

Deformable Registration of Biomedical Images Using 2D Hidden Markov Models

Renuka Shenoy, *Student Member, IEEE*, Min-Chi Shih, *Student Member, IEEE*, and Kenneth Rose, *Fellow, IEEE*

Abstract—Robust registration of unimodal and multimodal images is a key task in biomedical image analysis, and is often utilized as an initial step on which subsequent analysis techniques critically depend. We propose a novel probabilistic framework, based on a variant of the 2D hidden Markov model, namely, the turbo hidden Markov model, to capture the deformation between pairs of images. The hidden Markov model is tailored to capture spatial transformations across images via state transitions, and modality-specific data costs via emission probabilities. The method is derived for the unimodal setting (where simpler matching metrics may be used) as well as the multimodal setting, where different modalities may provide very different representations for a given class of objects, necessitating the use of advanced similarity measures. We utilize a rich model with hundreds of model parameters to describe the deformation relationships across such modalities. We also introduce a local edge-adaptive constraint to allow for varying degrees of smoothness between object boundaries and homogeneous regions. The parameters of the described method are estimated in a principled manner from training data via maximum likelihood learning, and the deformation is subsequently estimated using an efficient dynamic programming algorithm. Experimental results demonstrate the improved performance of the proposed approach over the state-of-the-art deformable registration techniques, on both unimodal and multimodal biomedical data sets.

Index Terms—Registration, deformable, multimodal, hidden Markov models.

I. INTRODUCTION

MODERN advances in imaging technology have enabled the collection of huge amounts of image data of complex biological systems. The extraction of information from

this image data and subsequent analysis and interpretation on the information are the central tasks in the field of biomedical image computing. Due to the large size of these datasets, manual annotation and analysis is usually not feasible - the development of robust automated and semi-automated techniques is critical for analysis and diagnosis. Image registration [1], which aims at accurately aligning structures or regions across related images, is an important problem in biomedical image computing, and is an active area of current research.

Image registration is used in a variety of applications in biomedical image analysis. Assessing the efficacy of treatments requires registration in order to accurately compare pre-treatment and post-treatment scans. Atlas-based methods, which analyze subject data in comparison to one or more standard models, critically depend on accurate registration. Often, different types of information, eg., structural and functional, can be extracted from different image modalities, and registering these multimodal images is an important step in combining the complementary sources of information. Alignment of structures across temporal or depth-based volumes are important in time-lapse and 3D network reconstruction. Hence, accurate deformable registration [2] is the need of the hour.

This paper is organized as follows: In section II, we present an overview of related work in biomedical image registration, ranging from early work to recent techniques. Section III presents a method for deformable registration with a generic theoretical formulation based on the turbo hidden Markov model [3]. The method is first derived for unimodal data and then extended to the multimodal, multi-channel setting. We also describe an edge-adaptive constraint which allows for variation in degree of smoothness across the image. In section IV, we provide experimental evaluation of the proposed algorithm, for unimodal registration of MRI data as well as multimodal registration of connectome data, in comparison with related techniques in literature, and discuss experimental results. Finally, in section V, we provide concluding remarks.

II. BACKGROUND AND RELATED WORK

Registration is a classic problem in biomedical image analysis and has been widely studied over the past three decades. Early research in unimodal registration focused on rigid alignment [4], [5]. Due to the nature of the unimodal registration problem, the sum of squared differences (SSD) [6]–[8] and the mean squared difference (MSD) [9], [10] of image intensities have been widely and successfully used to measure data similarity.

Manuscript received November 11, 2015; revised May 2, 2016; accepted June 21, 2016. Date of publication July 18, 2016; date of current version August 19, 2016. This work was supported by the National Science Foundation (NSF) through the Division of Information and Intelligent Systems under Grant III 0808772, in part by NSF through the Office of Integrative Activities under Grant OIA 0941717, and in part by NSF through the Division of Computing and Communication Foundations under Grant CCF 1320599. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Oleg V. Michailovich.

R. Shenoy was with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA. She currently resides in Santa Clara, CA 95051 USA (e-mail: renuka@ece.ucsb.edu).

M.-C. Shih was with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA. He is now with MediaTek Inc., San Jose, CA 95134 USA (e-mail: minchi_shih@ece.ucsb.edu).

K. Rose is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: rose@ece.ucsb.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/TIP.2016.2592702

Early work on multimodal registration used intensity levels [11] and joint entropy [12] as the measure of information across modalities. The correlation ratio [13] and normalized cross correlation [14] are amongst the popular similarity measures in computing optimal transformations. Mutual information (MI), which measures the statistical dependence of two random variables, was proposed in [15]–[17] for rigid alignment problems and quickly gained popularity as a measure of similarity for both rigid and deformable registration, and in both unimodal and multimodal settings. (A survey of various MI-based methodologies for medical image registration is available in [18].) In [19], a deformable registration technique was introduced which uses both a global affine transformation and local transformations, with MI as a measure of information. The local transformations are modeled with Free Form Deformations (FFD) on a non-rigid lattice of control points using cubic B-splines for interpolation. Several subsequent papers leveraged FFDs, extending their use to multimodal data [20], [21] and using variants of MI (such as Conditional MI (CMI) in [22] and Normalized MI (NMI) in [23]). In [24] and subsequently [25], pixelwise matching is performed using a cost function based on MI under a global smoothness constraint. Dense deformable registration is performed by modeling it as a minimal cost graph problem on a Markov random field (MRF) built upon the FFD framework in [26] and this idea is further developed in [27]. Smoothness constraints are imposed through connectivity of nodes and labels correspond to deformations. For the multimodal case, MI (or a variant of MI) is used as the matching criterion.

A significant drawback of these techniques lies in the estimation of the joint histogram between modalities where one or both of the modalities proffers multiple channels of information. A natural extension of MI-based approaches to include data with multiple channels (e.g., RGB data) would be to use multivariate MI. However, the complexity of populating higher dimensional joint histograms grows exponentially with the number of channels, and these methods quickly become impractical for multi-channel modalities. Further, inadequate population of such high dimensional histograms due to sparse availability of data could lead to inaccuracies in the inferred deformation.

Some notable prior research has been devoted to the problems associated with estimation of multivariate MI. A simplifying approximation of the general multivariate MI was proposed as a similarity measure in [28]. In [29], an entropic graph-based implementation was used to estimate α -MI of multiple channels.

