

Landmark/Image-based Deformable Registration of Gene Expression Data

Uday Kurkure¹, Yen H. Le¹, Nikos Paragios², James P. Carson³, Tao Ju⁴, and Ioannis A. Kakadiaris¹

¹University of Houston, Houston, TX, USA ²Ecole Centrale Paris, France

³Pacific Northwest National Laboratory, Richland, WA, USA

⁴Washington University, St. Louis, MO, USA

<http://cbl.uh.edu>

Abstract

Analysis of gene expression patterns in brain images obtained from high-throughput *in situ* hybridization requires accurate and consistent annotations of anatomical regions/subregions. Such annotations are obtained by mapping an anatomical atlas onto the gene expression images through intensity- and/or landmark-based registration methods or deformable model-based segmentation methods. Due to the complex appearance of the gene expression images, these approaches require a pre-processing step to determine landmark correspondences in order to incorporate landmark-based geometric constraints. In this paper, we propose a novel method for landmark-constrained, intensity-based registration without determining landmark correspondences a priori. The proposed method performs dense image registration and identifies the landmark correspondences, simultaneously, using a single higher-order Markov Random Field model. In addition, a machine learning technique is used to improve the discriminating properties of local descriptors for landmark matching by projecting them in a Hamming space of lower dimension. We qualitatively show that our method achieves promising results and also compares well, quantitatively, with the expert's annotations, outperforming previous methods.

1. Introduction

With the developments in high-throughput *in situ* hybridization (HITISH) [3], gene expression patterns can be obtained at cellular resolution to explore the functional relationship between various genes and disease mechanisms. The gene expression images are generated with different gene probes that highlight different cells expressing genes at different levels. Determining the correspondence mapping in these images is necessary for any meaningful interpretation of multiple gene expression profiles within the cells. The images can be then organized into a database and queried for similarities in expression patterns to find poten-

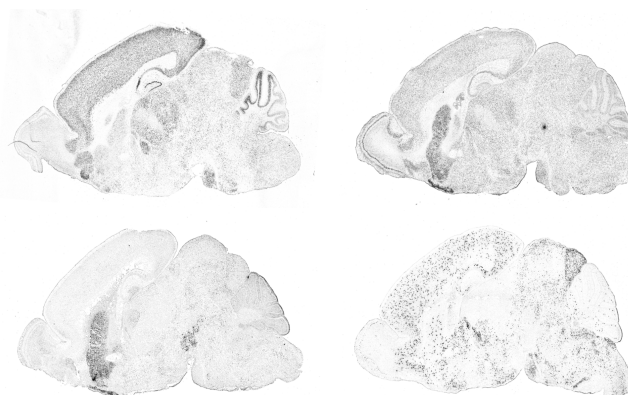


Figure 1. Example images depicting the complex appearance and shape patterns in gene expression images. Each depicted gene expression image is generated with a different gene probe that highlights different cells expressing gene at different levels.

tial interactive relationships between different genes in the same anatomical subregion. However, the expression images exhibit significant variations in appearance and shape, and do not have significant anatomical information (Fig. 1). In addition, the acquisition and sectioning process may introduce multiple artifacts related to smearing and missing parts. Nevertheless, it is required to determine the correspondences of anatomical regions/subregions (Fig. 2) in the expression images to the annotated anatomical atlas to gain knowledge about which genes are expressed with a particular expression pattern [2] in a specific anatomical region/subregion.

Existing methods for gene expression mapping can be broadly classified into approaches based on image registration and deformable model fitting. Due to the complexity of the appearance of the gene expression images, intensity-based registration approaches treat this problem as a multimodal registration problem [9]. However, due to the lack of information in many parts of the images, other approaches have additionally incorporated geometric constraints using signed-distance maps [17] and anatomical landmarks [7, 6].

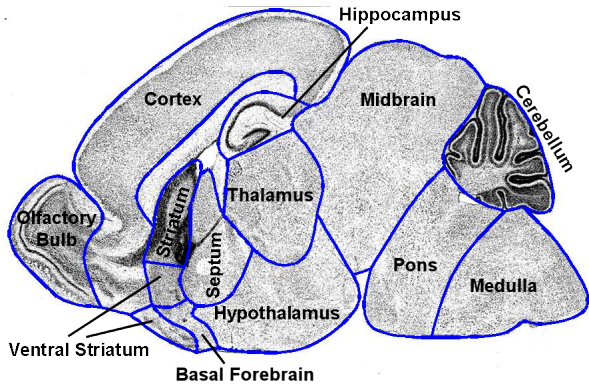


Figure 2. Overlay of the boundary contours depicting the anatomical regions annotated manually on the reference image (from [1]).

Though signed-distance maps are effective in constraining the solution at the boundaries, they lack information of internal anatomical regions. Bello *et al.* [1] proposed an atlas deformation method, where a deformable mesh was fitted using a statistical shape model, anatomical landmarks, and region boundaries in various stages, successively. The corresponding landmarks were detected prior to the fitting process via a classification method by computing features from regions arranged manually for each individual landmark. Landmarks provide anatomy specific constraints and guide the deformation process in regions with uneven information. However, most of the landmark-based approaches require to determine landmark correspondences *a priori* either interactively or by using an automated approach.

