of its main axes. Deterministic fiber tracking is then incorporated by starting from a seed voxel and then following the direction of the longest (or primary) eigenvector until a termination criterion (such as the decrease of fractional anisotropy below a threshold value) is satisfied [13]. This allows the creation of distinct connections between a set of starting and ending points, which are shown to correspond to known fiber tracts in human brains. Moreover, this connectivity can be quantified by relating the number of fibers originating from the set of seed points to the number of fibers which terminate in the set of ending points. We hypothesize that the connectivity strength can provide clues about the functional performance of a network and can also be used to delineate distinct functional areas within the cortex.

Most clustering algorithms can serve the obvious goal of white matter segmentation. These techniques employed traditional clustering methods such as agglomerative or divisive clustering, in which pairwise similarity measures between individual trajectories are identified for the clustering criteria. These measures can be in the native space of the DTI or some low-dimensional feature space thus reducing computational time. One of the earliest works proposed by Ding *et al.* [14] classified the fiber bundles into "K-most-similar" fiber bundles using

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

an algorithm similar to K-means [2]. The mean of Euclidean distance measure tracks the similarity between fiber tracts. The seed region of interest (ROI) is chosen on a 2-D slice, thus limiting the algorithm to work with certain structures of interest. Zhang and Laidlaw [15] proposed a nine-valued vector similarity identifying the centroid of the starting points, middle points, and ending points of a tract for hierarchical clustering. Maddah *et al.* [16] introduced an atlas-based clustering technique for effective delineation of fiber tracts. A reference brain is identified and landmarked into various hand-drawn regions depending on functional areas identified by underlying tracts. Spatial information is captured for each fiber path trajectory by fitting a B-Spline representation which can be compared to its corresponding location on the atlas using the spatial and geometric properties and labeled accordingly. Spectral clustering techniques [17], [18] find a local affinity matrix created from a set of eigenvectors that can be used to cluster the fiber traces. O'Donnell and Westin [17] employed the Nystrom method to interpolate the values of an eigenvector system identifying the embedding vectors from a tract-similarity matrix. The embedding vectors are then clustered using any clustering method. Jonasson *et al.* [18] used high angular resolution diffusion imaging (HARDI) data to create a co-occurrence matrix that lists the number of times two fibers share the same voxel. Because of the high orientational resolution of HARDI, the matrix creation is trivial though data acquisition is impractical when compared to DTI. A similar technique using a weighted undirected graph instead of a data matrix is presented by Brun *et al.* [19]. The technique uses normalized cuts [20] to partition a graph into clusters. Clustering based on local shape models and probabilistic models were explored by some authors that described a spatial model based on intrinsic shape properties [21] and a posterior distribution model [22] based on length, similarity angle, and continuity angle between candidate tracts and reference tracts.

Traditional partitioning-based clustering algorithms, such as *K*-means and *K*-medoids [3] are simple and efficient. However, their final results may be overly sensitive to initialization and the presence of outliers. In addition, these methods require the knowledge of K, which might not be possible for many applications. Hierarchical clustering algorithms [4], [5] do not require the number of clusters K as input, but they require a termination condition and do not support re-classification of objects to new clusters. Density-based algorithms [6]–[8] have good performance with respect to noise handling and do not require the specification of K. However, they do not consider the maximal connectivity strength condition between clusters. Spectral clustering algorithms [17], [18] can efficiently partition objects into clusters, but have similar limitations of not considering the cross-connectivity between clusters, and the need of manual specification of the number of clusters and other parameter values for each individual dataset.

Despite some success in clustering and delineating functional cortical areas using DTI, a systematic framework which allows functional parcellation of the neocortex into distinct functional units based on quantitative assessment of fiber tract connectivity has not yet been produced, and the relationship among functional territories, fiber tracts, and neuronal connections remains controversial. Most aforementioned work can only handle few structures of interest and are computationally expensive. Also, the dispersion effect of the fibers near the cortical regions can seriously affect similarity measures based on distance, hence reducing the clustering accuracy. Therefore, there is a pressing need to further develop advanced clustering algorithms that allow better characterization of brain connectivity patterns, and as a result, improve our understanding of process interactions in a complex biological system. Since our proposed BCA makes clustering decisions using both the cortical ending points and the connectivity traces, it can handle the dispersion effect as well as proximity of the fiber traces deep within the white matter structures.

It is important to note the major differences between this work on corticocortical connectivity analysis and previous work on thalamocortical connectivity analysis [23]. First, the thalamocortical problem requires two predefined classes for coclustering, the class of cortical voxels and the class of thalamic voxels, while the corticocortical problem does not have such an anatomical constraint to exploit, resulting in a much larger exploration space and a dramatic increase in computational complexity. Second, the cross-connectivity cost in thalamocortical connectivity is only determined by the connectivity across objects in different classes, while the cost in corticocortical connectivity has to consider connections between all objects. Third, the thalamocortical problem requires the number of clusters for each class of objects; while there are no such clues to predefine the number of clusters for objects in the corticocortical problem. As a result, the previous GCA algorithm [23] is inapplicable to the corticocortical problem. Since GCA and BCA address two fundamentally different problems and one algorithm will not work for the other problem, these two algorithms are incomparable by experiments.

III. Coclustering Model

In this section, we present our coclustering model, which models the corticocortical connectivity analysis problem. In this model, the structure of the cortex and corticocortical connections are represented as a graph $G = (V, F)$ (see Fig. 2), where V is the set of cortical voxels and F represents all the pairwise corticocortical connections. For each voxel $v \in V$, we use $F(v)$ to denote the other voxel that is connected to v. Although not required by our model, the working hypothesis is that "each cortical region is *strongly* connected to another specific cortical region and is *weakly* connected to other cortical regions." The goal of a coclustering procedure is to group cortical voxels into K clusters (where K will be determined automatically) while minimizing the cross-connectivity cost between those clusters, so that: 1) close voxels are within the same cluster, while distant voxels are in different clusters; 2) each cluster is strongly connected to another cluster, called its *spouse cluster*; and (3) the total cross-connectivity cost between each cluster and its any other nonspouse cluster is minimized.

To achieve the aforementioned goals, we introduce the notions of *spouse cluster* and *partner cluster* and develop a model that aims to minimize the cross-connectivity cost between

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

clusters by requiring that for each cluster, its spouse cluster should mostly *coincide* with its partner cluster. We firstl define the notion of *association strength* of cluster C_j with respect to cluster C_i.

Definition 3.1 [Association Strength $\theta(C_i, C_j)$]: Given a partition $C = \{C_1, \ldots, C_K\}$ of $G(V, F)$, the association strength of C_j w.r.t. C_i is defined as

$$\theta(C_i, C_j) = \frac{N_{ij}}{|C_i|} \tag{1}$$

where N_{ij} denotes the total number of connections between C_i and C_j and $|C_i|$ is the total number of voxels in C_i.

The spouse cluster of C_i is the cluster that has the maximum association strength w.r.t. C_i, which is defined formally as follows.

Definition 3.2 (Spouse Cluster): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the spouse cluster of a cluster C_i is defined as

$$\mathrm{SC}(C_i) = \{C_k | C_k \in C \text{ and } \forall_{C_j \in C}, \theta(C_i, C_k) \geq \theta(C_i, C_j)\}. \tag{2}$$

In the case that multiple clusters have the same maximum association strengths w.r.t. C_i, our model will *nondeterministically* select one of them as the spouse cluster of C_i. A cluster and its spouse cluster forms a *cocluster*.

Definition 3.3 (Cocluster): Given a partition C of $G(V, F)$, $C = \{C_1, C_2, \ldots, C_K\}$, $\langle C_i, C_j \rangle$ is called a cocluster if C_j is the spouse cluster of C_i.