The concept of self-similarity was introduced as a means for multimodal registration in two recent papers, [30] and [31], which introduce the modality independent neighborhood descriptor (MIND) and the self-similarity context (SSC) descriptor respectively. These descriptors exploit local structural similarities between multimodal image pairs and are calculated on a defined spatial search region in each modality. SSD is used as a measure of distance between descriptors, and final deformation is estimated using Gauss-Newton optimization. However, as acknowledged in [30], a limitation of such approaches is that they require an anatomical

feature to be present in both modalities. While this may be a reasonable requirement in the case of structurally similar modalities such as MRI/CT, some multimodal datasets, such as connectome [32] data, do not exhibit such distinct features between their modalities. In addition, these methods require modification to be used with multi-channel inputs.

Our preliminary work [33] exploited the probabilistic nature of the 2D hidden Markov model (HMM) framework to perform multimodal fusion on connectome data. To the best of our knowledge, this work represents the first effort to apply 2D HMMs to the deformable registration problem. (A recent paper [34] utilizes HMMs for deformable registration; however the HMM is used in its conventional 1D form, and the method is domain-specific to computed tomography angiography.) In this paper, we provide a detailed view of the registration algorithm as well as additional experimental validation. We begin by developing the HMM-based registration framework for the basic setting of deformable registration, i.e., the single modality problem. We then show how the framework may be extended to the multimodal setting. Further, we show how to add the capability to handle local differences in the degree of smoothness within an image.

Our primary motivation was to develop a method for 2D-2D deformable registration which was applicable to the general case of multimodal, multi-channel input images. The hidden Markov model framework provides a platform to estimate parameters in a principled manner even from unlabeled training data (i.e., source-target image pairs without labeled dense point-wise correspondences). We note the state-of-the-art approaches for the deformable registration problem which we compare against ([27]–[31]) require parameters of the registration model to be manually set. However, in the proposed approach, only the hyperparameters of the HMM framework need to be specified. The data-specific parameters of the model, describing both the inter-modal data similarity and the degree of smoothness, are automatically learned from training data. This allows us to define rich models (with the number of model parameters ranging in the order of hundreds). Therefore, we can effectively register even challenging datasets having complex multimodal relationships, such as the cellular microscopy modality described in Section IV-C.

III. PROPOSED APPROACH

Our aim is to find the deformation that best explains the relation between one image (the “source”) and a second image, from the same or different modality (the “target”). We propose a probabilistic method that estimates the global deformation with a set of local deformations. There is a clear trade-off between flexibility in local deformations so that high accuracy is achieved in discerning the true structural relationship between the source and the target, and the need to impose global coherence and avoid highly “non-smooth” deformations. Moreover, every local transformation results in a mapping between the two modalities and there must be a way to measure and maximize the goodness of this match within the constraints of the deformation framework.

Translation consistency in neighborhoods and cross-modality matching costs are embedded into a 2D HMM built

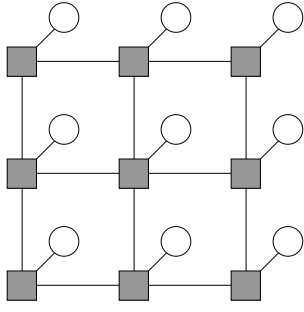


Fig. 1. A two-dimensional hidden Markov model build on a first order Markov mesh random field, with hidden states (grey squares) and observations (circles).

on a first order Markov mesh random field (MMRF). Local translations are “hidden” and their impact is felt through the corresponding data matching costs. The parameters of these data matching costs are data dependent, and are learned from source-target pairs of training images. In addition, the degree of smoothness of the optimal deformation can differ considerably across different types of data, and hence the training images are also utilized in learning smoothness parameters.

A. 2D Hidden Markov Models and the Turbo Approximation

Conventional (1D) HMMs have been widely used in various applications. This has been feasible due to the existence of computationally efficient algorithms for the central problems associated with HMMs [35] - inferring the optimal state sequence given observation data and a model, and learning the parameters of the model. However, direct extension of these to 2D leads to exponential increase in complexity and is intractable in most practical applications.

There has been considerable interest in developing algorithms that approximate the performance of 2D HMMs at a lower complexity. Most approaches approximate 2D HMMs with one [36] or more [3], [37] 1D HMMs. The turbo hidden Markov model (T-HMM) introduced in [3] consists of horizontal and vertical 1D HMMs that are decoded separately but “communicate” in a manner similar to decoding of turbo codes [38]. Each row of the 2D HMM is represented by a 1D HMM during a horizontal pass, and each column by a 1D HMM during a vertical pass (see Fig. 2). Alternate horizontal and vertical passes induce prior probabilities on each other, and are repeated until they converge to a required degree of agreement. Optimal state sequences are estimated via delayed decisions on posterior probability vectors at each node, rather than a greedy “winner-take-all” scheme. The T-HMM framework provides efficient approximations for both learning [39] and inference [3], and has been used in applications such as speech presence detection [40] and face recognition [41]. In addition, it has been shown [3] to outperform other 2D HMM approximations. The T-HMM is used for both learning and inference in the proposed approach.

B. Data Similarity Measure

The probability of matching a feature vector in the source image to a feature vector in the target image is captured by the emission probabilities of the HMM.

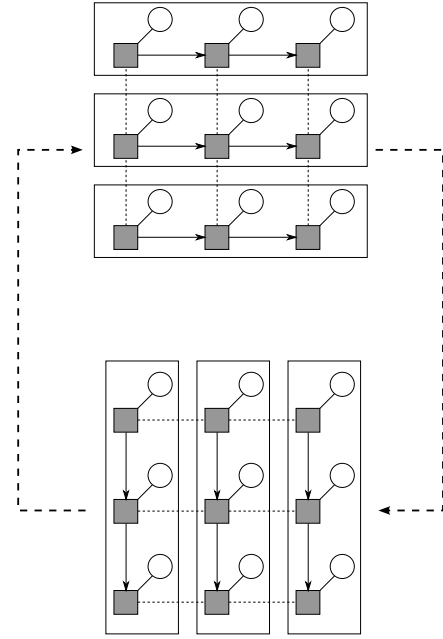


Fig. 2. Turbo decoding of a 2D HMM, where alternating horizontal and vertical passes consist of separately decoded rows and columns that induce priors on each other.

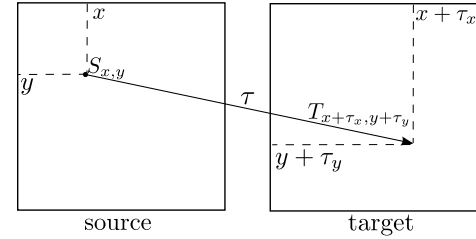


Fig. 3. Mapping feature vector $S_{x,y}$ at point (x, y) in the source image to a feature vector $T_{x+\tau_x, y+\tau_y}$ at point $(x + \tau_x, y + \tau_y)$ in the target image using a translation τ .