In this paper, we present a novel method to determine the correspondences for automated region annotation in gene expression images using a landmark-constrained registration approach based on the Markov Random Field (MRF) framework. The key difference between our approach and previous approaches is that it solves the landmark correspondence and iconic registration problems simultaneously while being rigid transform invariant. Our method does not require to determine the landmark correspondences *a priori*. It only needs to be provided with a few landmark candidates among which there exists at least one desired candidate. Our approach is based on an optimization step where both landmark correspondences and iconic registration are optimized through inter-connected variables. We construct a single graphical model that incorporates intensity-based image registration in one layer and landmark matching in another layer. Both layers are connected through a neighborhood system to impose geometric constraints on each other. The rigid transform invariant formulation is obtained by imposing higher order constraints in the landmark layer through a prior geometric model that is learned from the relative statistics of higher order geometry. The landmark candidates are determined based on their local descriptors' similarity with the landmarks in a reference image. However,

owing to the variability within the gene expression images, it is difficult to define appropriate descriptors. Toward this, Hamming embeddings for the descriptors are learned using similarity sensitive hashing [10] for efficient matching.

Our method partially shares the philosophy in terms of the interaction between the landmark and the deformation space with that of Sotiras *et al.* [13]. However, we incorporate robust higher order constraints through a learned prior model instead of simple pair-wise regularization constraints. Second, our approach is formulated to be translation, rotation, and scale invariant, and therefore, removes the need of global registration, unlike [13]. Finally, landmark candidate selection is context specific and determined through similarity sensitive hashing, which enhances the distinctiveness of the landmark descriptors specific to a particular problem.

The remainder of the paper is structured as follows. In Section 2 we describe the methods for the landmark candidates detection and the construction of the graphical model to solve the correspondence problem. Experiments on the gene expression images and the validation results are presented in Section 3, and Section 4 concludes the paper.

2. Methods

The problem of annotation of anatomical regions and subregions of gene expression images is formulated as a correspondence matching problem with respect to an annotated reference image. In this section, we describe our formulation of Markov Random Field model that simultaneously optimizes for landmark correspondences and iconic registration in a single two-layer graphical model. First, we describe our approach to generate landmark candidates from the gene expression images. Then, we present the details of the construction of each layer of the model and their inter-connections.

2.1. Landmarks and Descriptors

We define landmarks as points in the gene expression images that are locally salient in terms of appearance or shape characteristics irrespective of the gene being expressed. Two types of landmarks were chosen through visual inspection [1]: appearance-based landmarks and boundary-based landmarks. Figure 3 depicts the chosen landmarks on a Nissl-stained reference image.

Appearance-based landmarks: We use local image descriptors computed from gradient orientation histograms to represent the appearance-based landmark features [14]. In this representation, commonly known as DAISY, the polar Gaussian pooling approach is used to construct the histograms. The measure of (dis)similarity between any two points is defined as the distance between the distributions of the oriented gradients in their descriptors. It has been shown

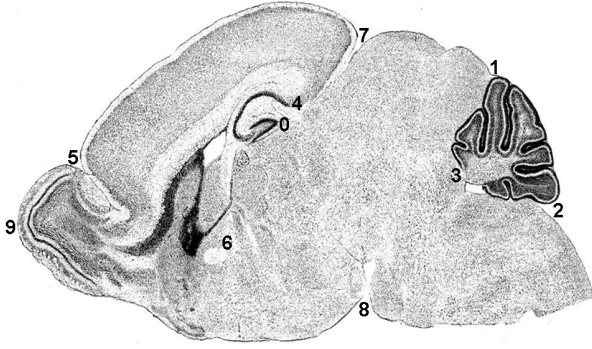


Figure 3. Manually selected landmarks in tissue sections that are locally salient in terms of appearance or shape characteristics. The numbers represent the IDs assigned to each landmark location (from [1]).

to outperform other discriminative local image descriptors [16]. In addition, it can be computed very efficiently.

Figure 4(a) depicts sample patches from the gene expression images for landmark ‘0’, whereas Fig. 4(b) depicts the distance map computed with reference to the landmark point in the reference image. The best candidate for this landmark can be selected as the point that corresponds to the minimum distance in the distance map. However, it can be observed from Fig. 4(b) that it may result in false matches, and may not identify the real landmark point in the image because of complex appearance information, which varies from image to image.

We employ a machine learning approach to construct efficient similarity measures to reduce both types of errors to enhance the (dis)similarity of the descriptor between (dis)similar points. An ensemble classifier based on boosting technique is trained on *pairs* of point descriptors instead of single point descriptors. The *paired* learning is performed to select similarity-relevant features, and to binarize them such that the Hamming distance between the descriptors of a pair of points is small, if they correspond to the same landmark.

