# Global Image Registration Based on Learning the Prior Appearance Model

Ayman El-Baz
Bioimaging Lab., Dept. of Bioengineering,
University of Louisville,
Louisville, KY, 40292, USA.
aselba01@louisville.edu

Georgy Gimel'farb
Dept. of Computer Science, Tamaki Campus,
The University of Auckland,
Auckland, New Zealand.
g.gimelfarb@auckland.ac.nz

## Abstract

*A new approach to align an image of a textured object with a given prototype (learned reference object) is proposed. Visual appearance of the images, after equalizing their signals, is modeled with a Markov-Gibbs random field with pairwise interaction. Similarity to the prototype (learned reference object) is measured by a Gibbs energy of signal co-occurrences in a characteristic subset of pixel pairs derived automatically from the prototype. An object is aligned by an affine transformation maximizing the similarity by using an automatic initialization followed by gradient search. To get accurate appearance model, we developed a new approach to automatically select the most important cliques (neighborhood system) that describe the visual appearance of a texture object. Experiments confirm that our approach aligns complex objects better than popular conventional algorithms.*

## 1. Introduction

Image registration aligns two or more images of similar objects taken at different times, from different viewpoints, and/or by different sensors. The images are geometrically transformed to ensure their close similarity. Registration is a crucial step in many applied image analysis tasks, e.g. to fuse various data sources (such as computer tomography (CT) and MRI data in medical imaging) for image fusion, change detection, or multichannel image restoration; form and classify multi-band images in remote sensing; update maps in cartography, perform automatic quality control in industrial vision, and so forth. Co-registered medical images provide more complete information about the patient, help to monitor tumor growth and verify treatment, and allow for comparing the patient's data to anatomical atlases. For the state of the art of registration methods we refer the reader to [1].

Most of the known registration methods fall into two main categories: feature-based and area-based tech-

niques [1]. Feature based techniques rely on salient local structures extracted from images, e.g. specific areas such as water reservoirs and lakes [2, 3], buildings [4], or urban areas [5], specific lines like straight segments [6, 7, 8], object contours [9, 10, 11], coast lines [12, 13], rivers, or roads [14, 15], and specific points, e.g. road crossings [16], centroids of water areas, or oil and gas pads [17]. Scale invariant feature transform (SIFT) proposed by Lowe [18] is most popular at present because it reliably determines a number of point-wise correspondences between two images differing by affine transformation and local contrast / offset signal deviations. But these methods can be used only if objects have distinctive and non-repetitive local features.

Area-based methods such as the classical least square correlation match directly image signals to avoid feature extraction [19]. However, because the objects are assumed identical to within spatially uniform signal deviations, the correlation is too sensitive to non-uniform and spatially interdependent deviations of the corresponding signals due to sensor noise, illumination variations, and/or different sensor types. Alternative phase correlation and spectral-domain (e.g. Fourier-Mellin transform based) methods [20, 21, 22, 23, 24] are more robust with respect to the correlated and frequency dependent noise and non-uniform time varying illumination. However, these methods typically allow for only very limited geometric transformations.

More powerful mutual information (MI) based image registration [25, 26] exploits a probabilistic similarity measure that allows for more general types of signal deviations than correlation. The statistical dependency between two data sets is measured by comparing a joint empirical distribution of the corresponding signals in the two images to the joint distribution of the independent signals. Because the MI-based registration performs the best with multi-modal images [26], it is used in many of medical imaging applications. The joint distribution is estimated using Parzen windows [27] or discrete histograms [28]. The main advantage of the MI is insensitivity to monotone variations of correspondence between the object and prototype signals,

but the objects should be of almost identical shape apart from their affine geometrical and monotone signal transformations. The MI allows also for some non-monotone signal correspondence variations although they may change the visual appearance too much and hinder registration accuracy.

Recently, the researchers used Markov-Gibbs Random Field (MGRF) for modeling image deformation, e.g. Roy and Govindu [29], Boykov et al. [30], Kumar et al. [31], Winn and Jojic [32], Kovtun [33], and Shekhovtsov et al. [34], **and, to the best of our knowledge, we are the first authors who propose to use MGRF as a new similarity measure for global image alignment.**

We consider a more general case of registering a textured object to a prototype with similar but not necessarily identical visual appearance under their relative 2D affine transformations and monotone variations of signal correspondences. The variations are suppressed by equalizing signals in the images. The co-registered equalized images are described with a characteristic subset of signal co-occurrence statistics. The description implicitly "homogenizes" the images, i.e. considers them as spatially homogeneous patterns with the same statistics. In contrast to the feature-based registration, the statistics characterize the whole object. In contrast to the conventional area-based techniques, similarities between the statistics rather than pixel-to-pixel correspondences are measured. Section 2 represents the equalized object and prototype images as samples of a generic Markov-Gibbs random field (MGRF) with pairwise pixel interaction. Gibbs potentials are analytically estimated from co-occurrence statistics for the prototype. Similarity between an affinely transformed object and the prototype is measured with a total Gibbs energy for a characteristic pixel neighborhood. A new algorithm for selecting the neighborhood for the MGRF model is introduced. After an automatic initialization, the affine transformation aligning the object with the prototype is found by the gradient search for the maximum Gibbs energy of the transformed object. Experiments in Section 3 confirm that our method is more efficient for complex textured objects than more conventional SIFT and MI based registration techniques.