In our design, each state, q , of the HMM corresponds to a specific, unique translation τ relating the source and target, and whose components are τ_x and τ_y in the x - and y -directions, respectively. A state with translation τ maps a point (x, y) in the source to $(x + \tau_x, y + \tau_y)$ in the target. Therefore, the emission probability $b_{x,y}^\tau$ represents the probability of matching the source feature vector at (x, y) , given by $S_{x,y}$, to the target feature vector at the given translation, given by $T_{x',y'}$. (see Fig. 3.)

$$b_{x,y}^\tau = P(T_{x',y'} | S_{x,y}) \quad (1)$$

This probability can be modeled in various ways in single modality problems, for example, using a simple similarity metric such as sum of absolute differences in intensity or correlation on patches from source and target images. In our experiments, we use a term based on SSD of intensities to model emission probability.

$$b_{x,y}^\tau = \frac{1}{\sqrt{2\pi}\sigma_{SSD}} \exp \left\{ -\frac{\epsilon_{x,y}^\tau}{2\sigma_{SSD}^2} \right\} \quad (2)$$

where $\epsilon_{x,y}^\tau$ is the sum of squared differences in sliding windows of size $(2W + 1) \times (2W + 1)$ centered at the points

of interest in the source and target images.

$$\epsilon_{x,y}^{\tau} = \sum_{x_w, y_w = -W}^W (\mathbf{S}_{x+x_w, y+y_w} - \mathbf{T}_{x'+x_w, y'+y_w})^2 \quad (3)$$

C. Extending to Multimodal Data

Since different modalities may vary greatly in the way they represent objects from the same class, we cannot directly apply intensity-based similarity measures for registration of multimodal data. Instead, we assume that the feature vectors of the source and target images at specific locations are not directly related, but rather, related only through the object type at the corresponding locations in the underlying “true” arrangement. In other words, the source feature vector, the underlying object type at the corresponding location in the source ($\omega_{x,y}^S$), the underlying object type at the location after translation in the target ($\omega_{x',y'}^T$) and the target feature vector form a Markov chain.

$$\mathbf{S}_{x,y} \longleftrightarrow \omega_{x,y}^S \longleftrightarrow \omega_{x',y'}^T \longleftrightarrow \mathbf{T}_{x',y'} \quad (4)$$

In our model, the underlying object types and their spatial relations are hidden, information from them can only be extracted from the observable features in the images from each modality.

We learn the distribution of the source feature vectors, and rather than making a hard decision on object type at each node, we obtain its posterior probability. Thus, for data consisting of objects belonging to M classes, we learn $P(\omega_m | \mathbf{S}_{x,y})$ for each class $m \in \{1, 2, \dots, M\}$ at every node. Applying the law of total probability to (1) under the Markov assumption (4) results in:

$$b_{x,y}^{\tau} = \sum_{m=1}^M P(\omega_m | \mathbf{S}_{x,y}) P(\mathbf{T}_{x',y'} | \omega_m) \quad (5)$$

Since linear combinations of Gaussians can approximate arbitrarily shaped densities, we use a mixture of K Gaussians to model $P(\mathbf{T}_{x',y'} | \omega_m)$ for each object class ω_m .

$$P(\mathbf{T}_{x',y'} | \omega_m) = \sum_{k=1}^K w_m^k P(\mathbf{T}_{x',y'} | \omega_m^k) \quad (6)$$

where mixture component weights must satisfy the constraint:

$$\sum_{k=1}^K w_m^k = 1 \quad \forall m \in \{1, 2, \dots, M\} \quad (7)$$

Each individual component density is a Gaussian having dimensionality D , equal to that of the target feature space.

$$P(\mathbf{T}_{x',y'} | \omega_m^k) = \frac{\exp\{-\frac{1}{2}(\mathbf{T}_{x',y'} - \mu_m^k)^T \Sigma_m^{k-1} (\mathbf{T}_{x',y'} - \mu_m^k)\}}{(2\pi)^{\frac{D}{2}} |\Sigma_m^k|^{\frac{1}{2}}} \quad (8)$$

where μ_m^k and Σ_m^k are the mean and covariance of the Gaussian respectively.

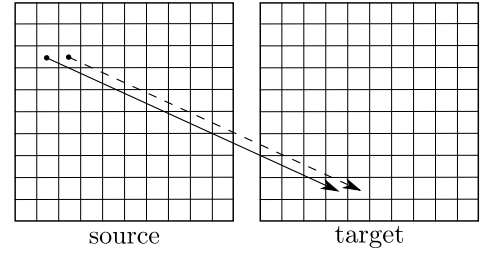


Fig. 4. Neighboring translations are correlated.

D. Deformation Smoothness Model

Each state of the HMM corresponds to a translation τ relating the source and target. The translations of neighboring nodes are correlated as quantified by the transition probabilities of the HMM (see Fig. 4). Equivalently, in a first order Markov mesh random field (MMRF), the state of a node depends on the state of its adjacent neighbors in the horizontal and vertical directions. While an arbitrary transition probability matrix may be used, we introduce some assumptions in order to reduce the number of free parameters in the system. These are outlined below.

Assuming a stationary HMM, the transition probabilities of the horizontal and vertical 1D HMMs are given by

$$a^H(\tau, \tau') = P(q_{x,y} = \tau' | q_{x,y-1} = \tau)$$

and

$$a^V(\tau, \tau') = P(q_{x,y} = \tau' | q_{x-1,y} = \tau)$$

respectively. In the case of cellular microscopy images, we make the simplifying assumption that parameters of horizontal and vertical HMMs are identical as these images typically do not exhibit directionality along coordinate axes (as is often the case in faces [39], man-made scenes, and certain natural images):

$$a^V(\tau, \tau') = a^H(\tau, \tau') = a(\tau, \tau') \quad (9)$$

We also impose shift invariance so that probability of moving from one state to another only depends on the difference in the corresponding translations:

$$a(\tau, \tau') = a(\tau' - \tau) = a(\delta\tau) \quad (10)$$

where $\delta\tau = [\delta\tau_x \ \delta\tau_y]^T$ and $\delta\tau_x$ and $\delta\tau_y$ are, respectively, the horizontal and vertical components of the difference in translations.

Further, we restrict ourselves to parametric transition probabilities to increase robustness. For a Gaussian model with a covariance matrix Σ , the general expression is

$$a(\delta\tau) \propto \exp\left\{-\frac{1}{2}(\delta\tau)^T \Sigma^{-1} (\delta\tau)\right\} \quad (11)$$

Noting again the lack of consistent directionality in microscopy data, we simplify to an isotropic model with a single variance parameter σ^2 .