Specifically, we randomly selected eight gene expression images in addition to the reference image for learning the binary embedding model. We constructed sets of positive pairs and negative pairs for each landmark from a local patch centered at the landmark points. We followed a boosting-based approach [10] to learn the transformation of a high dimensional feature space into a reduced space of binary features. The n -dimensional binary Hamming embedding is represented as $\xi(x) = (\xi_1(x), \dots, \xi_n(x))$, where each dimension is computed by a binary function parameterized by a projection function $\varphi : X \rightarrow \mathbb{R}$, or

$$\xi_i(x) = \begin{cases} 0, & \text{if } \varphi_i(x) \leq 0 \\ 1, & \text{otherwise} \end{cases}.$$

Each such function defines a weak binary classifier on pairs

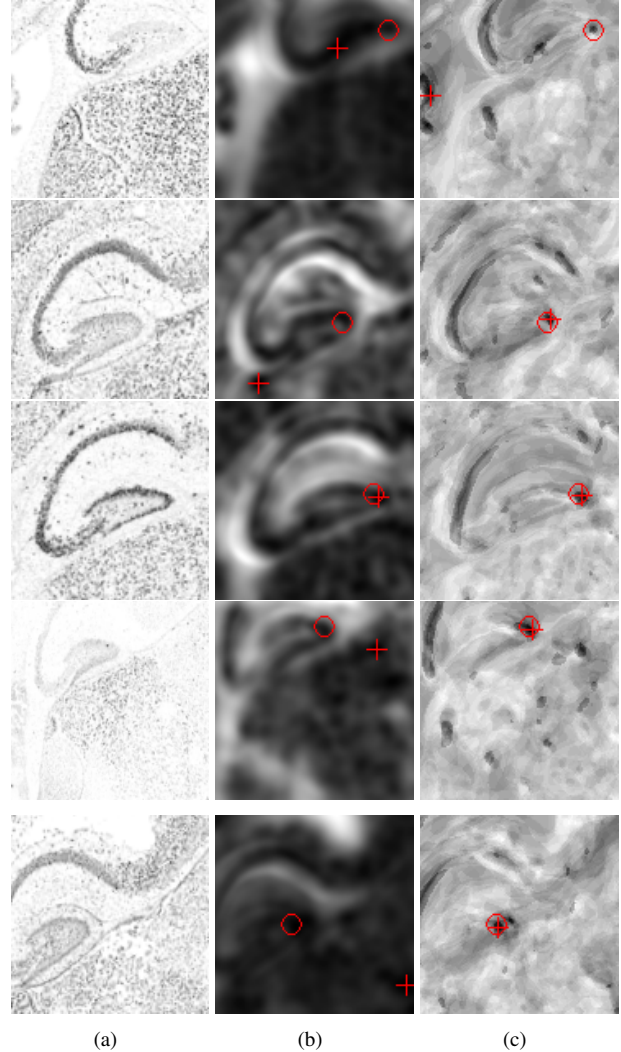


Figure 4. Depiction of a single candidate selection for landmark ‘0’ in a search window: (a) gene expression image (left column), (b) distance maps using daisy descriptor (middle column), and (c) distance maps using binary embedding (right column). The distance maps are computed by comparing features of each pixel in the search window with the features of the landmark in the reference image. The symbol ‘o’ depicts the expected location and ‘+’ depicts the obtained location as the minima of the distance map.

of points,

$$h_i(x, y) = \begin{cases} +1, & \text{if } \xi_i(x) = \xi_i(y) \\ -1, & \text{otherwise} \end{cases},$$

or simply,

$$h_i(x, y) = \text{sign}(\varphi_i(x)) \cdot \text{sign}(\varphi_i(y)).$$

Thus, the Hamming metric between the embeddings $\xi(x)$ and $\xi(y)$ of a pair of points x and y can be computed as

weighted summation of output of the binary classifiers,

$$d(\xi(x), \xi(y)) = \frac{1}{2} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \alpha_i h_i(x, y), \quad (1)$$

where $\alpha_i > 0$ is the weight for the i^{th} dimension. In each iteration of the boosting, we select a Hamming embedding ξ_i for which the binary weak classifier h_i maximizes the weighted correlation,

$$r_i = \sum w_{ik} s_k h_i(x_k, y_k),$$

of labels s_k for all pairs. Note that the Hamming embedding ξ_i is a function of the projection map φ_i which is represented by a single feature from the descriptors and a threshold a_i for that feature. Figure 4(c) depicts the Hamming distance map computed using Eq. (1) with reference to the landmark ‘0’ in the reference image (the expected location and the minimum of the distance map are depicted as ‘o’ and ‘+’, respectively). It can be observed that for the depicted images, the candidates selected using the Hamming embeddings are closer to the expected locations than the candidates selected using the original features. Additional candidates can be selected for each landmark by finding additional local minima points in the Hamming distance maps.

Boundary-based landmarks: The curvature value of the boundary contour at the landmark points is used to represent the boundary-based landmarks. First, the gene expression image is segmented by applying histogram thresholding, flood-filling and morphological operations. Then, the boundary of the segmented brain image is obtained using a border tracing algorithm [12]. The resulting boundary is further smoothed with a moving average filter. The curvature κ_i at each point on the smoothed boundary is computed as:

$$\kappa_i = \frac{\dot{x}_i \ddot{y}_i - \dot{y}_i \ddot{x}_i}{(\dot{x}_i^2 + \dot{y}_i^2)^{3/2}},$$

where, (\dot{x}_i, \dot{y}_i) and (\ddot{x}_i, \ddot{y}_i) are the first order and second order derivatives, respectively. The candidate points for a boundary landmark in a given image are determined as the points with maximum convex or concave curvature (depending on the landmark curvature type in the reference image) in the local search window.