## 2. MGRF Based Image Registration

### 2.1. Basic notation.

We denote $\mathcal{Q} = \{0, \ldots, Q-1\}$; $\mathbf{R} = [(x,y) : x = 0, \ldots, X-1; y = 0, \ldots, Y-1]$, and $\mathbf{R_p} \subset \mathbf{R}$ a finite set of scalar image signals (e.g. gray levels), a rectangular arithmetic lattice supporting digital images $g : \mathbf{R} \to \mathcal{Q}$, and its arbitrary-shaped part occupied by the prototype, respectively. A finite set $\mathcal{N} = \{(\xi_1, \eta_1), \ldots, (\xi_n, \eta_n)\}$ of $(x,y)$-coordinate offsets defines neighbors $\{((x+\xi, y+\eta), (x-\xi, y-\eta)) : (\xi, \eta) \in \mathcal{N}\} \wedge \mathbf{R_p}$ interacting with each pixel $(x,y) \in \mathbf{R_p}$. The set $\mathcal{N}$ yields a neighborhood graph on $\mathbf{R_p}$ to specify translation invariant pairwise interactions with $n$ families $\mathcal{C}_{\xi,\eta}$ of cliques $c_{\xi,\eta}(x,y) = ((x,y), (x+\xi, y+\eta))$. Interaction strengths are given by a vector $\mathbf{V}^{\mathsf{T}} = \left[ \mathbf{V}_{\xi,\eta}^{\mathsf{T}} : (\xi, \eta) \in \mathcal{N} \right]$ of potentials $\mathbf{V}_{\xi,\eta}^{\mathsf{T}} = \left[ V_{\xi,\eta}(q, q') : (q, q') \in \mathcal{Q}^2 \right]$ depending on signal co-occurrences; here $\mathsf{T}$ indicates transposition.

### 2.2. Image normalization.

To account for monotone (order-preserving) changes of signals (e.g. due to different illumination or sensor characteristics), the prototype and object images are equalized using the cumulative empirical probability distributions of their signals on $\mathbf{R_p}$.

### 2.3. MGRF based appearance model.

In line with a generic MGRF with multiple pairwise interaction [36, 37] (see Fig. 1), the Gibbs probability $P(g) \propto \exp(E(g))$ of an object $g$ aligned with the prototype $g^\circ$ on $\mathbf{R_p}$ is specified with the Gibbs energy:

$$E(g) = |\mathbf{R_p}| \mathbf{V}^{\mathsf{T}} \mathbf{F}(g) \qquad (1)$$

where $\mathbf{F}^{\mathsf{T}}(g)$ is the vector of scaled empirical probability distributions of signal co-occurrences over each clique family: $\mathbf{F}^{\mathsf{T}}(g) = [\rho_{\xi,\eta} \mathbf{F}_{\xi,\eta}^{\mathsf{T}}(g) : (\xi, \eta) \in \mathcal{N}]$ where $\rho_{\xi,\eta} = \frac{|\mathcal{C}_{\xi,\eta}|}{|\mathbf{R_p}|}$ is the relative size of the family and $\mathbf{F}_{\xi,\eta}(g) = [f_{\xi,\eta}(q, q'|g) : (q, q') \in \mathcal{Q}^2]^{\mathsf{T}}$; here, $f_{\xi,\eta}(q, q'|g) = \frac{|\mathcal{C}_{\xi,\eta;q,q'}(g)|}{|\mathcal{C}_{\xi,\eta}|}$ are empirical probabilities of signal co-occurrences, and $\mathcal{C}_{\xi,\eta;q,q'}(g) \subseteq \mathcal{C}_{\xi,\eta}$ is a subfamily of the cliques $c_{\xi,\eta}(x,y)$ supporting the co-occurrence $(g_{x,y} = q, g_{x+\xi,y+\eta} = q')$ in $g$. The co-occurrence distributions and the Gibbs energy for the object are determined over $\mathbf{R_p}$, i.e. within the prototype boundary after an object is affinely aligned with the prototype. To account for the affine transformation, the initial image is resampled to the back-projected $\mathbf{R_p}$ by interpolation.