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Incorporating the isotropic model into (11) results in a simplified expression for transition probability, given below.

$$a(\delta\tau) \propto \exp \left\{ -\frac{1}{2} \left(\frac{\delta\tau_x^2 + \delta\tau_y^2}{\sigma^2} \right) \right\} \quad (12)$$

E. Edge-Adaptive Smoothness Constraint

For data consisting of multiple objects, it is often observed that large translations occur more frequently near object boundaries than well inside object. This is usually due to the choice of the model rather than a true biological mapping across images. Similar behavior is noticed while computing optical flow in images with multiple objects, and has been addressed by including an structure-adaptive regularization constraint in the cost function [42], [43]. We model this variation in smoothness by introducing a spatially varying transition matrix, parametrized by two Gaussians. The transition probability matrix at each point is calculated using local “edgeness” at that point.

To calculate the edgeness at a given point (x, y) , we consider the set of all estimated object labels in a window centered at that point. By normalizing the histogram of all unique labels l in the window, we obtain $[P(l)]$, the local label probability vector. The label entropy around the point (x, y) is given by

$$\mathcal{H}_{x,y} = - \sum_l P(l) \log_2 P(l) \quad (13)$$

The label entropy in each window measures label uncertainty in the window and is used to estimate edgeness. We thus quantify edgeness \mathcal{E} at (x, y) as

$$\mathcal{E}_{x,y} = 1 - e^{-\mathcal{H}_{x,y}} \quad (14)$$

Note that this measure of edgeness approaches zero at low entropy, and approaches one at high entropy. The transition matrix at each point is modeled as a linear combination of two Gaussian-parametrized matrices.

$$a_{x,y}(\delta\tau) = \mathcal{E}_{x,y} a^E(\delta\tau) + (1 - \mathcal{E}_{x,y}) a^I(\delta\tau) \quad (15)$$

where $a^E(\delta\tau)$ denotes the transition matrix for points on cell edges and $a^I(\delta\tau)$, that of interior points.

$$a^E(\delta\tau) \propto \exp \left\{ -\frac{1}{2} \left(\frac{\delta\tau_x^2 + \delta\tau_y^2}{\sigma_E^2} \right) \right\} \quad (16)$$

$$a^I(\delta\tau) \propto \exp \left\{ -\frac{1}{2} \left(\frac{\delta\tau_x^2 + \delta\tau_y^2}{\sigma_I^2} \right) \right\} \quad (17)$$

Hence, we allow for flexibility in the transition probability matrix according to the edgeness of the point in question.

F. Estimation of Deformation Field

We infer the optimal state sequence using the Viterbi algorithm with the modified forward-backward iterations described in [3].

G. Parameter Estimation

Baum-Welch Training: The parameters of the HMM are estimated in an unsupervised manner from source-target image pairs. In the Baum-Welch algorithm ([44], [45]) HMM parameters are iteratively optimized to increase the value of likelihood. Re-estimation formulas for these parameters are derived by maximizing Baum’s auxiliary function, given by

$$\mathcal{Q}(\lambda'|\lambda) = \sum_Q P(Q|S, T, \lambda) \log P(S, Q|T, \lambda') \quad (18)$$

with respect to λ' , where λ denotes the current estimate of HMM parameters, λ' the model re-estimate and Q , a sequence of states $Q = \{q_{x,y}, x = 1, 2, \dots, X, y = 1, 2, \dots, Y\}$. S and T denote corresponding source and target images. As noted in [35], this re-estimation procedure can be interpreted as an implementation of the Expectation-Maximization (EM) algorithm [46] and is known to monotonically increase likelihood. The Baum-Welch algorithm was originally derived for parameter estimation in 1D HMMs, but is easily extended to the 2D case with the modified forward-backward iterations of the T-HMM.

During the expectation step, we estimate the occupancy probabilities of the horizontal and vertical 1D HMMs,

$$\gamma_{x,y}^{H,\tau} = P(q_{x,y}^H = \tau | S, T, \lambda)$$

$$\gamma_{x,y}^{V,\tau} = P(q_{x,y}^V = \tau | S, T, \lambda)$$

and the overall occupancy probability

$$\gamma_{x,y}^\tau = \frac{\gamma_{x,y}^{H,\tau} + \gamma_{x,y}^{V,\tau}}{2}$$

We also estimate the ancillary training variables,

$$\xi_{x,y}^H(\tau, \tau + \delta\tau) = P(q_{x,y+1} = \tau + \delta\tau, q_{x,y} = \tau | S, T, \lambda)$$

$$\xi_{x,y}^V(\tau, \tau + \delta\tau) = P(q_{x+1,y} = \tau + \delta\tau, q_{x,y} = \tau | S, T, \lambda)$$

During the maximization-step, we maximize Baum’s auxiliary function with respect to each parameter to derive re-estimation formulas.

For the SSD-based metric, there is only one emission parameter, σ_{SSD} , whose re-estimation equation is given by

$$\hat{\sigma}_{SSD}^2 = \frac{\sum_{x,y,\tau} \gamma_{x,y}^\tau \epsilon_{x,y}^\tau}{\sum_{x,y,\tau} \gamma_{x,y}^\tau} \quad (19)$$

To update emission parameters for the multimodal case, we must first use the updated variables from the expectation step to calculate the per-component posterior probability at each node, given by

$$\phi_{x,y,m}^{\tau,k} = \frac{P(\omega_m | \mathbf{S}_{x,y}) w_m^k P(\mathbf{T}_{x',y'} | \omega_m^k)}{\sum_{m=1}^M P(\omega_m | \mathbf{S}_{x,y}) \sum_{k=1}^K w_m^k P(\mathbf{T}_{x',y'} | \omega_m^k)} \quad (20)$$

where $x' = x + \tau_x$ and $y' = y + \tau_y$.

Emission parameters are re-estimated using the following update equations:

$$\hat{w}_m^k = \frac{\sum_{x,y,\tau} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k}}{\sum_{x,y,\tau,k} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k}} \quad (21)$$

$$\hat{\mu}_m^k = \frac{\sum_{x,y,\tau} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k} \mathbf{T}_{x',y'}}{\sum_{x,y,\tau} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k}} \quad (22)$$

$$\hat{\Sigma}_m^k = \frac{\sum_{x,y,\tau} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k} (\mathbf{T}_{x',y'} - \hat{\mu}_m^k)(\mathbf{T}_{x',y'} - \hat{\mu}_m^k)^T}{\sum_{x,y,\tau} \gamma_{x,y}^\tau \phi_{x,y,m}^{\tau,k}} \quad (23)$$

For transition probability matrices parameterized by a single Gaussian, the update equation is given by:

$$\hat{\sigma}^2 = \frac{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')] [\delta\tau]^2}{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')]} \quad (24)$$