Example 3.4: In Fig. 2, C_2 is C_1's spouse cluster and C_1 is C_7's spouse cluster. As a result, $\langle C_1, C_2 \rangle$ and $\langle C_7, C_1 \rangle$ form two coclusters. Note that the notion of cocluster is nonsymmetric, i.e., $\langle C_i, C_j \rangle$ is a cocluster does not imply $\langle C_j, C_i \rangle$ is a cocluster. In our example, $\langle C_1, C_7 \rangle$ is not a cocluster although $\langle C_7, C_1 \rangle$ is a cocluster.

We define the partner cluster of a cluster C_i as the cluster consisting of all the voxels that are connected to voxels in C_i.

Definition 3.5 (PC): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the partner cluster for a cluster $C_i \in C$ is defined as

$$\mathrm{PC}(C_i) = \{v' | v' \in V \text{ and } \exists v \in C_i, (v', v) \in F\}. \tag{3}$$

Obviously, $\mathrm{PC}(C_i)$ has the largest connection strength to C_i ($\theta(C_i, \mathrm{PC}(C_i)) = 1$). In an ideal case, if all partner clusters and spouse clusters coincide exactly as shown in Fig. 3(a), then the overall association strength is maximized as the association strength between each cluster and its spouse cluster (which is also its partner cluster) is maximized. However, this is not the case in general, as shown in Fig. 3(b). Therefore, maximizing the overall association strength can be reduced to maximizing the coincidence between each cluster's spouse cluster and partner cluster. Below, we quantify the deviation of a cluster's partner cluster from its spouse cluster by *partner within-cluster variation* PWCV. In this way, we can further reduce the problem of maximizing the overall association strength to minimizing the *total* PWCV (TPWCV).

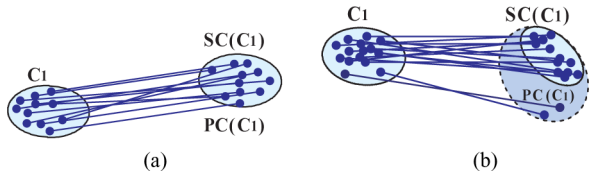


Fig. 3. Partner clusters in (a) an ideal case and (b) general case. (a) $\mathrm{SC}(C_1) = \mathrm{PC}(C_1)$. (b) $\mathrm{SC}(C_1) \neq \mathrm{PC}(C_1)$.

In order to calculate PWCV, we first define the *centroid* of a cluster and its WCV to quantify the similarity of objects within one cluster.

Definition 3.6 (Centroid): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the centroid of cluster C_i is defined as

$$\overrightarrow{\mu_i} = \frac{\sum_{\overrightarrow{X_n} \in C_i} \overrightarrow{X_n}}{|C_i|}. \tag{4}$$

Definition 3.7 (WCV): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the WCV of cluster C_i is defined as

$$\mathrm{WCV}(C_i) = \sum_{\overrightarrow{X_n} \in C_i} d(\overrightarrow{X_n}, \overrightarrow{\mu_i}) \tag{5}$$

where $d(\overrightarrow{X_n}, \overrightarrow{\mu_i})$ is the Euclidean distance between the voxel $\overrightarrow{X_n}$ and the centroid $\overrightarrow{\mu_i}$ of cluster C_i.

Definition 3.8 (PWCV): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the PWCV of a partner cluster $\mathrm{PC}(C_i)$ is defined as

$$\mathrm{PWCV}(\mathrm{PC}(C_i)) = \sum_{\overrightarrow{X_n} \in PC(C_i)} d(\overrightarrow{X_n}, \overrightarrow{\nu_i}) \tag{6}$$

where ν_i is the centroid of $\mathrm{SC}(C_i)$.

Note that, instead of using the centroid of the partner cluster, the centroid of the corresponding spouse cluster is used to calculate $\mathrm{PWCV}(\mathrm{PC}(C_i))$. The intuition is that, in an ideal partitioning, the partition $\{\mathrm{PC}(C_1), \ldots, \mathrm{PC}(C_K)\}$ should mostly coincide with $\{\mathrm{SC}(C_1), \ldots, \mathrm{SC}(C_K)\}$.

Then, we define the TWCV and TPWCV to quantify the overall quality of a partitioning.

Definition 3.9 (TWCV): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the TWCV of C is defined as

$$\begin{aligned} \mathrm{TWCV}(C) &= \sum_{i=1}^{K} \mathrm{WCV}(C_i) \\ &= \sum_{i=1}^{K} \sum_{\overrightarrow{X_n} \in C_i} \sum_{d=1}^{D} (X_{n_d} - \mu_{i_d})^2 \\ &= \sum_{i=1}^{K} \sum_{d=1}^{D} X_{n_d}^2 - \sum_{i=1}^{K} \frac{1}{|C_i|} \sum_{d=1}^{D} (\mathrm{SCF}_{i_d})^2 \end{aligned} \tag{7}$$

where SCF_{i_d} is the sum of the ith feature of all voxels in C_i.

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

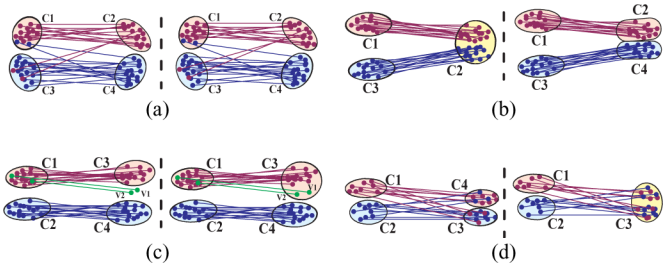


Fig. 4. Compare coclustering results with traditional clustering results in the aforesaid four cases. (a) *Coincide* case. (b) *Split* case. (c) *Transfer* case. (d) *Merge* case.

Definition 3.10 (TPWCV): Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the TPWCV of a partition $\{C_1, \ldots, C_K\}$ is defined as

$$\begin{aligned}&\mathrm{TPWCV}(\mathrm{PC}(C_1), \ldots, \mathrm{PC}(C_K)) \\ &\quad = \sum_{i=1}^{K} \mathrm{PWCV}(\mathrm{PC}(C_i)) \\ &\quad = \sum_{i=1}^{K} \sum_{\overrightarrow{X_n} \in PC(C_i)} \sum_{d=1}^{D} (X_{n_d} - \nu_{i_d})^2. \end{aligned} \tag{8}$$

Finally, the coclustering problem can be formally stated as follows: given a graph $G = (V, F)$, a distance metric d for nodes between $v_i \in V$ and $v_j \in V$ $(i \neq j)$, and α as a weight value, partition G into K coclusters, $\{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_K, \mathrm{SC}(C_K)\rangle\}$ such that the following objective function OWCV is minimized

$$\begin{aligned}\mathrm{OWCV}(C) = {} & \alpha \times \mathrm{TWCV}(C) \\ & + (1-\alpha) \times \mathrm{TPWCV}(\mathrm{PC}(C_1), \ldots, \mathrm{PC}(C_K)).\end{aligned} \tag{9}$$

Note that our objective function does not include a component that directly maximizes the overall association strength. This is so because the association strength and TWCVs have different units and their direct aggregation is problematic. However, by minimizing TPWCV, which is a component of the objective function, the overall deviation of partner clusters from spouse clusters is minimized, and thus, the overall association strength is maximized. In particular, when all partner clusters coincide exactly with spouse clusters, the overall association strength reaches its maximum.