The appearance model consists of the neighborhood $\mathcal{N}$ and the potential $\mathbf{V}$ to be learned from the prototype.
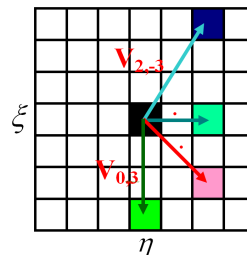


Figure 1. Pairwise pixel interaction MGRF model.

### 2.4. Learning the potentials.

The MLE of $\mathbf{V}$ is proportional in the first approximation to the scaled centered empirical co-occurrence distributions

for the prototype. We use here our novel analytical maximum likelihood estimator for the Gibbs potentials of the pair-wise MGRF [35]:

$$V_{\xi,\eta}(q,q') = \lambda \rho_{\xi,\eta}[f_{\xi,\eta}(q,q') - f(q)f(q')] \qquad (2)$$

where $f_{(.)}$ and $f_{\xi,\eta}(q,q')$ be a joint empirical probability distribution of pixel intensities and of intensity co-occurrences, respectively. The common scaling factor $\lambda$ is also computed analytically; it is approximately equal to $Q^2$ if $Q \gg 1$, $f(q) \approx 1/Q$, and $\rho_{\xi,\eta} \approx 1$ for all $(\xi,\eta) \in \mathcal{N}$. In our case it can be set to $\lambda = 1$ because the registration uses only relative potential values and energies.

## 2.5. Learning the characteristic neighbors.

To find the characteristic neighborhood set $\mathcal{N}$, the relative energies $E_{\xi,\eta}(g^\circ) = \rho_{\xi,\eta}\mathbf{V}_{\xi,eta}^\top\mathbf{F}_{\xi,\eta}(g^\circ)$ for the clique families, i.e. the scaled variances of the corresponding empirical co-occurrence distributions, are compared for a large number of possible candidates. Figure 2 shows a zebra prototype and its Gibbs energies $E_{\xi,\eta}(g^\circ)$ for 5000 clique families with the inter-pixel offsets $|\xi| \leq 50$; $0 \leq \eta \leq 50$.

To automatically select the characteristic neighbors, we consider an empirical probability distribution of the energies as a mixture of a large "non-characteristic" low-energy component and a considerably smaller characteristic high-energy component: $P(E) = \pi P_{\text{lo}}(E) + (1 - \pi)P_{\text{hi}}(E)$. Both the components $P_{\text{lo}}(E)$ and $P_{\text{hi}}(E)$ are of arbitrary shape and thus are approximated with linear combinations of positive and negative discrete Gaussians (efficient EM-based algorithms introduced in [36] are used for both the approximation and estimation of $\pi$, Fig. 3 shows all steps of density estimation using the modified EM algorithm [36]). To the best of our knowledge, we are the first authors who developed an automatic way to estimate the neighborhood system from the learned data (see Fig. 4(a) for zebra, Fig. 13(a) for the kidney, and Fig. 18(a) for the lung).
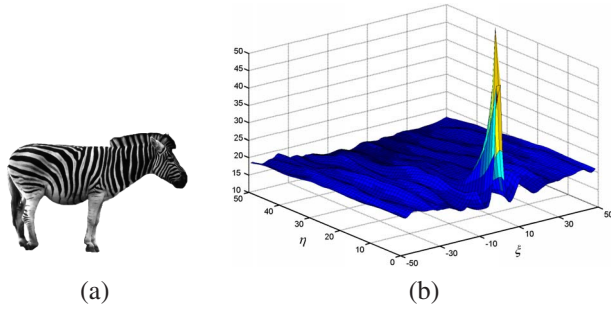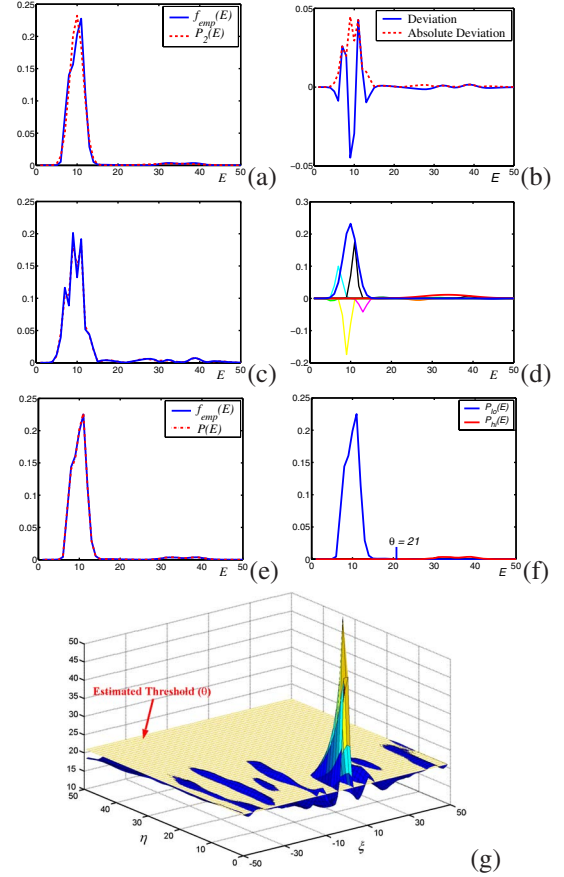