2.2. Iconic Registration

Consider an image $I : \Omega \rightarrow R$ to be registered to another image J . Using energy minimization principles, the spatial correspondences between the two images can be obtained by recovering an optimal transformation $T(\mathbf{x})$ from:

$$E_1(T) = \int_{\Omega} \psi_1(J(\mathbf{x}), I(T(\mathbf{x}))) d\mathbf{x},$$

where $\psi_1(\cdot)$ is a similarity function that defines the relationship between the intensity patterns in the two images. For

nonlinear registration, the transformation $T(\mathbf{x})$ is defined in terms of the deformation field $D(\mathbf{x})$, or $T(\mathbf{x}) = \mathbf{x} + D(\mathbf{x})$. To impose smoothness on the deformation field, a regularization function is included in the energy function as:

$$E_2(T) = \int_{\Omega} \psi_2(\nabla T(\mathbf{x})) d\mathbf{x},$$

where $\psi_2(\cdot)$ is a smoothness function. Considering a deformation grid of control points, \mathcal{C} , super-imposed on the image, the deformation field, $D(\mathbf{x})$, at any point in the image can be interpolated from the deformation vectors of the control points:

$$D(\mathbf{x}) = \sum_{\mathbf{c} \in \mathcal{C}} \eta(|\mathbf{x} - \mathbf{c}|) \mathbf{d}_{\mathbf{c}},$$

where $\eta(\cdot)$ is a weighting function describing the contribution of the control point \mathbf{c} at any point \mathbf{x} in the image and $\mathbf{d}_{\mathbf{c}}$ is the displacement vector of the control point \mathbf{c} . The appearance-based energy function thus can be redefined as:

$$E_1(T) = \sum_{\mathbf{c} \in \mathcal{C}} \int_{\Omega} \hat{\eta}(\mathbf{x} - \mathbf{c}) \psi_1(J(\mathbf{x}), I(T(\mathbf{x}))) d\mathbf{x},$$

where $\hat{\eta}(\cdot)$ determines the influence of a point \mathbf{x} on the control point \mathbf{c} .

Next, we describe the construction of the image registration-based graph layer for the MRF model following the formulation of Glocker *et al.* [4]. Consider a graph, G^a (here a refers to the registration layer), whose nodes are the M^a control points \mathcal{C} from the registration grid. Let $L^a = \{l_1^a, \dots, l_{H^a}^a\}$ be a discrete set of H^a labels for that layer. These labels correspond to a quantized deformation space $D = \{\mathbf{d}_1, \dots, \mathbf{d}_{H^a}\}$. For a particular node i , a label assignment l_k^a corresponds to the displacement of the node by $\mathbf{d}_{l_k^a}$. Thus, the goal of the registration is to find a mapping $F^a : G^a \rightarrow L^a$ that is optimal given some criterion to recover the transformation $T(\mathbf{x})$ defined in terms of the deformation field $U(\mathbf{x})$, or $T(\mathbf{x}) = \mathbf{x} + U(\mathbf{x})$. The MRF provides an elegant and efficient mathematical framework for solving such discrete labeling problems. Any possible assignment $f^a = \{f_1^a, \dots, f_{M^a}^a\}$ of labels to the random variables is called a configuration of F^a , and is essentially a realization of the field. Note that every configuration f^a defines a labeling and \mathbb{F}^a denotes the set of all possible configurations. We also define a neighborhood system $N = \{N_i \mid \forall i \in \mathcal{C}\}$ for the set of control point nodes \mathcal{C} , where N_i is the set of all neighbors of the node $i \in \mathcal{C}$. The MRF energy function for image registration is defined as:

$$E^a(f^a) = \sum_{i \in \mathcal{C}} V_i^a(f_i^a) + \sum_{i \in \mathcal{C}} \sum_{j \in N_i} V_{ij}^a(f_i^a, f_j^a),$$

where $V_i^a(\cdot)$ and $V_{ij}^a(\cdot, \cdot)$ are the first- and second-order clique potential functions representing the data and the

smoothness terms, respectively. The unary or the first-order potentials are defined as:

$$V_i^a(f_i^a) = \int_{\Omega} \hat{\eta}(\mathbf{x} - \mathbf{c}) \psi_1(J(\mathbf{x}), I(\mathbf{x} + U^{t-1}(\mathbf{x}) + \mathbf{d}_{f_i^a})) d\mathbf{x},$$

where $\psi_1(\cdot)$ is an image similarity measure. Though the gene expression images belong to the same modality, owing to the variations in the observed information caused by the different types of probing, we can consider that they are acquired from different modalities. Such a consideration leads us to use the popular multi-modality similarity function defined in terms of normalized mutual information [8, 15]. In this work, we computed the normalized mutual information using a simple histogram-based method with 128 bins. The regularization function for smoothness in the label domain is defined as a function of distance between the deformation vectors of neighboring control points, or

$$V_{ij}^a(f_i^a, f_j^a) = \lambda^a \exp\left(\frac{-\|\mathbf{d}_{f_i^a} - \mathbf{d}_{f_j^a}\|^2}{2(\sigma_a)^2}\right),$$

where λ^a is a weighting parameter balancing the effects of the appearance-based similarity and the smoothness of the deformation field.