For edge-adaptive transition probabilities, the update equations are given by:

$$\hat{\sigma}_E^2 = \frac{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')] \left(\frac{\mathcal{E}_{x,y}}{a_{x,y}(\delta\tau)} \right) [\delta\tau]^2}{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')] \left(\frac{\mathcal{E}_{x,y}}{a_{x,y}(\delta\tau)} \right)} \quad (25)$$

$$\hat{\sigma}_I^2 = \frac{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')] \left(\frac{1 - \mathcal{E}_{x,y}}{a_{x,y}(\delta\tau)} \right) [\delta\tau]^2}{\sum_{x,y,\tau,\delta\tau} [\xi_{x,y}^H(\tau, \tau') + \xi_{x,y}^V(\tau, \tau')] \left(\frac{1 - \mathcal{E}_{x,y}}{a_{x,y}(\delta\tau)} \right)} \quad (26)$$

Transition matrices are initialized uniformly. Parameters relating to emission probabilities may be initialized either by learning the data from the target feature vectors independently of the source image, or by performing rigid matching between the source and the target to estimate the parameters. While training multiple Gaussians per cell class, we perform standard Expectation-Maximization for Gaussian mixture models (EM-GMM) on each class to learn parameters of each desired component, and use these for initialization. $P(\omega_m | \mathbf{S}_{x,y})$ is learned by applying EM-GMM on the source modality.

H. Complexity

The complexity of training and decoding under the T-HMM approximation is $O(N^2XY)$, where N is the number of states of the HMM and XY is the number of nodes in the HMM. N , in turn, is proportional to Δ^2 , where Δ is the maximum translation allowed per direction. The complexity of algorithm is hence $O(\Delta^4XY)$.

In order to reduce complexity, we employ a multi-resolution coarse-to-fine scheme, approximating a group of 4 nodes at

each resolution with a single node at the nearest coarser resolution. For The complexity at each level is $O(n^2X_LY_L)$, where L is the index of each level, taking values from 1 (the finest level) to L_{max} (the coarsest level), n is the number of states at each level of resolution (a constant), and X_LY_L is the number of nodes at the given level. Therefore,

$$X_LY_L = \frac{XY}{4^{(L-1)}}.$$

The total complexity considering all levels is $O(n^2XY(\sum_{L=1}^{L_{max}} 4^{-(L-1)}))$. Since $[\sum_{L=1}^{L_{max}} 4^{-(L-1)}] \approx 1.33$, the complexity of the hierarchical approach is $O(n^2XY)$, independent of Δ .

IV. EXPERIMENTS AND RESULTS

A. Evaluation Metrics

We present the performance of the proposed approach on two biomedical image datasets, both of which have manually-guided expert labeled ground truth segmentations. To evaluate registration quality, we generated automated segmentations by warping the source image segmentations using the transformation obtained from each method. The resulting automated segmentation (S_{trans}) was compared to the manual target segmentation (T) using the Dice similarity measure (DSC) [47] as a measure of overlap. The DSC is given by:

$$\text{DSC} = \frac{2|S_{trans} \cap T|}{|S_{trans}| + |T|} \quad (27)$$

where $|\bullet|$ denotes cardinality in terms of number of pixels.

In addition, manually labeled landmark pairs across source and target images are available for one of the datasets, and registration accuracy is quantified in terms of the target registration error (TRE) metric. Consider a landmark located at a point (x_S, y_S) in the source and (x_T, y_T) in the target. Let the landmark be translated to a point (x'_S, y'_S) under a given transformation. Then, the TRE for the landmark point pair is given by the Euclidean distance between the translated source location and the target location.

$$\text{TRE} = \sqrt{(x'_S - x_T)^2 + (y'_S - y_T)^2} \quad (28)$$

To measure the statistical significance of improvement in results, we obtained p -values by performing two-sided Wilcoxon tests [48] on DSC and TRE values obtained using the proposed approach paired with corresponding values from each of the other methods. A value of $p < 0.05$ was considered to indicate statistical significance.

B. Dataset 1: Multi-Subject Brain Data

The MR brain data sets were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at the Neuroimaging Informatics Tools and Resources Clearinghouse (<https://www.nitrc.org/projects/ibsr/>). The dataset consists of MR brain images of dimensions $256 \times 256 \times 128$, along with manual expert segmentations of white and grey matter, for 16 subjects. For each subject, the T1-weighted volumetric images have

TABLE I
PERFORMANCE COMPARISON OF SINGLE-CHANNEL REGISTRATION METHODS MULTI-SUBJECT BRAIN MRI DATA,
MEASURED BY DSC OF GREY AND WHITE MATTER BETWEEN SUBJECTS AFTER WARPING

Method	Grey Matter				White Matter			
	Mean	Median	Std Dev	p -value	Mean	Median	Std Dev	p -value
MIND [30]	0.7469	0.7358	0.0276	$7 \cdot 10^{-3}$	0.7213	0.7233	0.0285	$3 \cdot 10^{-4}$
SSC [31]	0.7513	0.7493	0.0263	$3 \cdot 10^{-2}$	0.7203	0.7240	0.0265	$1 \cdot 10^{-4}$
DROP [27]	0.7192	0.7350	0.0808	$1 \cdot 10^{-2}$	0.6589	0.6727	0.0773	$8 \cdot 10^{-5}$
Proposed Method	0.7756	0.7788	0.0244	—	0.7612	0.7645	0.0194	—

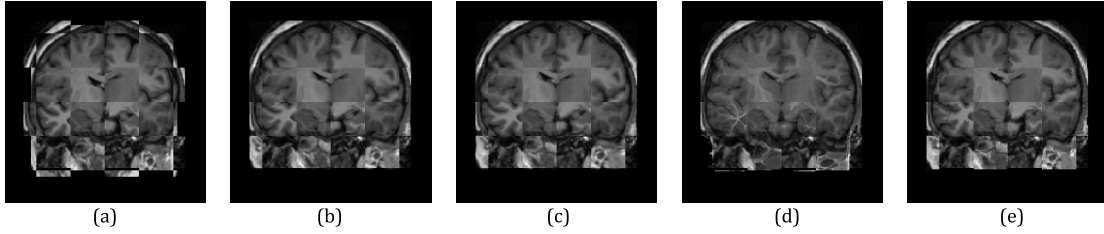


Fig. 5. Visual results on multi-subject MRI data. Results are shown as a checkerboard, where neighboring tiles come from different subjects. (a) Source-target pair before registration (b-e) After registration using (b) MIND (c) SSC (d) DROP and (e) Proposed method.