While a traditional clustering paradigm [2], [3] does not consider the connection patterns between clusters, our coclustering model considers both the distances and connections between voxels. As a result, our coclustering model can provide several salient advantages over traditional clustering in partitioning graphs in general and in corticocortical connectivity analysis in particular. Fig. 4 illustrates the comparison between the partitioning results from our coclustering model and from the traditional clustering paradigm. In Fig. 4(a), the coclustering result from our model (right) coincides with the traditional clustering results (left), in which objects are grouped into four clusters from C_1 to C_4, and similar objects are in one cluster, while dissimilar objects are in different clusters. In this coclustering partition, each cluster is strongly connected to its spouse cluster and each partner cluster coincides with the corresponding spouse cluster. In Fig. 4(b), while a traditional clustering procedure that only considers TWCV groups C_2 and C_4 into one cluster (left), our clustering model can potentially split that cluster into C_2 and C_4 by considering that most voxels in C_2 are connected to C_1 and most voxels in C_4 are connected to C_3 (right). In Fig. 4(c), while a traditional clustering procedure might consider v_1 and v_2 as outliers, or classify them to C_4 simply because they are closer to C_4 (left), our coclustering model can potentially reassign them into C_3 by observing that these two voxels, like other voxels in C_3, are mainly connected to C_1 (right). Finally, in Fig. 4(d), while a traditional clustering procedure partitions two clusters C_3 and C_4 as they fall apart to some degree in distances (left), our coclustering model can potentially merge them into one cluster C_3 by observing that all voxels in C_1 and C_2 are connected to C_3 (right).

The proposed coclustering model is different from the model [23] for thalamocortical connectivity in the following aspects: 1) a cluster and its *spouse cluster* are all from cortical voxels, while previously one has to come from cortical voxels and the other from thalamic voxels; 2) we introduce the concept of *association strength* to quantify each cluster's spouse cluster; and 3) the new concepts PC, PWCV, and TPWCV are introduced to determine the cross-connectivity costs between clusters and there is no constraint on the number of clusters, as well as the class that a cluster belongs to.

IV. The BCA Coclustering Algorithm

In this section, we propose the BCA to solve the coclustering problem. The BCA starts with a density-based initialization, which generates an initial number of cortical clusters. Then, it produces a better solution from the current solution by applying the following three operators, viz. *split*, *transfer*, and *merge* sequentially. This procedure is iterated until a certain terminating condition is reached, e.g., after certain number of iterations or when the improvement between consective iterations is within some user-specified threshold. In the following, we describe the initialization, split, transfer, and merge operators in detail.

A. Density-Based Initialization

Our density-based initialization is based on the idea of DBSCAN [24], in which clusters are identified by a density-based spational clustering algorithm that uses two global density parameters: ε, the maximum radius of a voxel, and δ, the minimum number of voxels in an ε-neighbor of a voxel v, denoted $N_\varepsilon(v)$, for the voxel to be a *core voxel*. A voxel v' is *directly density-reachable* from a voxel v w.r.t. ε, δ if $v' \in N_\varepsilon(v)$. A voxel v' is *density-reachable* from a voxel v w.r.t. ε, δ if there is a chain of voxels $v_1 = v, \ldots, v_n = v'$ such that v_{i+1} is directly density-reachable from v_i. A voxel v' is *density-connected* to v w.r.t. ε, δ if there exists a voxel v'' such that both v and v' are density-reachable from v'' w.r.t. ε and δ. A cluster C_k is formed

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

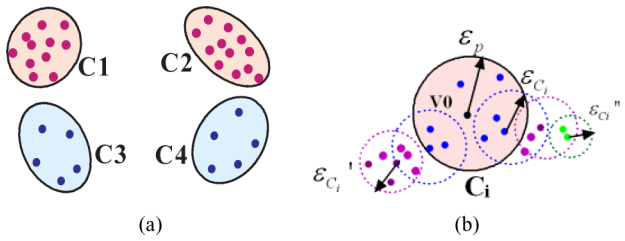


Fig. 5. (a) Coclustering with various densities; (b) adaptive determination of ε_{C_j} and threshold δ_{C_j}.

in DBSCAN by starting a core voxel v and adding all voxels that are density-reachable from v or density-connected to v to C_k.

While a density-based clustering initialization has the advantage of being able to identify clusters of arbitrary shapes (as cortical regions can be of different shapes), there are two major issues we need to resolve: 1) different brains differ in their sizes and the numbers of voxels; hence, applying the same values of ε and δ to each brain will fail to identify some key corresponding regions across brains; and 2) different cortical regions in a brain have different densities; hence, applying the same values of ε and δ to a whole brain fails to identify some sparse clusters.

To resolve the first issue, we have hypothesized that, *although each brain might differ in the total number of voxels, the proportion of voxels that fall in each corresponding anatomical region is the same*. This hypothesis is formulated as

$$\frac{\delta_1}{N_1} = \frac{\delta_2}{N_2} = \cdots = \frac{\delta_N}{N_N} = C_\delta \tag{10}$$

where N_p $(p = 1, \ldots, N)$ is the total number of voxels for brain B_p, and C_δ is a brain-independent constant with an empirical value of 0.00435 for the best results in our study. In this way, δ_p can be calculated as $C_\delta \times N_p$. To calculate ε_p, we apply Chebyshev's inequality [25]. Let v_i be a random cortical voxel in brain B_p, let μ_p and θ_p^2 be the mean and variance of all voxels' shortest distances in B_p, respectively. We have

$$P\{|v_i - \mu_p| < \varepsilon_p\} \geq 1 - \frac{\theta_p^2}{\varepsilon_p^2}. \tag{11}$$

In our study, we use $\varepsilon_p = 5 \times \theta_p$ and get $P\{|v_i - \mu_p| < 5\theta_p\} \geq 0.96$, which implies that v_i has a probability of at least 96% within the radius of five standard deviations around the mean. One important property of Chebyshev's inequality is that the aforesaid equation holds regardless of the exact distribution of a random variable, and only the knowledge of mean μ_p and variance θ_p^2 is required.

To resolve the second issue, which arises due to the fact that different cortical regions have different densities of voxels [see Fig. 5(a)], instead of using the radius ε_p and threshold δ_p to expand all distance-reachable voxels from $N_{\varepsilon_p}(v)$ for all clusters that have diverse densities, we apply ε_{C_i} and threshold δ_{C_i} to locally determine the extension of each cluster. The assumption is that each cortical region has a similar local density while different cortical regions can have dramatically different densities. In order to dynamically and adaptively determine the *local* density parameter values for ε_{C_i} and δ_{C_i} for each working cluster C_i, first, a core voxel v and its $N_{\varepsilon_p}(v)$ are identified as an initial cluster C_i. Then, the value of $\varepsilon_{C_i} = 5\theta_{C_i}$ and $\delta_{C_i} = C_\delta |C_i|$ are calculated in a way similar to ε_p and δ_p, except that we use the current voxels in C_i instead of the voxels of the whole brain. For each voxel v', its $N_{\varepsilon_{C_i}}(v')$ is added to C_i. The procedure is repeated until we reach $N_{\varepsilon_{C_i}}(v) \leq \delta_{C_i}$. In Fig. 5(b), a new cluster C_i is formed by ε_p and δ_p. Then, ε_{C_i} is used to expand C_i within the radius of ε_{C_i} (blue circles), if the current number of voxels $N_{\varepsilon_{C_i}}(v)$ is no less than δ_{C_i}, a new version of ε_{C_i} (purple circles) is recalculated. This procedure continues until $N_{\varepsilon_{C_i}}(v) \leq \delta_{C_i}$ (green circle).

After the above procedure, some voxels might be still unclassified due to the following reasons: 1) some voxels cannot form a new cluster by ε_p and δ_p; 2) some voxels cannot be assigned to any cluster by an expansion based on ε_{C_i} and δ_{C_i}; or 3) some voxels can be noises and are distant to all other voxels. We proceed with these voxels as follows. 1) For each voxel v, let v' be the closest voxel to v that has been classified to some cluster C_i. If $dist(v, v') \leq \varepsilon_p$, then v is assigned to C_i as well. This step is repeated until no more voxels can be classified using this step. 2) For those voxels which are still unclassified after step (2), they are considered as outliers. All cocluster outliers are then eliminated from graph $G(V, F)$, forming $G(V', F')$.