2.3. Landmark Correspondences

Landmarks are generally used in the registration process to constrain the solution at specific known locations in the image. The landmark-based constraints are imposed on the deformation field by adding an additional energy term corresponding to the sum of distances between the corresponding landmarks. Note that in such methods the location of the landmarks in the reference image and the given image are assumed to be known *a priori* along with their correspondences. If the landmark correspondences are not known *a priori*, a pre-processing step is performed to select the best landmark candidate in the given image corresponding to the landmark in the reference image.

The problem of selecting a single best candidate for each landmark can be formulated as a discrete labeling problem. Thus, for the landmark-based graph layer, consider a graph G^b (b refers to the landmark layer) whose nodes are the M^b landmark points $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_{M^b}\}$ in the reference image, and the discrete labels L^b correspond to the H^b candidate points $\{\mathbf{p}_1^i, \dots, \mathbf{p}_{H^b}^i\}$ in the given image for each landmark \mathbf{q}_i . A label assignment $f_i^b = l_k^b$ to a landmark \mathbf{q}_i corresponds to the selection of the candidate \mathbf{p}_k^i as the best candidate for that landmark. The MRF energy function for landmark matching is defined as:

$$E^b(f^b) = \sum_{i \in \mathcal{Q}} V_i^b(f_i^b) + \sum_{c \in \mathcal{C}} V_c^b(\mathbf{f}_c^b),$$

where $V_i^b(\cdot)$ and $V_c^b(\cdot)$ are the unary and higher order potential functions, respectively, for the landmark graph layer

and \mathbf{f}_c^b is the configuration of the triplets forming a clique c from a set of cliques \mathcal{C} .

The unary potential for the landmarks is defined by $\psi_3(\cdot)$ which measures the candidacy strengths of the landmark candidates based on the Hamming metric or the curvature depending on the type of the landmark:

$$V_i^b(f_i^b) = \psi_3\left(\mathbf{p}_{f_i^b}^i\right).$$

The higher order potential encodes the penalty function of assigning a triplet of labels to a triplet of connected nodes. The interdependencies between the locations of the various landmark triplets can be modeled using the probability distribution of the relative lengths of triangles formed between the landmarks. Specifically, given a triplet of landmark points i, j , and k belonging to a clique c , the shape of the triangle formed by them can be defined by the relative lengths of any two sides of the triangle. The relative length for a particular side is defined by the ratio of the length of that side over the total length of all three sides. Thus, we represent a triplet forming the clique c by a two element descriptor $\mathbf{r}_c = (r_{ij}, r_{jk})$, where r_{st} is the relative length of the side defined by the points s and t .

Using such representation, we can capture spatial variability that is translation, rotation, and scale invariant by learning the distributions for each \mathbf{r}_c from a set of training images. Thus, we learn the distribution, $\rho_c(\mathbf{r}_c)$ using a multivariate Gaussian distribution, $\mathcal{N}(\mathbf{r}_c | \mu_c, \Sigma_c)$, where μ_c and Σ_c are the mean and the covariance matrix learned from the training set, respectively. This prior captures the local relationship between the landmark locations and constrains the space in which a triplet of landmarks can co-exist. This is especially advantageous when the strength of landmark candidacy is not strong enough. Thus, the higher order clique potential is defined to incorporate this prior knowledge to impose a global prior cost on the locations of the landmark candidates as:

$$V_c^b(\mathbf{f}_c^b) = \lambda^b (1 - \rho_c(\mathbf{r}_c(\mathbf{f}_c^b))),$$

where λ^b is a positive weight, $\mathbf{r}_c(\mathbf{f}_c^b)$ maps the locations of the triplet c to the two element descriptor formed from relative lengths of the sides.

2.4. Combining Image Registration and Landmark Correspondences

In this section, we describe how to connect the two graph layers such that the control points and the landmark points influence each other to obtain optimal labeling configuration. The MRF framework allows us to add an additional layer to the graph. For a given landmark point \mathbf{q}_i in the reference image, selection of a particular candidate point \mathbf{p}_k^i determines the displacement of that point in the given image. In other words, the point defined as a landmark

in the reference image is then known to be displaced by $\mathbf{u}_{k,i} = \mathbf{p}_k^i - \mathbf{q}_i$ in the given image. Thus, the control points that are in the vicinity of the landmark point should be assigned labels that correspond to the displacement that is most similar to $\mathbf{u}_{k,i}$ in terms of magnitude and direction. Similarly, the assignments of the displacement labels to the control points in the vicinity of the landmark points should also influence the selection of the landmark candidate based on its displacement vector. Such interactions between the landmark points and the control points can be defined by a pair-wise potential function which is a function of the distance between the displacements of the control points and the landmark points. Thus, we define the pairwise potential function between a landmark point \mathbf{q}_i and a control point \mathbf{c}_j as:

$$V_{ij}^{ab}(f_i^b, f_j^a) = \lambda^{ab} \phi(\mathbf{q}_i, \mathbf{c}_j) \exp\left(\frac{-\|\mathbf{u}_{k,i} - \mathbf{d}_{f_j^a}\|^2}{2(\sigma_{ab})^2}\right),$$

where,

$$\phi(\mathbf{q}_i, \mathbf{c}_j) = 1 - \exp\left(\frac{-\|\mathbf{q}_i - \mathbf{c}_j\|^2}{2(\sigma_r)^2}\right),$$

is a weighing function based on the distance between the landmark point and the control point. The weighting function $\phi(\cdot)$ can also be used to determine the edges between the two layers of the graph. We define a local neighborhood system for the landmarks with respect to the control points to impose geometrical constraints in order to preserve the local deformations. Specifically, a landmark point \mathbf{q}_i is paired with all the control points for which $N^{ab}(\mathbf{q}_i) = \{\mathbf{c}_j | \phi(\mathbf{q}_i, \mathbf{c}_j) < \epsilon\}$, where ϵ is a threshold. Based on the threshold ϵ , we can increase or decrease the number of edges between the two layers of the graph. Moreover, such neighborhood system avoids undesired regularization, especially at the lower grid resolution.

2.5. MRF Optimization

Finally, we need to define a way to optimize the proposed higher-order MRF energy function. Toward this, we adopt the general framework proposed by Ishikawa [5]. In this method, the optimization is performed by: (i) transforming the multi-label MRF problem to a binary-MRF problem, and (ii) transforming the higher-order MRF to a first-order MRF. The optimization of the binary MRF is performed using a well-known quadratic pseudo-boolean optimizer (QPBO) [11] implemented by Vladimir Kolmogorov.

3. Results and Discussion

We evaluated our method on the 2D gene expression images that were acquired as part of gene expression study [2] and they were provided by Bello *et al.* [1]. These images are

sagittal sections of postnatal day 7 mouse brains at standard section 9, each revealing the expression of a single gene after *in situ* hybridization. The images were acquired using a light microscope at $3.3 \mu\text{m}$ per pixel resolution resulting in approximately 2400×4000 pixels image size. The images were scaled down to 25% of the original size for computation purposes.

To capture the local deformations, the control point grid resolution was successively increased at each iteration. At each grid resolution, multiple optimization cycles were used with successively decreasing maximum displacement range for each control point. The maximum displacement range was sampled to provide a total of 25 labels or possible displacements for each control point. The landmark candidate points in a given test image were selected by dividing the local search window into 3×3 subregions and picking the two best candidates from each subregion that have the minimum Hamming distance with respect to the landmark point in the reference image.

To demonstrate the advantage of the dual graph layers for registration, we compare, qualitatively and quantitatively, the proposed method with iconic registration performed without the landmark layer. Figure 5 depicts selected qualitative results for both methods using checkerboard visualization of the registered gene expression images.

Table 1 depicts the mean distance error in pixels for each landmark averaged over two cohorts of images - all the 100 images (Cohort A) and a subset of 53 images that excludes training images and images with severe boundary deformity or tear (Cohort B). Note that the proposed method performs comparable to or better than the iconic registration method. The errors for the boundary landmarks, especially landmarks 1 and 2, are higher than those for the appearance-based internal landmarks. This can be attributed to the high variability in the local shape and position of the anatomical structures in the mouse brain.

We also compare the performance of the proposed method with the results provided by Bello *et al.* [1] in terms of landmark distance errors in pixels. Figure 6 depicts the scatter plots of distance errors for selected landmarks comparing the two methods for Cohort B gene expression images. In the scatter plots, the data points below the diagonal represent the images where the proposed method outperforms the previous method. For few boundary landmarks, our method has slightly higher error in individual images but the overall mean error is still significantly lower than the overall mean error of the previous method. Since the boundary landmarks are described by a weak curvature-based feature, there is possibility of unreliable candidates which may produce worse results. Thus the proposed method has a landmark bias which can be put to good use by improving the candidate detection (e.g., by fusing the shape and appearance information for describing the boundary land-