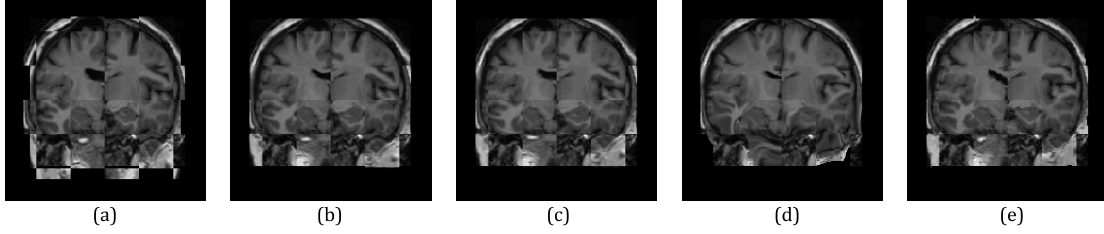


Fig. 6. Visual results on multi-subject MRI data. Results are shown as a checkerboard, where neighboring tiles come from different subjects. (a) Source-target pair before registration (b-e) After registration using (b) MIND (c) SSC (d) DROP and (e) Proposed method.

been positionally normalized into the Talairach orientation (rotation only). We randomly select one subject as the target and register images from each of the other 15 subjects to the selected target image on a slice-wise basis. The SSD-based metric (2) was used as similarity measure, with a window size of 5. We use leave-one-out cross validation for training and testing data per subject.

We compare the performance of the proposed method with that of two self-similarity based registration approaches - MIND [30] and SSC [31] - as well as dense (iconic) registration based on discrete optimization, DROP [27]. For DROP, SSD was used as a measure of similarity since it resulted in the highest DSC values, and the weighting factor λ was empirically found to be 0.01. The optimal regularization term α for MIND and SSC was found to be 0.1. We used three levels of resolution for all methods to compare performance.

The quantitative results on the IBSR dataset are presented in Table I, and visual results are shown in Fig. 5 and Fig. 6. We observe that the proposed approach shows statistically significant improvement over related approaches.

C. Dataset 2: Rabbit Retinal Connectome Data

We test our algorithm on the RC1 connectome, an open access retinal volume described in [32]. The volume contains

images of a 0.25 mm diameter column of the Inner Plexiform Layer (IPL) of rabbit retina. The Viking viewer [49] can be used to view the volume and annotations on it. The entire volume consists of 370 slices of image data, each corresponding to a physical section. 340 slices are captured by an automated electron transmission microscope ATEM and the other 30 slices are captured using Computational Molecular Phenotyping (CMP), a form of light microscopy. Each CMP slice is obtained by probing for either an amino acid (Glutamate, Glutamine, Glycine, Taurine or 4-aminobutyrate (GABA)) or the excitation marker 1-amino-4-guanidobutane(AGB). The resolution of these images is both modalities is 2.18 nm/pixel. The dataset consists of 1132 total cells belonging to 6 major types and various subtypes.

We focus our attention on the top “capstone” section of the volume, which consists of six slices of CMP (each obtained using a unique marker) followed by an ATEM slice. The set of 6 CMP slices together provide “protein signatures” for different cell types (i.e., functional information) which are useful in cell classification. However, cell structure is not clearly discernible in CMP images. ATEM is a complementary source of information which clearly shows the structure of cells including sub-cellular bodies, but does not

TABLE II
PERFORMANCE COMPARISON OF MULTI-CHANNEL REGISTRATION METHODS ON CONNECTOME DATA, MEASURED BY DSC OF BETWEEN ATEM IMAGES AND WARPED CMP IMAGES AND TRE OF LANDMARKS

Method	Dice Similarity Coefficient				Target Registration Error (μm)				Running Time(s)
	Mean	Std Dev	Median	p -value	Mean	Std Dev	Median	p -value	
MIND [30]	0.7661	0.1547	0.7896	$3.9 \cdot 10^{-3}$	1.3532	0.6124	1.3014	$3.7 \cdot 10^{-3}$	463
SSC [31]	0.7682	0.1536	0.7927	$6.4 \cdot 10^{-3}$	1.3112	0.5984	1.2644	$9.0 \cdot 10^{-3}$	404
α -MI [29]	0.7708	0.1436	0.7863	$1.1 \cdot 10^{-3}$	1.3781	0.7581	1.3097	$9.6 \cdot 10^{-3}$	1822
CAMIR [28]	0.7810	0.1495	0.7924	$5.2 \cdot 10^{-3}$	1.3465	0.7401	1.2518	$1.8 \cdot 10^{-2}$	1102
Proposed Method	0.8185	0.1338	0.8345	—	0.9982	0.7144	0.8400	—	598

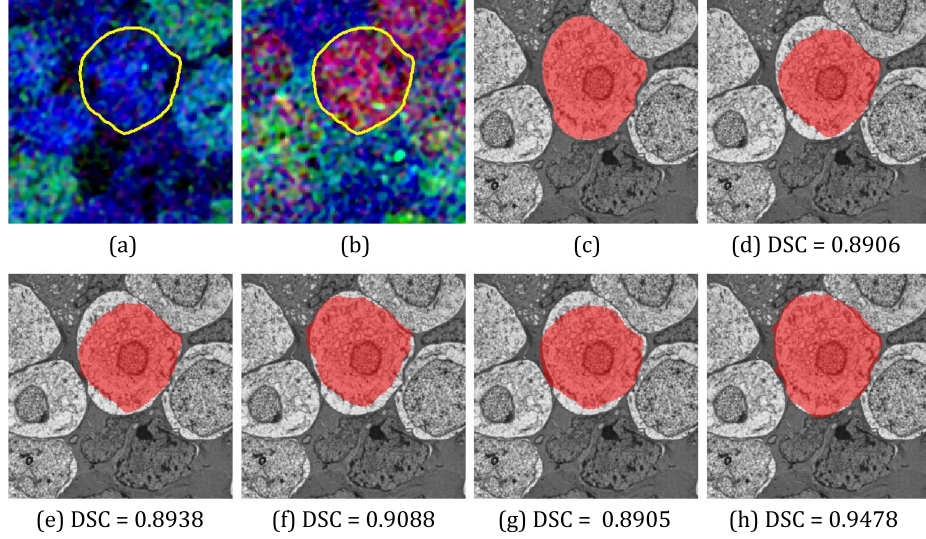


Fig. 7. (Best in color) Visual results on a relatively easy cell. (a) RGB visualization of 3 CMP channels, with ground truth of the cell of interest outlined in yellow (b) RGB visualization of the remaining 3 CMP channels, with outlined ground truth (c) ATEM image corresponding to the same region, with ground truth overlaid in red (d-h) ATEM image with results from various registration methods overlaid in red, along with the corresponding DSC. (d) MIND (e) SSC (f) α -MI (g) CAMIR (h) Proposed method.