B. The Split Operator

The goal of the split operator is to eliminate the situation that a cluster hierarchically contains subclusters that have different spouse clusters, as illustrated in Fig. 4(b), by an attempt to split a cluster into two clusters when such a split will improve the result of coclustering that is characterized by the following split condition:

Definition 4.1 (Split Condition): Given a coclustering $\mathrm{CC} = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_i, \mathrm{SC}(C_i)\rangle, \ldots, \langle C_K, \mathrm{SC}(\mathrm{C_K})\rangle\}$, let C_{i_1} be the set of voxels in $C_i \in \mathrm{CC}$ that are connected to $\{\mathrm{SC}(C_i)$, C_{i_2} be $C_i - C_{i_1}$, and $\mathrm{CC}' = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_{i_1}, \mathrm{SC}(C_{i_1})\rangle, \langle C_{i_2}, \mathrm{SC}(C_{i_2})\rangle, \ldots, \langle C_K, \mathrm{SC}(C_K)\rangle\}$, then we say that C_i satisfies the split condition iff

1) $|C_{i_1}| \geq \delta_{C_{i_1}}$ and $|C_{i_2}| \geq \delta_{C_{i_2}}$;
2) $\mathrm{OWCV}(\mathrm{CC}') \leq \mathrm{OWCV}(\mathrm{CC})$;
3) $\theta(C_i, \mathrm{SC}(C_i)) \leq \theta(C_{i_2}, \mathrm{SC}(C_{i_2}))$.

Intuitively, the split condition ensures that after a split, 1) the number of voxels in each new cluster is still greater than or equal to δ_{i_1} and δ_{i_2}; 2) the OWCV value for the new coclustering will not increase; and 3) the association strengths of the two new clusters C_{i_1} and C_{i_2} will be no less than the association strength of the original cluster C_i. This is always true for C_{i_1}, and thus, we only need to require $\theta(C_i, \mathrm{SC}(C_i)) \leq (C_{i_2}, \mathrm{SC}(C_{i_2}))$ in the aforesaid definition of the split condition. Algorithm Split is sketched in Fig. 6. Basically, it iteratively splits the colustering result until no more cluster satisfies the aforegiven split condition.

C. The Transfer Operator

The transfer operator attempts to reassign certain voxels to a new cluster, as illustrated in Fig. 4(c), in order to improve

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

(1) **Algorithm:** Split
(2) **Input:** $CC = \{\langle C_1, SC(C_1)\rangle, \cdots, \langle C_K, SC(C_K)\rangle\}$
(3) **Output:** a new version of CC in which no more cluster satisfies the split condition
(4) **Begin**
(5) **While** there exists a cluster $C_i \in CC, \theta(C_i, SC(C_i)) \neq 1$ satisfying the split condition **do**
(6) Split C_i into C_{i_1} and C_{i_2};
(7) Recalculate the spouse cluster for each cluster in CC' according to Definitions in Section III;
(8) **End while**
(9) **End Algorithm**

Fig. 6. Algorithm Split.

(1) **Algorithm:** Transfer
(2) **Input:** $CC = \{\langle C_1, SC(C_1)\rangle, \cdots, \langle C_K, SC(C_K)\rangle\}$
(3) **Output:** a new version of CC in which no more cluster satisfies the transfer condition
(4) **Begin**
(5) **While** there exists a voxel $v \in C_i$ satisfying the transfer condition **do**
(6) Transfer v from C_i to C_j such that $F(v) \in SC(C_j)$;
(7) Recalculate the spouse cluster for each cluster in CC' according to Definitions in Section III;
(8) **End while**
(9) **End Algorithm**

Fig. 7. Algorithm Transfer.

the result of coclustering that is characterized by the following transfer condition:

Definition 4.2 (Transfer Condition): Given a coclustering $\mathrm{CC} = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_i, \mathrm{SC}(C_i)\rangle, \ldots, \langle C_j, \mathrm{SC}(C_j)\rangle, \ldots, \langle C_K, \mathrm{SC}(C_K)\rangle\}$, let $v \in C_i, F(v) \notin \mathrm{SC}(C_i)$, and C_j be another cluster such that $F(v) \in \mathrm{SC}(C_j)$. After transferring v from C_i to C_j, C_i becomes C_i', C_j becomes C_j', and CC becomes $\mathrm{CC}' = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_i', \mathrm{SC}(C_i')\rangle, \ldots, \langle C_j', \mathrm{SC}(C_j')\rangle, \ldots, \langle C_K, \mathrm{SC}(C_K)\rangle\}$, we say that v satisfies the transfer condition iff

1) $|C_i'| \geq \delta_{C_i'}$;
2) $v \in C_i, F(v) \notin \mathrm{SC}\}(C_i)$ and $F(v) \in \mathrm{SC}(C_j)$;
3) $\mathrm{OWCV}(\mathrm{CC}') \leq \mathrm{OWCV}(\mathrm{CC})$;
4) $\theta(C_i, \mathrm{SC}(C_i)) \leq \theta(C_i', \mathrm{SC}(C_i'))$ and $\theta(C_j, \mathrm{SC}(C_j)) \leq \theta(C_j', \mathrm{SC}(C_j'))$.

Intuitively, the transfer condition ensures that after a transfer, 1) C_i still contains at least δ_p voxels; 2) we transfer v from C_i to C_j only if the connectivity involving v is a cross-connectivity, i.e., v does not connect to the spouse cluser of C_i but connects to the spouse cluster of C_j; 3) the OWCV value for the new coclustering will not increase; and 4) the association strengths of the two affected clusters will not decrease. Algorithm Transfer is sketched in Fig. 7. Basically, it attempts to assign voxels to a new cluster if it satisfies the transfer condition. The procedure terminates when no more voxel satisfies the aforementioned transfer condition.

D. The Merge Operator

Finally, the merge operator attempts to merge two clusters, illustrated in Fig. 4(d), if such a merge will improve the result of coclustering that is characterized by the following merge condition:

(1) **Algorithm:** Merge
(2) **Input:** $CC = \{\langle C_1, SC(C_1)\rangle, \cdots, \langle C_K, SC(C_K)\rangle\}$
(3) **Output:** a new version of CC in which no more cluster satisfies the transfer condition
(4) **Begin**
(5) **While** there exists $C_i, C_j \in CC$ satisfying the merge condition **do**
(6) Merge C_i and C_j into C_m;
(7) Recalculate the spouse cluster for each cluster in CC' according to Definitions in Section III;
(11) **End While**
(13) **End Algorithm**

Fig. 8. Algorithm Merge.

Definition 4.3 (Merge condition): Given a coclustering $\mathrm{CC} = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_i, \mathrm{SC}(C_i)\rangle, \ldots, \langle C_j, \mathrm{SC}(C_j)\rangle, \ldots, \langle C_K, \mathrm{SC}(C_K)\rangle\}$, and two clusters $C_i, C_j \in \mathrm{CO}$, we merge C_i and C_j into C_m and derive a new coclustering $\mathrm{CC}' = \{\langle C_1, \mathrm{SC}(C_1)\rangle, \ldots, \langle C_m, \mathrm{SC}(C_m)\rangle, \ldots, \langle C_{K-1}, \mathrm{SC}(C_{K-1})\rangle\}$. Then, we say that C_i and C_j satisfy the merge condition iff

1) $\mathrm{OWCV}(\mathrm{CC}') \leq \mathrm{OWCV}(\mathrm{CC})$;
4) $\theta(\mathrm{SC}(C_i), C_m) \leq \theta(\mathrm{SC}(C_m), C_i)$ and $\theta(\mathrm{SC}(C_j), C_m) \leq \theta(\mathrm{SC}(C_m), C_j)$.

Intuitively, the merge condition ensures that after a merge, 1) the OWCV value for the new coclustering will not increase and 2) the association strength of $\mathrm{SC}(C_i)$ (SC of C_i) and $\mathrm{SC}(C_j)$ (SC of C_j) are not reduced, with their new spouse cluster C_m. The Algorithm Merge is sketched in Fig. 7. Basically, it merges two clusters into one if the two clusters satisfy the aforementioned merge condition. The algorithm terminates when no more pair of clusters satisfy the aforementioned merge condition.