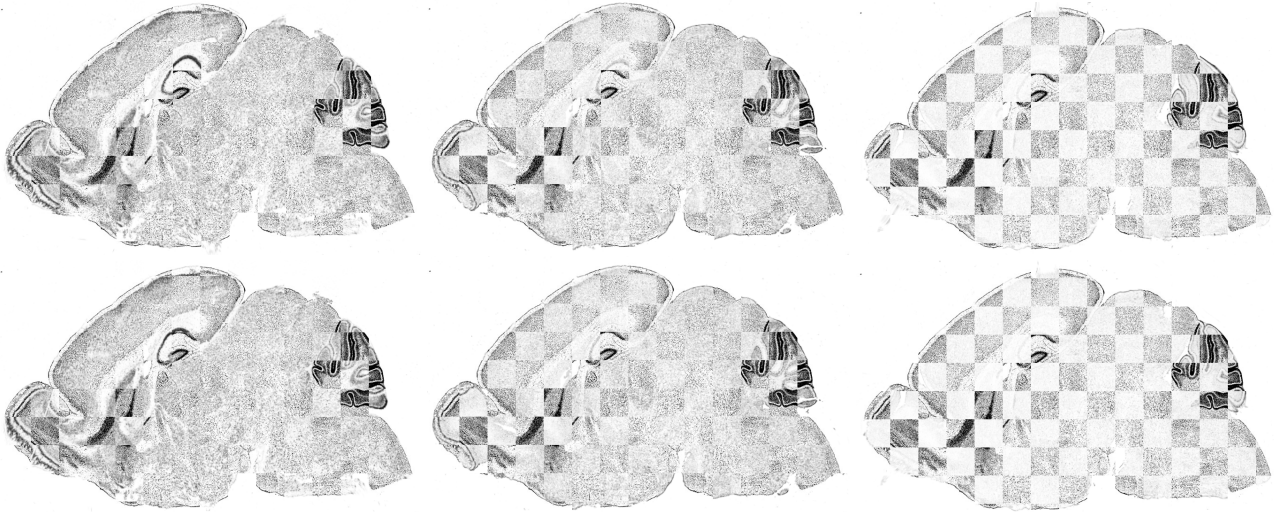


Figure 5. Checkerboard visualization of the registered images obtained from the image registration without (top) and with (bottom) the landmark graph layer on selected gene expression images.

ID	REG		LM+REG	
	A	B	A	B
0	9.49	9.60	5.33	5.84
1	20.73	7.68	16.04	7.00
2	17.29	15.04	13.70	10.98
3	10.92	6.83	9.29	6.18
4	9.44	9.35	6.25	6.48
5	9.19	7.38	9.52	7.63
6	12.17	11.11	9.78	8.62
7	8.80	9.71	5.99	6.34
8	9.42	6.63	8.29	5.34
9	5.48	6.19	5.51	6.19

Table 1. Average distance error (in pixels) for each landmark location compared to the manual annotations. The corresponding locations of the landmarks were obtained using only the registration layer (REG) and both registration and landmark layers in tandem (LM+REG). Columns A and B refer to results from cohorts A and B, respectively.

marks). Incorporating measures of confidence on reliability of the candidates and suitably modifying the graph structure to accommodate missing landmarks can alleviate the problem of missing correspondences.

4. Conclusion

In this paper, we have presented a novel method to obtain correspondences between gene expression images using a landmark-constrained image registration method. We formulate the registration and landmark matching problem in a single MRF model as a discrete labeling problem. Our method does not assume the landmark correspondences to be known prior to registration. The geometric relationships

of the landmarks are coded as higher-order spatial priors imposing translation, rotation, and scale-invariant constraints that are learned from few training images. Furthermore, a landmark specific similarity model is learnt using a boosting approach enhancing the descriptors' discriminative properties for landmark candidates selection. Finally, our method achieves lower errors for correspondence mapping as compared to other methods on a challenging dataset of gene expression images.

References

- [1] M. Bello, T. Ju, J. P. Carson, J. Warren, W. Chiu, and I. A. Kakadiaris. Learning-based segmentation framework for tissue images containing gene expression data. *IEEE Transactions on Medical Imaging*, 26:728–744, 2007. 1090, 1091, 1094, 1096
- [2] J. P. Carson, T. Ju, M. Bello, C. Thaller, J. Warren, I. A. Kakadiaris, W. Chiu, and G. Eichele. Automated pipeline for atlas-based annotation of gene expression patterns: Application to postnatal day 7 mouse brain. *Methods*, 50:85–95, Aug. 2010. 1089, 1094
- [3] J. P. Carson, C. Thaller, and G. Eichele. A transcriptome atlas of the mouse brain at cellular resolution. *Current Opinion in Neurobiology*, 12(5):562–565, Oct. 2002. 1089
- [4] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios. Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis*, 12(6):731–741, Dec. 2008. 1092
- [5] H. Ishikawa. Transformation of general binary MRF minimization to the first order case. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Mar. 2010. 1094
- [6] M. Jagalur, C. Pal, E. Learned-Miller, R. Zoeller, and D. Kulp. Analyzing in situ gene expression in the mouse brain with image registration, feature extraction and block clustering. *BMC Bioinformatics*, 21:1–21, 2007. 1089