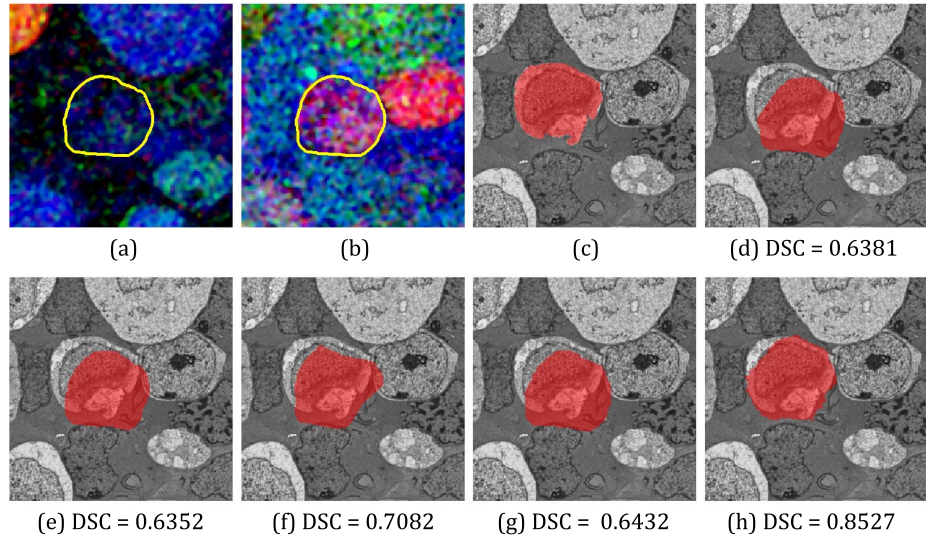


Fig. 8. (Best in color) Visual results on a challenging cell. (a) RGB visualization of 3 CMP channels, with ground truth of the cell of interest outlined in yellow (b) RGB visualization of the remaining 3 CMP channels, with outlined ground truth (c) ATEM image corresponding to the same region, with ground truth overlaid in red (d-h) ATEM image with results from various registration methods overlaid in red, along with the corresponding DSC. (d) MIND (e) SSC (f) α -MI (g) CAMIR (h) Proposed method.

capture functional information well. We evaluate the results against ground truth on a completely annotated region in the capstone section, consisting of 289 manually segmented cells with expert annotations. 40 pairs of expert annotated landmark

locations in CMP and ATEM are also available and used for evaluation of registration efficacy.

For this dataset, the multimodal data similarity measure described in Section III-C was used. CMP was used as the

source and ATEM as the target modality. The feature vector for the target image is the average pixel intensity in a 5×5 neighborhood, and hyper-parameters are set to $M = 7$ and $K = 2$. Training and testing was done using a 4-fold cross-validation strategy. We compare our method with CAMIR [28] embedded in the FFD [19] framework, the α -MI-based approach outlined in [29], MIND [30] and SSC [31]. The optimal value of parameters for the α -MI approach were empirically found to be $\alpha = 0.99$ and $k = 7$ respectively. Three levels of resolution were used for all approaches.

Quantitative results on the RC1 connectome dataset are presented in Table II. Visual results are shown in Figs. 7 and 8. We see the proposed method performs well in both easy as well as challenging scenarios. The most improvement over related methods is observed when there is a large deformation between modalities. We believe this is due to the fact that our approach can account for large deformations during the training phase, by iteratively optimizing emission and transition parameters. A relatively easy scenario is shown in Fig. 7. and an example of a cell with large changes across modalities can be seen in Fig. 8. We observe that the proposed approach shows both quantitative and qualitative improvement of registration accuracy in comparison to competing approaches.

D. Implementation Details

The described method was implemented in C++, with a MATLAB wrapper for ease of interfacing. All experiments were run on a dual-core 3.2 GHz Intel Core i3-550 processor with 8 GB RAM. The OpenMP C++ API was used to implement parallelization with 4 hyperthreads for the proposed approach. Running time for pair of a 1024×1024 pixel images from the connectome dataset are provided in Table II.

V. DISCUSSIONS AND CONCLUSION

We have presented a novel approach for registration of unimodal as well as multimodal image data, with the deformation system embedded in the probabilistic framework of a 2D HMM and solved using the T-HMM approximation. The formulation is general and different types of transformation models may be used. For registration of images with multiple objects, we allow flexibility in the smoothness of the transformation by allowing local adaptation in the transition probability matrix. Multi-channel input data, if available, is utilized in an efficient manner by incorporating it into the emission probabilities of the HMM. Further, we use an efficient approximation to train and decode the T-HMM at reduced complexity. The results of our method show substantial gains over state-of-the-art deformable registration techniques on both intra-modal and inter-modal problems.

We finally note that an important extension of the approach would be to address the domain of 3D-3D registration problems. Such an extension, while non-trivial, appears within reach by leveraging a corresponding extension of the turbo decoding approximation to iterate between three (per-dimension) 1D-HMMs while inducing priors on each other. Nevertheless, while the theory of turbo approximation for 2D-HMM has been described in detail in [3], its formal

extension to higher dimensions, as well as that of the complete approach herein, fall outside the scope of this paper and represent an interesting direction for future work.

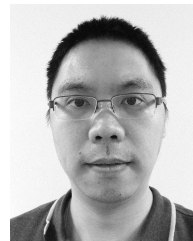
REFERENCES

- [1] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.
- [2] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013.
- [3] F. Perronnin, J.-L. Dugelay, and K. Rose, "Iterative decoding of two-dimensional hidden Markov models," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 3, Apr. 2003, pp. III-329–III-332.
- [4] R. P. Woods, S. R. Cherry, and J. C. Mazziotta, "Rapid automated algorithm for aligning and reslicing PET images," *J. Comput. Asst. Tomogr.*, vol. 16, no. 4, pp. 620–633, 1992.
- [5] P. Thevenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, Jan. 1998.
- [6] K. Ø. Noe, K. Tanderup, J. C. Lindegaard, C. Grau, and T. S. Sørensen, "GPU accelerated viscousfluid deformable registration for radiotherapy," *Stud. Health Tech. Inform.*, vol. 132, no. 16, p. 327–332, 2008.
- [7] M. Foskey *et al.*, "Large deformation three-dimensional image registration in image-guided radiation therapy," *Phys. Med. Biol.*, vol. 50, no. 24, p. 5869, 2005.
- [8] J. R. McClelland *et al.*, "A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy," *Med. Phys.*, vol. 33, no. 9, pp. 3348–3358, 2006.
- [9] R. Shekhar, P. Lei, C. R. Castro-Pareja, W. L. Plishker, and W. D. D'Souza, "Automatic segmentation of phase-correlated CT scans through nonrigid image registration using geometrically regularized free-form deformation," *Med. Phys.*, vol. 34, no. 7, pp. 3054–3066, 2007.
- [10] D. Yang, W. Lu, D. A. Low, J. O. Deasy, A. J. Hope, and I. El Naqa, "4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling," *Med. Phys.*, vol. 35, no. 10, pp. 4577–4590, 2008.
- [11] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm," *J. Comput. Asst. Tomogr.*, vol. 17, no. 4, pp. 536–546, 1993.
- [12] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3D multi-modality medical image registration using feature space clustering," in *Computer Vision, Virtual Reality and Robotics in Medicine*. Berlin, Germany: Springer, 1995, pp. 193–204.
- [13] G. Roche, A. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted Intervention*. Berlin, Germany: Springer, 1998, pp. 1115–1124.
- [14] J. N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image registration by template matching using normalized cross-correlation," in *Proc. IEEE Int. Conf. Adv. Comput. Control Telecommun. Technol.*, Dec. 2009, pp. 819–822.
- [15] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3-D registration of MR and CT images of the head," *Med. Image Anal.*, vol. 1, no. 2, pp. 163–175, 1996.
- [16] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.
- [17] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, Sep. 1997.
- [18] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.
- [19] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
- [20] J. A. Schnabel *et al.*, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Medical Image Computing and Computer-Assisted Intervention*. Berlin, Germany: Springer, 2001, pp. 573–581.
- [21] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 120–128, Jan. 2003.