Figure 3. (a) Empirical density of the Gibbs energy $f_{emp}(E)$ and the estimated two dominant Gaussian components $P_2(E)$ which represent low-energy component and high-energy component, (b) Deviation and absolute deviation between the empirical density and the two dominant components, (c) The estimated density for the absolute deviation, (d) The final LCDG components which are used in the estimation of mixed density shown in (e) and the marginal estimated density (f) giving the energy threshold $\theta$ to select the neighborhood system as shown in (g).



Figure 2. Zebra prototype (a) and relative interaction energies (b) for the clique families in function of the offsets $(\xi, \eta)$.

The intersection of the approximate mixture components gives an energy threshold $\theta$ for selecting the characteristic neighbors: $\mathcal{N} = \{(\xi,\eta) : E_{\xi,\eta}(g^\circ) \geq \theta\}$ where $P_{\text{hi}}(\theta) \geq P_{\text{lo}}(\theta)\pi/(1-\pi)$. The above example results in the threshold $\theta = 21$ producing 168 characteristic neighbors shown in
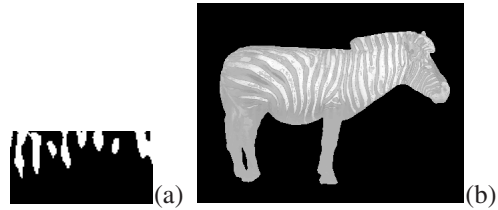


Figure 4. (a) Characteristic 168 neighbors among the 5000 candidates (a; in white) and the pixel-wise Gibbs energies (b) for the prototype under the estimated neighborhood.

Fig. 4 together with the corresponding relative pixel-wise energies $e_{x,y}(g^\circ)$ over the prototype:

$$e_{x,y}(g^\circ) = \sum_{(\xi,\eta)\in\mathcal{N}} V_{\xi,\eta}(g^\circ_{x,y}, g^\circ_{x+\xi,y+\eta})$$
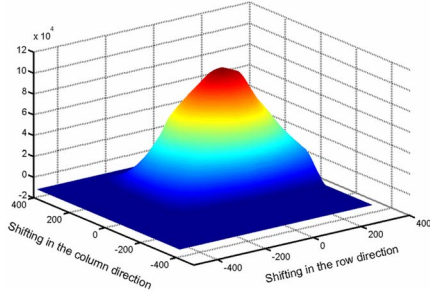
Figure 5. Gibbs energies for translations of the object with respect to the prototype.
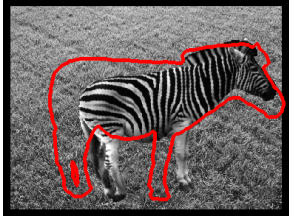


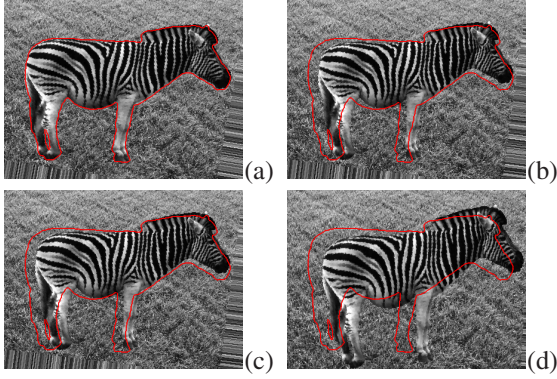Figure 6. Initial position of the object with respect to the prototype.



Figure 7. Our (a), MI-based (b), NMI-based (c), and SIFT-based (d) registration.

## 2.6. Appearance-based registration.

The object $g$ is affinely transformed to (locally) maximize its relative energy $E(g_{\mathbf{a}}) = \mathbf{V}^{\mathsf{T}}\mathbf{F}(g_{\mathbf{a}})$ under the learned appearance model $[\mathcal{N}, \mathbf{V}]$. Here, $g_{\mathbf{a}}$ is the part of the object image reduced to $\mathbf{R_p}$ by the affine transformation $\mathbf{a} = [a_{11}, \ldots, a_{23}]$: $x' = a_{11}x + a_{12}y + a_{13}$; $y' = a_{21}x + a_{22}y + a_{23}$. The initial transformation is a pure translation with $a_{11} = a_{22} = 1$; $a_{12} = a_{21} = 0$, ensuring the most "energetic" overlap between the object and prototype. The energy for different translations $(a_{13}, a_{23})$ of the object relative to the prototype is shown in Fig. 5; the chosen initial position $(a_{13}^*, a_{23}^*)$ in Fig. 6 maximizes this energy. Then the gradient search for the local energy maximum closest to the initialization selects the six parameters $\mathbf{a}$; Fig. 7 (a) shows the final transformation aligning the prototype contour to the object.