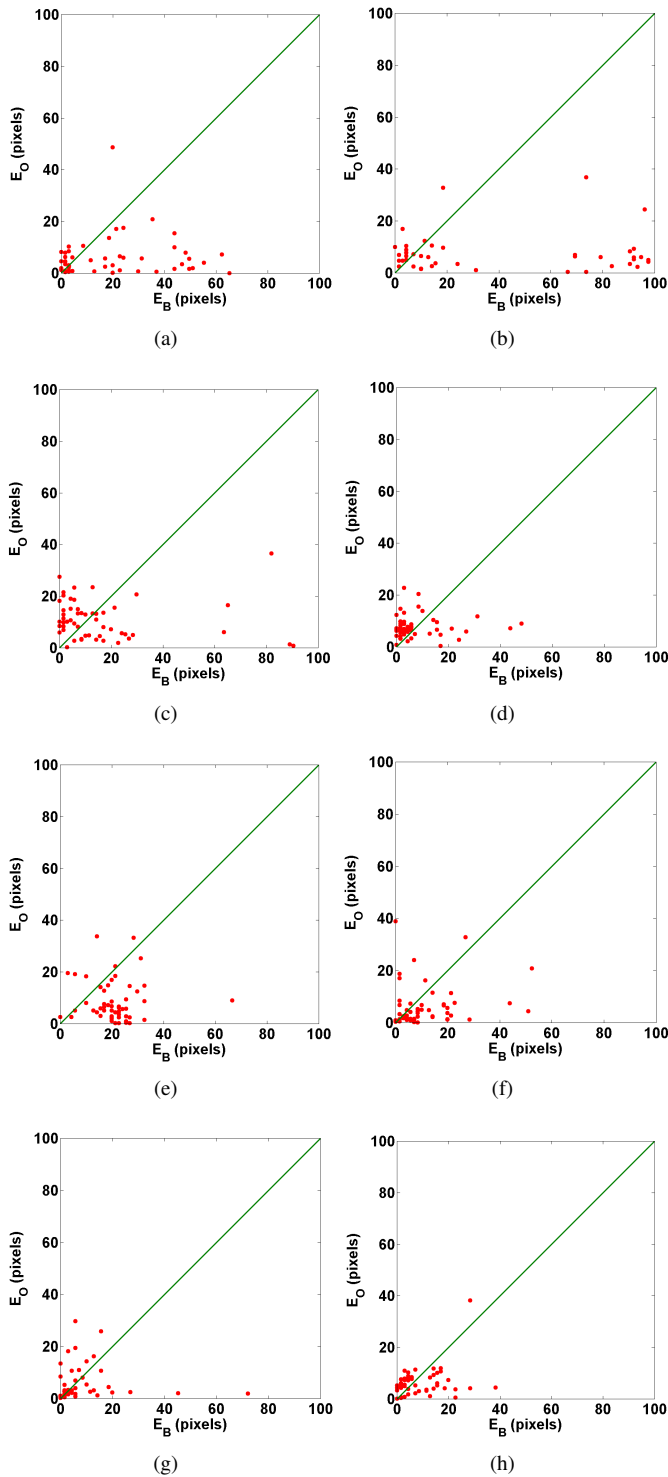


Figure 6. Comparison of the distance errors in pixels between our method (vertical axis) and [1] (horizontal axis) for: (a) landmark 0, (b) landmark 1, (c) landmark 2, (d) landmark 5, (e) landmark 6, (f) landmark 7, (g) landmark 8, and (h) landmark 9.

- [7] T. Lin, C. L. Guyader, E.-F. Lee, I. D. Dinov, P. M. Thompson, A. W. Toga, and L. A. Vese. Gene to mouse atlas registration using a landmark-based nonlinear elasticity smoother. In *Proc. SPIE Medical Imaging: Image Processing*, pages 1–16, Lake Buena Vista, FL, Feb. 7–12 2009. 1089
- [8] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, April 1997. 1093
- [9] L. Ng, S. D. Pathak, C. Kuan, C. Lau, H. Dong, A. Sodt, C. Dang, B. Avants, P. Yushkevich, J. C. Gee, D. Haynor, E. Lein, A. Jones, and M. Hawrylycz. Neuroinformatics for genome-wide 3D gene expression mapping in the mouse Brain. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4(3):382–393, 2007. 1089
- [10] L. Ren, G. Shakhnarovich, J. Hodgins, H. Pfister, and P. Viola. Learning silhouette features for control of human motion. *ACM Transactions on Graphics*, 24(4):1303–1331, 2005. 1090, 1091
- [11] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. Optimizing binary MRFs via extended roof duality. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Minneapolis, MN, June 17–22 2007. 1094
- [12] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis and machine vision*. Brooks/Cole Publishing Company, Pacific Grove, USA, 2nd edition, 1999. 1092
- [13] A. Sotiras, Y. Ou, B. Glocker, C. Davatzikos, and N. Paragios. Simultaneous geometric-ionic registration. In *Proc. Medical Image Computing and Computer-Assisted Intervention*, pages 676–683, Beijing, China, 2010. 1090
- [14] E. Tola, V. Lepetit, and P. Fua. DAISY: An efficient dense descriptor applied to wide baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):815–830, May 2010. 1090
- [15] P. A. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, Sep. 1997. 1093
- [16] S. Winder and M. Brown. Learning local image descriptors. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Minneapolis, MN, Jun. 18–23 2007. 1091
- [17] Z. Yi and S. Soatto. Correspondence transfer for the registration of multimodal images. In *Proc. IEEE International Conference on Computer Vision*, pages 1–8, Rio de Janeiro, Brazil, Oct. 14 - 21 2007. 1089