- [22] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 19–29, Jan. 2010.
- [23] M. Modat *et al.*, "Fast free-form deformation using graphics processing units," *Comput. Methods Prog. Biomed.*, vol. 98, no. 3, pp. 278–284, 2010.
- [24] H. Hirschmüller, "Accurate and efficient stereo processing by semiglobal matching and mutual information," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 807–814.
- [25] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008.
- [26] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through MRFs and efficient linear programming," *Med. Image Anal.*, vol. 12, no. 6, pp. 731–741, 2008.
- [27] B. Glocker, A. Sotiras, N. Komodakis, and N. Paragios, "Deformable medical image registration: Setting the state of the art with discrete methods," *Annu. Rev. Biomed. Eng.*, vol. 13, pp. 219–244, 2011.
- [28] G. K. Rohde, S. Pajevic, C. Pierpaoli, and P. J. Basser, "A comprehensive approach for multi-channel image registration," in *Biomedical Image Registration*. Berlin, Germany: Springer, 2003, pp. 214–223.
- [29] M. Staring, U. A. van der Heide, S. Klein, M. A. Viergever, and P. W. Pluim, "Registration of cervical MRI using multifeature mutual information," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1412–1421, Sep. 2009.
- [30] M. P. Heinrich *et al.*, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image Anal.*, vol. 16, no. 7, pp. 1423–1435, 2012.
- [31] M. P. Heinrich, M. Jenkinson, B. W. Z. Papież, M. Brady, and J. A. Schnabel, "Towards realtime multimodal fusion for image-guided interventions using self-similarities," in *Medical Image Computing and Computer-Assisted Intervention*. Berlin, Germany: Springer, 2013, pp. 187–194.
- [32] J. R. Anderson *et al.*, "Exploring the retinal connectome," *Molecular Vision*, vol. 17, p. 355, Feb. 2011.
- [33] R. Shenoy, M.-C. Shih, and K. Rose, "Hidden Markov model-based multi-modal image fusion with efficient training," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 3582–3586.
- [34] Y. Luo, J. Feng, M. Xu, J. Zhou, J. Min, and G. Xiong, "Registration of coronary arteries in computed tomography angiography images using hidden Markov model," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2015, pp. 1993–1996.
- [35] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [36] J. Li, A. Najmi, and R. M. Gray, "Image classification by a two-dimensional hidden Markov model," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 517–533, Feb. 2000.
- [37] S.-S. Kuo and O. E. Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 842–848, Aug. 1994.
- [38] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.
- [39] F. Perronnin, J.-L. Dugelay, and K. Rose, "A probabilistic model of face mapping with local transformations and its application to person recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 7, pp. 1157–1171, Jul. 2005.
- [40] D. H. T. Vu and R. Haeb-Umbach, "Using the turbo principle for exploiting temporal and spectral correlations in speech presence probability estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2013, pp. 863–867.
- [41] F. Perronnin, J.-L. Dugelay, and K. Rose, "Deformable face mapping for person identification," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 1, Sep. 2003, pp. I-661–I-664.
- [42] J. Weickert and C. Schnörr, "A theoretical framework for convex regularizers in PDE-based computation of image motion," *Int. J. Comput. Vis.*, vol. 45, no. 3, pp. 245–264, 2001.
- [43] A. Wedel, D. Cremers, T. Pock, and H. Bischof, "Structure- and motion-adaptive regularization for high accuracy optic flow," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 1663–1668.
- [44] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statist.*, vol. 41, no. 1, pp. 164–171, 1970.
- [45] L. R. Welch, "Hidden Markov models and the Baum–Welch algorithm," *IEEE Inf. Theory Soc. Newslett.*, vol. 53, no. 4, pp. 1–14, Dec. 2003.
- [46] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B (Methodol.)*, vol. 39, no. 1, pp. 1–38, 1977.
- [47] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [48] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.
- [49] J. R. Anderson *et al.*, "The Viking viewer for connectomics: Scalable multi-user annotation and summarization of large volume data sets," *J. Microscopy*, vol. 241, no. 1, pp. 13–28, 2011.



classification.



segmentation, and object



interested in the relations between information theory, estimation theory, and statistical physics, and their potential impact on fundamental and practical problems in diverse disciplines.

Prof. Rose was a co-recipient of the 1990 William R. Bennett Prize Paper Award of the IEEE Communications Society, and the 2004 and 2007 IEEE Signal Processing Society best paper awards.

Renuka Shenoy (S'14) received the B.E. degree in electronics and communication engineering from Visveswaraya Technological University, India, in 2010, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Santa Barbara, in 2012 and 2016, respectively.

She was a Research Intern with the GE Global Research, Niskayuna, NY, in 2015. Her research interests include biomedical image analysis, image registration, and object segmentation and

Min-Chi Shih (S'05) received the B.S. degree in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2005, the M.S. degree in communication engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2007, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, USA, in 2014.

He is currently with MediaTek USA as an Imaging Scientist. His research interests include image denoising, texture analysis, object tracking, video

Kenneth Rose (S'85–M'91–SM'01–F'03) received the Ph.D. degree from the California Institute of Technology, Pasadena, in 1991.

He was with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, where he is currently a Professor. His main research activities are in the areas of information theory and signal processing, including rate distortion theory, source and source-channel coding, audio-video coding and networking, pattern recognition, and non-convex optimization. He is