## 3. Experimental Results

Due to space limitations, we focus on zebra photos, dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) of human kidney, and low dose computed tomography (LDCT) of human lung commonly perceived as difficult for both the area- and feature-based registration. But the like results are obtained for several other types of complex objects (e.g., starfish photos or brain images). We compare our approach to three popular conventional techniques, namely, to the area-based registration using MI [26] or normalized MI [28] and to the feature-based registration establishing correspondences between the images with SIFT [18]. *Note that for both MI and NMI we used the implementations in ITK (Ver. 2.0).* Results are shown in Fig. 7.

To clarify why the MI- or NMI-based alignment is less accurate, Fig. 8 compares the MI / NMI and Gibbs energy values for the affine parameters that appear at successive steps of the gradient search for the maximum energy. Both the MI and NMI have many local maxima that potentially hinder the search, whereas the energy is practically unimodal in these experiments. The SIFT-based alignment fails because it cannot establish accurate correspondences between similar zebra stripes (Fig. 9).
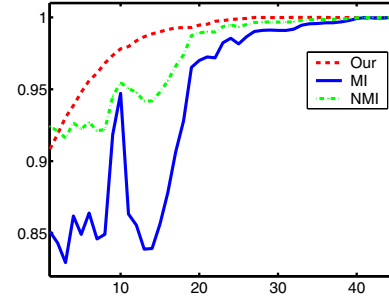


Figure 8. Gibbs energy, MI, and NMI values at the successive steps of the gradient search.
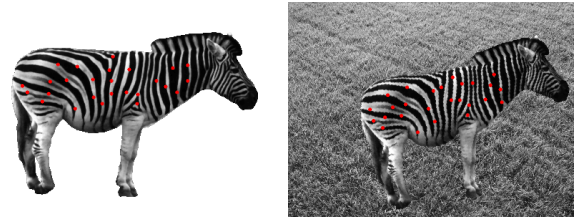


Figure 9. Corresponding points by SIFT.

In the above example the object aligned with the prototype has mainly different orientation and scale. Figure 10 shows more diverse zebras and their Markov-Gibbs appearance-based and MI-based alignment with the prototype in Fig. 2(a). Visually, the back-projection of the prototype contour onto the objects suggests the better performance of our approach. To quantitatively evaluate the accuracy, masks of the co-aligned objects obtained by manual
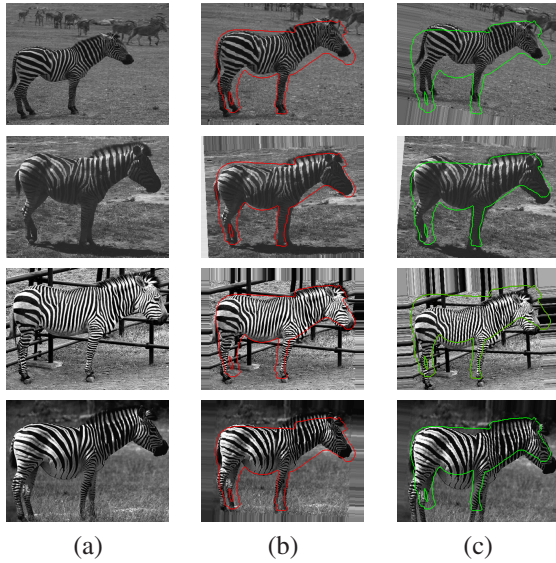
Figure 10. Original zebras (a) aligned with our (b) and the MI-based (c) approaches.
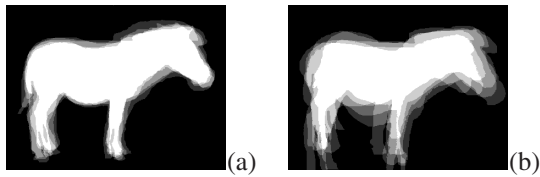


Figure 11. Overlap between the object masks aligned with our (a; 91.6%) and the MI-based approaches (b; 70.3%).
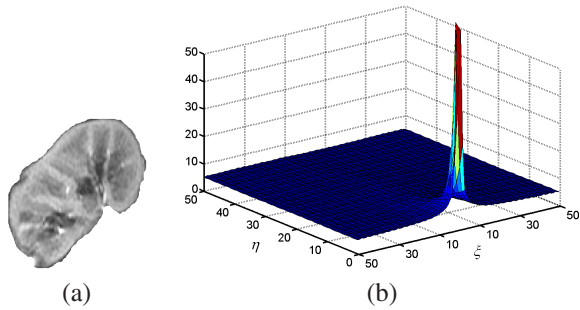