V. Image Processing and Experiments

To facilitate the cross-subject fiber tract analysis, we develop a comprehensive framework that includes three major software modules: 1) a fiber tracking module that is responsible for DTI preprocessing; 2) a coclustering processing module that implements the proposed BCA and automatically produces coclustering results and analytical results for each brain; 3) and a visualization and analysis module to visualize coclusters and conduct statistical analysis for control groups, patient groups, and their intergroups based on the postprocessed coclustering results. Modules 1) and 3) are developed in Visual C++ on the Windows platform, and module 2) is developed using C#. The visualization routines are referenced from the VTK library. A Pentium Core Duo processor with 2 GB RAM and an Nvidia 8600GT graphics card is used for all computations. The total time taken for processing a single dataset is around 20 minutes including human intervention.

This section details data acquisition and DTI fiber tracking for visualizing coclustering results, experiments and evaluation of the BCA coclustering results on normal subjects (controls), as well as the application of our framework to the diagnosis of Tourette syndrome (TS) patients.

A. Image Data Acquisition

MRI studies were performed on a GE 1.5T Signa unit (GE Medical Systems, Milwaukee, Wisconsin). Volumetric imaging

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

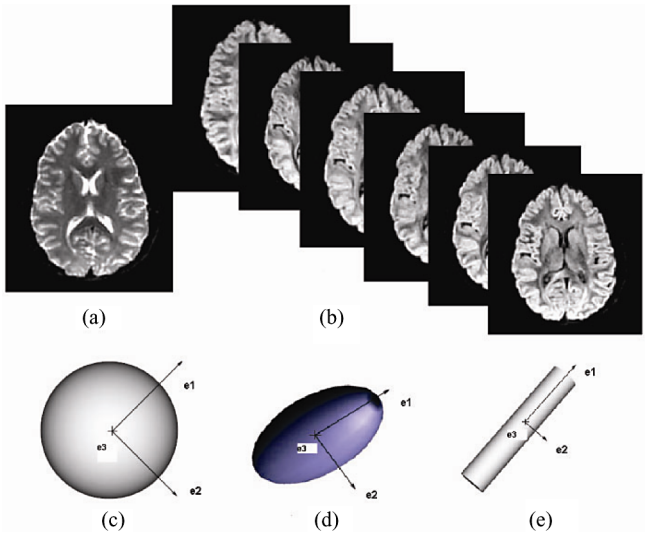


Fig. 9. (a) and (b) Nondiffusion-weighted image (T-1 image) and the six gradient-weighted images, respectively. (c)–(e) Various diffusion profiles ranging from an isotropic profile (spherical) to a highly anisotropic profile (cigar-shaped), with their corresponding eigendirections. $e1$ is the primary eigendirection.

was performed utilizing a T1-weighted spoiled gradient echo (SPGR) sequence. The 3-D SPGR technique generates 124 contiguous 1.5 mm sections of the entire head using a 35/5/1 (TR/TE/NEX) pulse sequence, flip angle of 35°, matrix size of 256 × 128, and FOV of 220 mm. Diffusion-weighted dual spin-echo single-shot echoplanar MR imaging was performed using the following parameters: 6000/110 (TR/TE), image matrix 256 × 256, FOV of 240 mm, 40 planes with a slice thickness of 3 mm covering the whole brain. The DTI sequence consists initially of an image volume with no diffusion weighting (b = 0 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 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.

B. DTI Computing

DTI [26] is a new technique that measures the diffusion of water molecules in human brain tissue and allows the *in vivo* assessment of local anisotropic water diffusion in human brains. Once the diffusion tensor is determined at each voxel location, the information can be used to create continuous fiber tracts by following the direction of the strongest diffusion, also referred to as the primary eigendirection [see Fig. 9(c)–(e)]. Such fiber tracts are subsequently used to estimate reciprocal connectivity between various subcortical structures. Moreover, quantitative measures extracted from the diffusion tensor provide a measurement of overall connectivity strength, and thus, may aid in the clinical management of patients with various neurological diseases.

In deterministic tractography, a single fiber model is assumed at each image voxel characterized by a diffusion tensor D with six independent elements. To obtain numerical values for the diffusion tensor, at least seven independent measurements are required: six diffusion-weighted images [Fig. 9(a)] along six independent diffusion sensitizing gradient directions x, and one nondiffusion-weighted image [Fig. 9(b)]. An eigenvalue decomposition of D results in an optimal rotation transformation from the reference coordinate system to a new basis system along the principal directions (eigenvectors $e1$, $e2$, and $e3$) of the diffusion profile. The three eigenvalues λ_1, λ_2, and λ_3 characterize the length of the main axis of the constructed ellipsoid and provide an orientation-independent measure of anisotropy which describes the structural geometry of the underlying microstructure. The fractional anisotropy (FA) in each image voxel is subsequently defined as

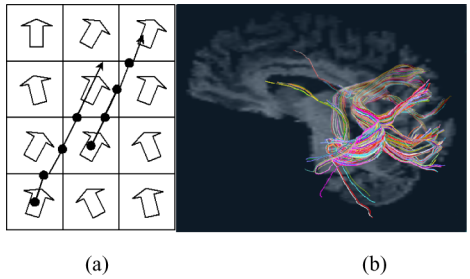


Fig. 10. (a) Example of a deterministic fiber tract; (b) fiber tracts for an ROI in the corpus callosum, with the deterministic tracking.

Definition 5.1 [Fractional Anisotropy (FA)]: Given the eigenvalues λ_1, λ_2, and λ_3, the fractional anisotropy of a fiber tract is defined as

$$\mathrm{FA}(\lambda_1,\lambda_2,\lambda_3)=\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{2}\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}}. \tag{12}$$

A high FA indicates a strong directionality of water diffusion in the underlying tissue, corresponding to an ellipsoidal diffusion profile. A low FA value indicates low-directionality corresponding to a spherical diffusion profile, as shown in Fig. 9(c). Since FA provides a quantitative measure of the anisotropy, it is used as a stopping criteria for fiber tracking in white matter.

Fig. 10(a) shows an illustrative example of the applied fiber tracking algorithm. The arrows represent the primary eigendirection $e1$ which corresponds to the strongest diffusion direction. Starting at the center of each voxel, the fiber is tracked within the voxel along the primary eigendirection until it hits the boundary between two voxels, after which the direction changes to the primary eigendirection of the next voxel. The fiber is tracked in both forward and backward directions and is connected together to form one continuous fiber tract. The procedure is repeated for all voxels within the brain volume. To terminate a fiber tract, one of two conditions needs to be met: either the voxel FA value falls below a predefined threshold or the angle between two consecutive eigendirections exceeds a predefined turning angle. The threshold value for FA is usually chosen as 0.15 and the threshold for the turning angle is typically chosen to be 70°. This fiber tracking procedure allows optimal tracking of prolate diffusion profiles, where $\lambda_1 \gg \lambda_2$ and λ_3. Empirical studies have shown that this tractography technique

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

TABLE I
BCA EXPERIMENTAL RESULTS IN A POPULATION OF PATIENT DATASETS

P_{ID}	N_V	$OWCV_{DCA}$	$OWCV_{BCA}$	ε_{DCA}	ε_{BCA}	δ_{DCA}	δ_{BCA}	$OWCV_{Init}$	$OWCV_{Split}$	$OWCV_{Transfer}$
A	12316	$3702493E+5$	$177984E+5$	15	11.67	55	53	$9462832E+4$	$5599247E+4$	$177984E+5$
B	11544	$1793856E+5$	$1638639E+4$	10	12.61	52	50	$7923895E+4$	$4720502E+4$	$4584369E+4$
C	10922	$4080492E+5$	$1718737E+4$	12	10.95	50	47	$7975345E+4$	$32943E+6$	$3237084E+4$
D	6830	$7357482E+4$	$4962117E+3$	10	12.35	25	29	$4388464E+4$	$213765E+5$	$2117771E+4$
E	8436	$1967691E+5$	$8620464E+3$	10	11.85	30	36	$340694E+5$	$1782064E+4$	$1782064E+4$

allows reliable tracking of major fiber bundles in the human brain [27]. For the purpose of our research, we were interested in major corticocortical fiber bundles connecting different functional cortical areas. As the pathways of these fiber bundles are well known, they are well suited to assess the feasibility of the proposed coclustering algorithm.

C. Experiments and Evaluation of Coclustering Results

Six normal subjects were chosen for our experiment. High-resolution MR Tl-weighted images as well as DTI image volumes were acquired for each subject and subsequently coregistered using an in-house-developed registration program. Moreover, Tl-weighted MR images were deskulled and the exposed cortical surface was then subdivided into 512 finite elements using a recursive parcellation scheme. Based on a user-defined depth parameter these finite elements were then extended into the cortex creating a cortical mantel of 15 mm depth. The depth parameter was chosen as 15 mm depth so that the border zone between the cortical gray matter and the underlying white matter was included. Fiber tracking was initiated from each voxel of the cortical mantel using the deterministic fiber tracking method described in the previous Section V-B. Once fiber tracking was completed, corticocortical fibers (fibers both originating and terminating in the cortical mantel) and their corresponding starting voxels were saved for subsequent coclustering. In order to exclude short U-shaped fibers, those fibers in which the starting and end voxel belonged to the same finite element were excluded. Moreover, to ensure optimal separation between various coclusters, a minimal fiber length of 100 mm was selected.

1) BCA Experimental Results: We have implemented BCA on a population of patient datasets, in which five coclustering results from patient P_A to P_E were chosen as shown in Table I (their real names are hidden for privacy protection). As DBSCAN is not directly applicable to produce coclustering results, we implemented DBSCAN to initialize the clustering results at the first and then generated coclustering results according to definitions in Section III. The distinct improvement of BCA from DBSCAN-based coclustering algorithm (DCA) can be seen in Fig. 11(a): For patient P_A, OWCV produced by DCA is 20.8 times greater than OWCV by BCA; For patient P_B, OWCV generated by BCA has achieved 90.87% decrease than its counterpart; Such improvement can also be seen from patient P_C to P_E, the values of OWCV are reduced by 95.79%, 93.26%, and 95.62%, respectively. In addition, ε and δ provided by our domain scientists for DCA are very close to the parameters that are automatically generated by BCA, as shown in Table I (P_{ID}: patient identifier; N_V: the number of cortical voxels for coclustering; $\mathrm{OWCV}_{\mathrm{DCA}}$, $\varepsilon_{\mathrm{DCA}}$ and δ_{DCA}: OWCV, ε and δ produced by DCA; $\varepsilon_{\mathrm{BCA}}$ and δ_{BCA}: ε and δ produced by BCA. $\mathrm{OWCV}_{\mathrm{Init}}$, $\mathrm{OWCV}_{\mathrm{Split}}$, $\mathrm{OWCV}_{\mathrm{Transfer}}$ and $\mathrm{OWCV}_{\mathrm{BCA}}$: OWCV produced by Initialization, Split, Transfer, and Merge operators). Overall, the experiments have shown that our proposed BCA can not only systematically optimize parameters for a population of data varied in size and density, but also produce much better coclustering results than those of DCA.

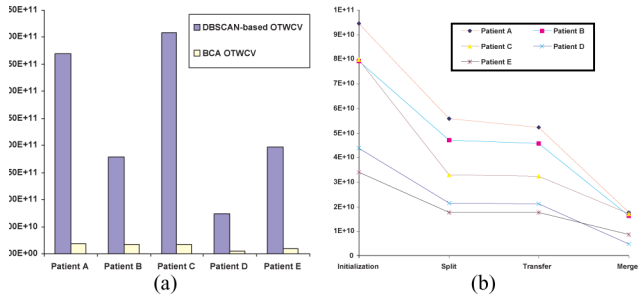


Fig. 11. (a) Comparison of OWCV between BCA and DCA; (b) evaluation of coclustering results at four phases.

Fig. 11(b) shows the decrease of OWCV after performing each operator of BCA. For patient P_A, OWCV reaches 9462832E + 4 at the initialization phase. It sharply goes down to 5599247E + 4 after running the Split operator, which is a 40.83% decrease from the initialized value. Then, OWCV continuously drops to 5230424E + 04 after the transfer phase, a 6.59% decrease from the prior phase. Finally, OWCV value is 65.97% reduced after applying the merge operator. Fig. 11(b) also shows OWCV at each phase for patient P_B to P_E and these results demonstrate the similar trend of the optimization performance as P_A. The average decrease rate by Split, Transfer, and Merge operator is around 47.79%, 1.12%, and 61.07%, respectively. It is expected that the transfer operator produces much smaller effects on OWCV than other operators, as only few voxels near the margin of clusters might need to be transferred to a new cluster. The reducing effects on OWCV by performing the split and merge operators are varied in datasets, but their standard derivation derived from the coclustering results are 0.076 and 0.119. It means that efforts of applying each operator remain consistent to different datasets. The consistent trend in decreasing OWCV across a population dataset also demonstrates that the consistent coclustering patterns exist in the population of human fiber tracts.

2) 3-D Visualization of the BCA Experimental Results: Fig. 12(a) shows all cortical voxels before coclustering and Fig. 12(b) shows the corticocortical fiber tracts following fiber