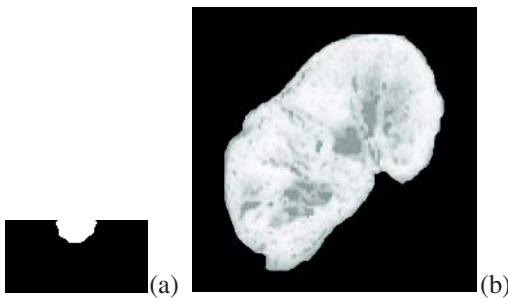Figure 12. Kidney image (a) and relative interaction energies (b) for the clique families in function of the offsets $(\eta, \xi)$.



Figure 13. (a) Most characteristic 76 neighbors among the 5000 candidates (a; in white) and the pixel-wise Gibbs energies (b) for the prototype under the estimated neighborhood.
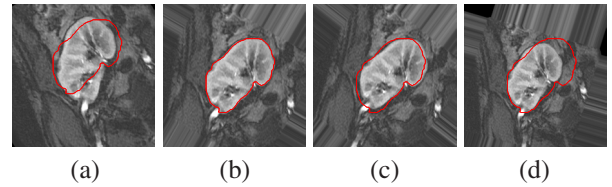


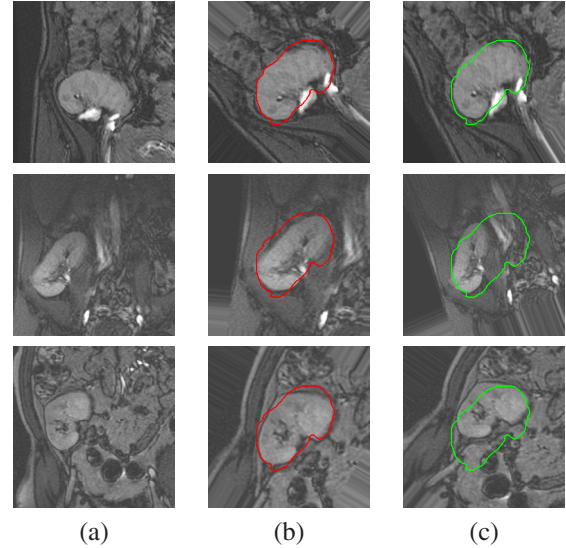Figure 14. Initialization (a) and our (b), MI-based (c), and SIFT-based (d) registration.



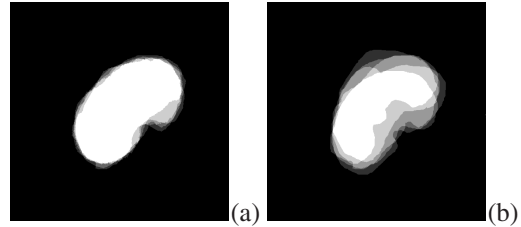Figure 15. Original kidneys (a) aligned with our (b) and the MI-based (c) approaches.



Figure 16. Overlap between the object masks aligned with our (a; 90.2%) and the MI-based approaches (b; 62.6%).
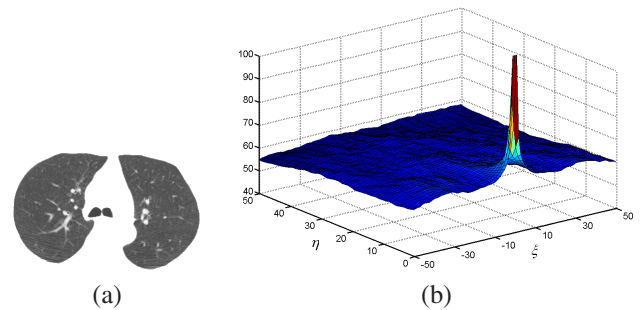


Figure 17. Lung image (a) and relative interaction energies (b) for the clique families in function of the offsets $(\eta, \xi)$.

segmentation are averaged in Fig. 11. The common matching area is notably larger for our approach (91.6%) than for the MI-based registration (70.3%). Similar results obtained for the kidney images are shown in Figs. 12–16: the com-
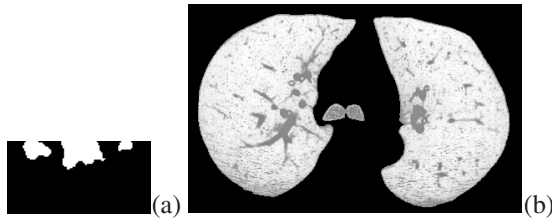
Figure 18. (a) Most characteristic 173 neighbors among the 5000 candidates (a; in white) and the pixel-wise Gibbs energies (b) for the prototype under the estimated neighborhood.
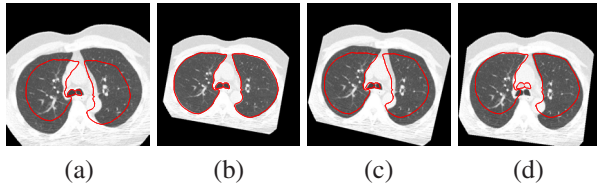