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

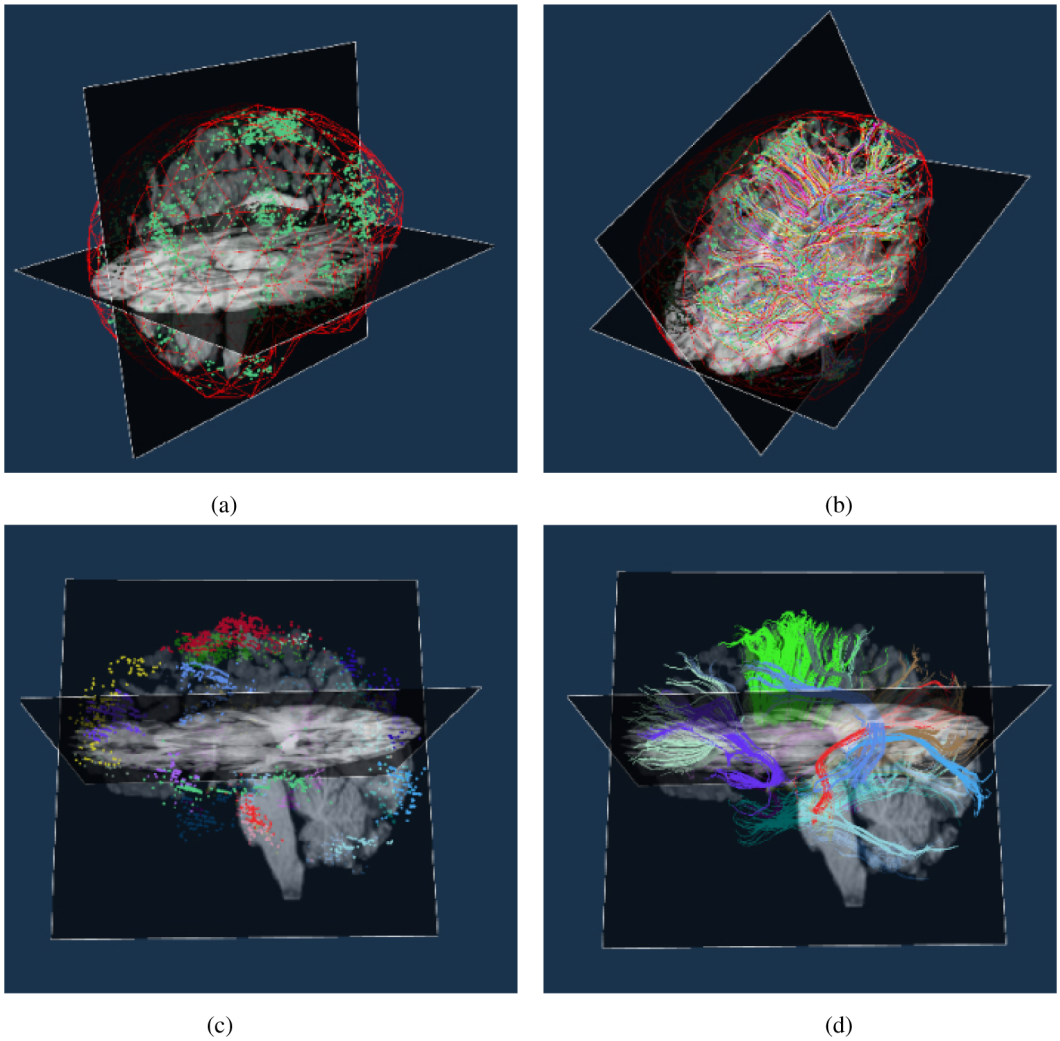


Fig. 12. (a) Cortical voxels with each voxel associated with a cortical volume element; (b) Cortico-cortical fiber connections of at least 100 mm long with both ending points being cortical voxels under depth of 15 mm; (c) All cortical voxels on a normal brain are visualized based on our coclustering results generated by BCA, and the voxels in one cocluster are rendered by one color; (d) all fiber tracts of the brain are visualized (in a coronal view) based on the coclustering results. No anatomical constraints are employed except for the side of the hemisphere.

tracking. The BCA algorithm partitions all fiber tracts into a number of automatically determined coclusters and each cluster is then rendered using one color, as shown in Fig. 12(c). Subsequently, all fibers constituting a particular cocluster are rendered in the same color, as displayed in Fig. 12(d). Our visualization results show that major fiber bundles are well identified using BCA even in the absence of anatomical constraints. For example, in Fig. 13, the cocluster created between regions 8 and 9 represents fiber tracts that connect the two hemispheres through the corpus callosum and the cocluster created between regions 1 and 3 represents the inferior longitudinal fasciculus. The quantitative evaluation of coclustering results is described in the subsequent sections.

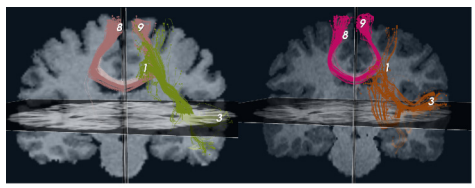
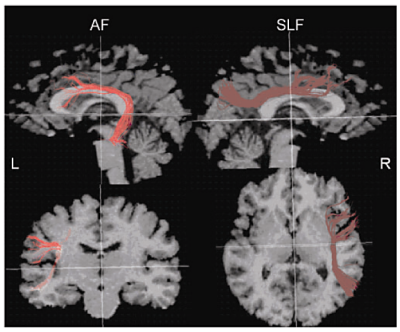


Fig. 13. Automatic labeling of the interested regions for groups of subjects. The cocluster between regions 8 and 9 show the connectivity patterns in the corpus callosum, and the cocluster between regions 1 and 3 show the connectivity patterns in inferior longitudinal fasciculus. Coclustering results are visualized and compared with different subjects.

3) Comparison of BCA Results to Manual Definition of Major Fiber Tracts: In order to assess the validity of the created coclusters, we compared the results of two major fiber tracts, the arcuate fasciculus (AF) and the superior longitudinal fasciculus (SLF) (Fig. 14). The fiber tracts were initially obtained from the BCA algorithm and subsequently derived based on manual definition using anatomical landmarks within the brain. In order to quantify the correlation between the two methods, we determined for each fiber tract the number of fibers originating/terminating in each of the 512 cortical elements. Subsequently, Pearsons correlation was calculated between the

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

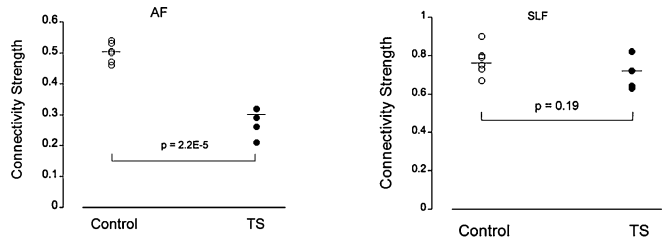


(a) (b)

Fig. 14. 3-D rendering of two major fiber tracts in the brain of a control subject. Fiber tracts derived from deterministic fiber tracking are displayed in the same space as high-resolution T1-weighted MR image volumes. The upper row shows sagittal views whereas the lower row shows coronal and axial views, respectively, for improved spatial orientation. The arcuate fasciculus (a) (on the left, rendered in red) connects cortical language areas while the superior longitudinal fasciculus (b) (SLF, on the right, rendered in brown) connects cortical areas responsible for selection and retrieval of spatial information.

TABLE II
CONNECTIVITY STRENGTH OF AF AND SLF FIBER TRACTS FOR FIVE TS SUBJECTS

P_{ID}	AF	SLF
A	0.21	0.72
B	0.32	0.63
C	0.26	0.64
D	0.29	0.82
E	0.32	0.82



(a) (b)

Fig. 15. Comparison of connectivity strength between the control and TS group for the arcuate fasciculus (left) and the SLF (right). Connectivity strength differs significantly between the two groups in the arcuate fasciculus but is similar in the SLF.

two 512-element vectors for both fiber tracts. The results of this correlation analysis are listed as follows. The correlation between the two methods was determined as 0.83 ± 0.09 for the AF indicating an excellent agreement between the BCA method and manual fiber tract definition. The correlation between the two methods for the SLF was found to be 0.65 ± 0.11. This value is lower, which is consistent with the greater physiological variability of this fiber bundle. These results indicate a very good overall agreement between the two methods.

4) Quantitative Evaluation of Coclustering Results: To allow quantitative assessment of the magnitude with which two cortical regions are connected, we defined a *connectivity strength* (CS) measure defined next. The CS measure allows for comparison of fiber tract connectivity among subjects and is a prerequisite for cross-subject analysis. Thus, given the mean and variance of the CS measure in a control group, abnormally high or low fiber connectivity can be objectively diagnosed in patients.

Definition 5.2 [Connectivity Strength (CS)]: Given a partition $C = \{C_1, C_2, \ldots, C_K\}$ of $G(V, F)$, the connectivity strength for a cocluster $\langle C_i, C_j \rangle$ is defined as

$$\mathrm{CS}(|C_i|, |C_j|, N_{C_{ij}}, N_{C_{ji}}) = \frac{(|C_i| + |C_j|)(N_{C_{ij}} + N_{C_{ji}})}{2|C_i||C_j|} \tag{13}$$

where $N_{C_{ij}}$ is the number of fibers from cluster C_i to cluster C_j and $N_{C_{ji}}$ is the number of fibers from cluster C_j to cluster C_i. $|C_i|$ and $|C_j|$ are the number of voxels in cluster C_i and C_j, respectively.

5) Application of the BCA Algorithm in Clinical Research: To evaluate the performance of the BCA in clinical research, we compared the connectivity strength of two major fiber tracts, the AF and the major part of the SLF between a group of children suffering from Tourette syndrome (TS) and an age-matched control group. TS is a childhood-onset neuropsychiatric disorder characterized by the presence of motor and vocal tics, together with a range of semicompulsive behaviors. Functional disturbances in the frontostriatal thalamic circuit [28] are thought to be pivotal in producing TS symptoms; however, it is likely that the arcuate fasciculus is implicated as well [29]. The arcuate fasciculus connects cortical areas responsible for language comprehension (Wernicke's area) and language generation (Broca's area) in the dominant (in right-handed subjects the left) hemisphere (see Fig. 14)), whereas the SLF connects the prefrontal with the parietal cortex and provides information regarding perception of visual space. As both the spatial attention and the regulatory mechanisms for selection and retrieval of spatial information are believed to be normal in TS, we expected that the connectivity strength of AF is an important parameter in differentiating normal and abnormal subjects, while the connectivity strength of the SLF should not differ between the two groups. Initially, coclusters representing the AF and the SLF were selected and the corresponding connectivity strength was calculated according to Definition 5.2.

To test our hypothesis, we performed cross-subject statistical analysis based on our coclustering results. Based on the statistical results on the normals, we identify the normative distribution for both the fiber tracts. The normal variability for AF has a mean of 0.50 with standard deviation of 0.03. Corresponding variablity of SLF has a mean of 0.77 with a standard derivation of 0.07. Any significant difference from this normative distribution is considered abnormal. Significant differences between the connectivity strengths obtained from the two groups were assessed using an independent sample *t*-test. Table II shows connectivity strength values determined in the two groups. As hypothesized, our results indicate a significantly lower connectivity strength of the left arcuate fasciculus between TS patients and controls (0.30 ± 0.09 versus 0.50 ± 0.03, $p = 2.2\mathrm{E} - 5$), but a similar connectivity strength between the two groups in

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.

the SLF (0.73 ± 0.09 versus 0.77 ± 0.08, $p = 0.19$), as shown in Fig. 15).

VI. Conclusion and Future Work

In this paper, we defined the coclustering problem and developed a coclustering algorithm, termed the BCA, which has been applied to *in vivo* fiber tract analysis of corticocortical connectivity patterns. Our study suggests that the developed BCA is well suited for the segmentation of major fiber tracts from DTI data, both in its computational stability and its processing speed. Moreover, the obtained results were consistent with well-known fiber tracts and correlated well with manually defined major fiber bundles. Our framework facilitates automatic analysis of tractography data for a large number of human subjects in clinical research and is likely to aid in the integration of genetic information with human image-derived phenotype data. Future research will also focus on translating the current version of the BCA to the analysis of fiber tracts derived from probabilistic fiber tracking methods, permitting also quantitative assessment of minor fiber tracts.

References

[1] R. Passingham, K. Stephan, and R. Kötter, "The anatomical basis of functional localization in the cortex," *Nat. Rev. Neurosci.*, vol. 3, no. 8, pp. 606–616, 2002.

[2] J. Han and M. Kamber, *Data Mining* (ser. I-55860-489-8). San Mateo, CA: Morgan Kaufmann, 2001, pp. 349–353.

[3] R. Ng and J. Han, "Efficient and effective clustering methods for spatial data mining," in *Proc. VLDB*, 1994, pp. 144–155.

[4] G. Karypis, E. Han, and V. Kumar, "Chameleon: A hierarchical clustering algorithm using dynamic modeling," *Computer*, vol. 32, no. 8, pp. 68–75, Aug. 1999.

[5] S. Guha, R. Rastogi, and K. Shim, "Cure: An efficient clustering algorithm for large databases," in *Proc. SIGMOD Conf.*, 1998, pp. 73–84.

[6] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, 1996, pp. 226–231.

[7] A. Hinneburg and D. Keim, "An efficient approach to clustering in large multimedia databases with noise," in *Proc. KDD*, 1998, pp. 58–65.

[8] M. Ankerst, M. Breunig, H. Kriegel, and J. Sander, "Optics: Ordering points to identify the clustering structure," in *Proc. SIGMOD Conf.*, 1999, pp. 49–60.

[9] C. Lin, S. Lu, D. Wu, J. Hua, and O. Muzik, "Coclustering based parcellation of human brain cortex using diffusion tensor MRI," in *Lecture Notes in Computer Science*. vol. 4463, New York: Springer-Verlag, 2007, pp. 539–550.

[10] D. L. Bihan, "Looking into the functional architecture of the brain with diffusion MRI," *Nat. Rev. Neurosci.*, vol. 4, pp. 469–480, 2003.

[11] T. Conturo, N. Lori, T. Cull, E. Akbudak, A. Snyder, J. Shimony, R. McKinstry, H. Burton, and M. Raichle, "Tracking neuronal fiber pathways in the living human brain," *Appl. Phys. Sci./Neurobiol.*, vol. 96, no. 18, pp. 10 422–10 427, 1999.

[12] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "*In vivo* fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 4, pp. 625–632, 2000.

[13] H. Jiang, P. VanZilj, J. Kim, G. Pearlson, and S. Mori, "Dtistudio: Resource program for diffusion tensor computation and fiber bundle tracking," *Comput. Methods Program. Biomed.*, vol. 81, no. 2, pp. 106–116, 2006.

[14] Z. Ding, J. Gore, and A. Anderson, "Classification and quantification of neuronal fiber pathways using diffusion tensor MRI," *Magn. Reson. Med.*, vol. 49, pp. 716–721, 2003.

[15] S. Zhang and D. Laidlaw, "DTI fiber clustering in the whole brain," in *Proc. IEEE Vis. 2004*, Oct., p. 28.

[16] M. Maddah, A. Mewes, S. Haker, W. Grimson, and S. Warfield, "Automated atlas-based clustering of white matter fiber tracts from DTMRI," in *Proc. MICCAI*, 2005, pp. 188–195.

[17] L. O'Donnell and C. Westin, "White matter tract clustering and correspondence in populations," in *Proc. MICCAI*, 2005, pp. 140–147.

[18] L. Jonasson, P. Hagmann, J. Thiran, and V. Wedeen, "Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering," in *Proc. ISMRM*, 2005, p. 1310.

[19] A. Brun, H. Knutsson, H. Park, M. Shenton, and C. Westin, "Clustering fiber traces using normalized cuts," in *Proc. MICCAI*, 2004, pp. 368–375.

[20] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[21] I. Corouge, S. Gouttard, and G. Gerig, "Towards a shape model of white matter bundles using diffusion tensor MRI," in *Proc. Int. Symp. Biomed. Imag.: Nano Macro*, 2004, pp. 344–347.

[22] J. Clayden, A. Storkey, and M. Bastin, "A probabilistic model-based approach to consistent white matter tract segmentation," *IEEE Trans. Med. Imag.*, vol. 26, no. 11, pp. 1555–1561, Nov. 2007.

[23] C. Lin, S. Lu, X. Liang, J. Hua, and O. Muzik, "Cocluster analysis of thalamo-cortical fiber tracts extracted from diffusion tensor MRI," *Int. J. Data Mining Bioinformatics*, vol. 2, no. 4, pp. 342–361, 2008.

[24] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, 1996, pp. 226–231.

[25] A. Hayter, *Probability and Statistics for Engineers and Scientists*, 2nd ed. ed. Belmont, CA: Duxbury Thomson Learning, 2002.

[26] P. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from NMR spin echo," *J. Magn. Reson.*, vol. 103, pp. 247–254, 1994.

[27] S. Mori, B. Crain, V. Chacko, and P. vanZijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, no. 2, pp. 265–269, 1999.

[28] H. Singer and K. Minzer, "Neurobiology of Tourette's syndrome: Concepts of neuroanatomic localization and neurochemical abnormalities," *Brain Dev.*, vol. 25, pp. 70–84, 2003.

[29] L. D. Nil, J. Sasisekaran, P. V. Lieshout, and P. Sandor, "Speech disfluencies in individuals with Tourette syndrome," *J. Psychosom. Res.*, vol. 58, no. 1, pp. 97–102, 2005.

Authors' photographs and biographies not available at the time of publication.

Authorized licensed use limited to: Wayne State University. Downloaded on July 23,2010 at 06:50:03 UTC from IEEE Xplore. Restrictions apply.