(a) (b) (c) (d)

Figure 19. Initialization (a) and our (b), MI-based (c), and SIFT-based (d) registration.
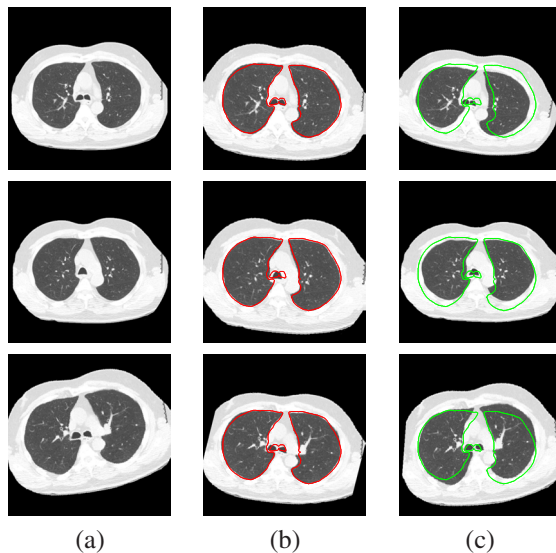


(a) (b) (c)

Figure 20. Original lungs (a) aligned with our (b) and the MI-based (c) approaches.
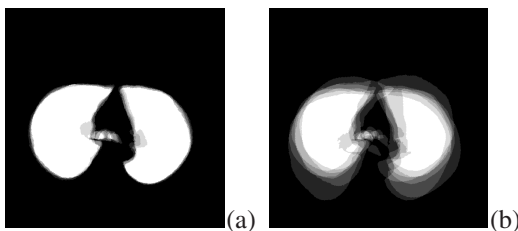


(a) (b)

Figure 21. Overlap between the object masks aligned with our (a; 96.8%) and the MI-based approaches (b; 54.2%).

mon matching area 90.2% is for our approach vs. 62.6% for the MI-based one. The final application which we introduced in this paper is the alignment of LDCT lung images. The results of LDCT lung images are shown in Figs. 17–21: the common matching area 96.8% is for our approach vs. 54.2% for the MI-based one.

## 4. Conclusions

In this paper we introduced a new approach to align an image of a textured object with a given prototype whose appearance is modeled with a Markov-Gibbs random field with pairwise interaction. Experimental results confirm that image registration based on our Markov-Gibbs appearance model is more robust and accurate than popular conventional algorithms. Due to the reduced variations between the co-aligned objects, our approach results in more accurate average shape models that are useful, e.g. in image segmentation based on shape priors.

As we mentioned in the experimental result section, the proposed approach is not only limited to zebra photos, kidney and lung images but also is suitable for registering starfish photos and brain images. The latter were not included in the paper because of the space limitations, but, the registration results for the starfishes and brain images are provided on our website[1].

Our future work will focus on generalizing the proposed appearance Markov-Gibbs random field model to learn the appearance model of 3D volumes instead of 2D images. Also, we will integrate our approach in a framework of shape based segmentation; which we believe will enhance the accuracy of segmentation results of the existing approaches.

## References

[1] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

[2] A. Goshtasby and G. Stockman, "Point pattern matching using convex hull edges," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 15, pp. 631–637, 1985.

[3] M. Holm, "Towards automatic rectification of satellite images using feature based matching," *In Proc. IGARSS*, Espoo, Finland, June 3–6, 1991, pp. 2439–2442.

[4] J. Hsieh, H. Liao, K. Fan, M. Ko, and Y. Hung, "Image registration using a new edge-based approach," *Computer Vision and Image Understanding*, vol. 67, pp. 112–130, 1997.

[5] M. Roux, "Automatic registration of SPOT images and digitized maps," *In Proc. ICIP*, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 625–628.

[6] Y. Hsieh, D. McKeown, and F. Perlant, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction," *IEEE Trans. PAMI*, vol. 14, pp. 214–237, 1992.

[1]http://uofl.edu/speed/bioengineering/faculty/bioengineering-full/dr-ayman-el-baz/elbazlab.html.

[7] S. Moss and E. Hancock, "Multiple line-template matching with EM algorithm," *Pattern Recognition Letters*, vol. 18, pp. 1283–1292, 1997.

[8] W. Wang and Y. Chen, "Image registration by control points pairing using the invariant properties of line segments," *Pattern Recognition Letters*, vol. 18, pp. 269–281, 1997.

[9] X. Dai and S. Khorram, "Development of a feature-based approach to automated image registration for multitemporal and multisensor remotely sensed imagery," *In Proc. IGARSS*, Singapore, August 3–8, 1997, pp. 243–245.

[10] V. Govindu, C. Shekhar, and R. Chellapa, "Using geometric properties for correspondenceless image alignment," *In Proc. ICPR*, Brisbane, Australia, August 16–20, 1998, pp. 37–41.

[11] H. Li, B. Manjunath, and S. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Processing*, vol. 4, pp. 320–334, 1995.

[12] H. Maitre and Y. Wu, "Improving dynamic programming to solve image registration," *Pattern Recognition*, vol. 20, pp. 443–462, 1987.

[13] D. Shin, J. Pollard, and J. Muller, "Accurate geometric correction of ATSR images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 35, pp. 997–1006, 1997.

[14] S. Li, J. Kittler, and M. Petrou, "Matching and recognition of road networks from aerial images," *In Proc. ECCV*, Santa Margherita Ligure, Italy, May 19–22, 1992, pp. 857–861.

[15] E. Mendoza, J. Santos, A. Rosa, and N. Silva, "Land use/land cover mapping in Brazilian Amazon using neural network with aster/terra data," *In Proc. Geo-Imagery Bridging Continents*, Istanbul, Turkey, 2004, pp. 123–126.

[16] S. Growe and R. Tonjes, "A knowledge based approach to automatic image registration," *In Proc. ICIP*, Santa Barbara, California, Oct. 26–29, 1997, pp. 228–231.

[17] J. Ton and A. Jain, "Registering landsat images by point matching," *IEEE Trans. Geoscience and Remote Sensing*, vol. 27, pp. 642–651, 1989.

[18] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comp. Vision*, vol. 60, pp. 91-110, 2004.

[19] P. Pope and J. Theiler, "Automated image registration (AIR) of MTI imagery," *In Proc. SPIE*, vol. 27, pp. 294–300, 2003.

[20] P. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform," *IEEE Trans. Geoscience Electronics*, vol. 8, pp. 353–368, 1970.

[21] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition," *IEEE Trans. PAMI*, vol. 16, pp. 1156–1168, 1994.

[22] B. Reddy and B. Chatterji, "An FFT-based technique for translation, rotation and scale-invariant image registration," *IEEE Trans. Image Processing*, vol. 5, pp. 1266–1271, 1996.

[23] G. Wolberg and S. Zokai, "Robust image registration using log-polar transform," *In Proc. ICIP*, Canada, Sept. 10–13, 2000, pp. 493–496.

[24] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Processing*, vol. 11, pp. 188–200, 2002.

[25] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geoscience and Remote Sensing*, vol. 42, pp. 2104–2120, Oct. 2004.

[26] P. Viola, "Alignment by maximization of mutual information," *Ph.D. dissertation, MIT*, Cambridge, MA, 1995.

[27] P. Thevenaz and M. Unser, "Alignment an efficient mutual information optimizer for multiresolution image registration," *In Proc. ICIP*, Chicago, Oct. 4–7, 1998, pp. 833–837.

[28] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–86, 1999.

[29] S. Roy and V. Govindu, "MRF solutions for probabilistic optical flow formulations," *In Proc. ICPR*, Barcelona, Spain, Sept. 3–8, 2000, vol III, pp. 1041-1047.

[30] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. PAMI*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[31] M. Kumar, P. Torr, and A. Zisserman, "Learning layered motion segmentation of video," *In Proc. ICCV*, Beijing, China, Oct. 17–20, 2005, pp. 33-40.

[32] J. Winn and N. Jojic, "Locus: Learning object classes with unsupervised segmentation," *In Proc. ICCV*, Beijing, China, Oct. 17–20, 2005, pp. 756-763.

[33] I. Kovtun, "Image segmentation based on sufficient conditions of optimality in NP-complete classes of structural labelling problem," *PhD thesis, IRTC ITS National Academy of Science Ukraine*, 2004.

[34] A. Shekhovtsov, I. Kovtun, V. Hlavac, "Efficient MRF deformation model for non-rigid image matching," *In Proc. CVPR*, Minneapolis, MN, June 18–23, 2007, pp. 1–6.

[35] G. Gimel'farb and D. Zhou, "Accurate identification of a Markov-Gibbs model for texture synthesis by bunch sampling", *In Proc. CAIP*, Vienna, Austria, 2007, pp. 979-986.

[36] A. El-Baz, "Novel stochastic models for medical image analysis," *Ph.D. dissertation, University of Louisville*, Louisville, KY, Chapter 3, pp. 61–94, 2006.

[37] G. Gimelfarb, *Image textures and Gibbs random fields,* Dordrecht, The Netherlands: Kluwer, 1999.