

Correspondence Detection Using Wavelet-Based Attribute Vectors

Zhong Xue, Dinggang Shen, and Christos Davatzikos

Section of Biomedical Image Analysis, Department of Radiology
University of Pennsylvania, Philadelphia, PA 19104
{zxue,dgshen,christos}@rad.upenn.edu

Abstract. Finding point correspondence in anatomical images is a key step in shape analysis and deformable registration. This paper proposes an automatic correspondence detection algorithm using wavelet-based attribute vectors defined on every image voxel. The attribute vector reflects the anatomical characteristics in a large neighborhood around the respective voxel. It plays the role of a morphological signature for each voxel and is therefore made as distinctive as possible. Correspondence is then determined via similarity of attribute vectors. Experiments with brain MR images show that the algorithm performs at least as well as human experts, even for complex cortical structures.

1 Introduction

Deformable registration of MR images is used actively for anatomical segmentation and labeling, and for morphological analysis via shape transformations [1, 2,3]. Many methods have been developed, including image similarity-based registration and feature-based registration. In image similarity-based methods, the registration is achieved by minimizing a cost function [4] that expresses a trade-off between similarity of the warped image and the target image, and smoothness of the transformation, *e.g.* maximizing mutual information (MI) [5] and local frequency representations [6]. However, image similarity-based methods capture only the local image information around voxels, and are therefore sensitive to initialization and particularly vulnerable to local minima. On the other hand, feature-based methods first detect anatomical features such as surfaces, landmark points, or ridges, and then determine a 3D spatial transformation so that corresponding features are mapped to each other, thus they are less susceptible to these shortcomings [7,8]. Although feature-based methods can rely on manual definition of point or feature correspondence, routine clinical use requires effective automated methods for finding correspondence that can subsequently guide a 3D transformation. In this paper, we combine the strengths of these two approaches by taking the notion of a feature to its extreme: every voxel is regarded as a feature point. In order to achieve accurate and automated feature identification and matching, we use an *Attribute Vector* (AV) that is defined for each voxel and serves as its morphological signature.

In this paper, we focus on the design and implementation of a distinctive and robust AV, which is calculated directly from the original image, so that correspondence between two images can be determined reliably using AV similarity. Generally, good AVs should possess three properties: i) they must be translation and rotation-invariant; ii) they must be distinctive in order to be spatially specific; iii) they must be robust to morphological variability. Under these ideal conditions, the AV similarities can be used to find out correspondence between different images of different individuals. In practice, we seek representations that make AVs as distinctive as possible.

In our method, the voxel-wise AVs are designed for correspondence detection in a multi-resolution framework. To construct the AVs, a Radial Profiling (RP) method is proposed to extract rotation-invariant feature vectors from the feature images, which are calculated from MR images using Discrete Wavelet Transformation (DWT). For each voxel, the DWT of the image data within a sliding window centered on it is performed, which yields a series of subimages, and the feature images are formed by combining these wavelet subimages. The Wavelet-based AV (WAV) is constructed by applying RP on the feature images. Although WAV is not rotation-invariant, it can tolerate a small rotation angle, and it gives detailed representation of anatomical features around a voxel. In practice, some rigid registration methods can be used first if the rotational angle between two brains is large, then, the WAV is further applied to determine the correspondence.

Automatic correspondence detection for 3D MR images is achieved based on evaluating the AV similarity. Three AV similarity criteria are examined: direct AV matching, matching based on the statistics of AVs, and matching by incorporating shape prior constraints. Experiments on finding anatomic correspondence in human brain MR images are carried out to validate the effectiveness and robustness of the proposed AV and AV similarity algorithms. The similarity of the AVs of the voxels selected from different sulci is evaluated by using a statistical analysis method, revealing a very high spatial specificity of the proposed AV. Moreover, the three AV similarity criteria are compared, with the landmark points generated by two experts being used as a reference for comparison purposes. Experimental results indicate that the proposed algorithm performs at least as well as experienced raters in defining pairs of anatomically corresponding points, while being fully automated.

2 Correspondence Detection Using Attribute Vectors

We approach automatic correspondence detection as a procedure that requires three steps: i) design of attribute vectors to capture the anatomical characteristics around each voxel, ii) calculation of the similarity between AVs, and iii) selection of the optimal matching voxel.

2.1 Design of the Attribute Vector

Radial Profiling: Extracting Rotation-Invariant Features. We first present a feature extraction algorithm named “Radial Profiling” (RP), which can be used to extract rotation-invariant features from images. An image, here, can represent not only the intensity of the original input MR image, but also features extracted from it. In general, the proposed RP algorithm can be regarded as a special case of the rotation-invariant moments.

Denote $D(\mathbf{x}) (\mathbf{x} \in R^3)$ as the feature image calculated from the original image $I(\mathbf{x})$. For the voxel of interest \mathbf{x}_0 , $D(\mathbf{x})$ can also be represented using polar coordinates, $D_{\mathbf{x}_0}^{polar}(r, \theta, \phi)$ by setting \mathbf{x}_0 as the new origin. Then, RP calculates the means and variances for a series of rings centered on \mathbf{x}_0 :

$$u(n) = \text{Mean}\{D_{\mathbf{x}_0}^{polar}(r, \theta, \phi) |_{n\Delta r \leq r < (n+1)\Delta r, 0 \leq \theta < \pi, 0 \leq \phi < 2\pi}\}, \quad (1)$$

$$w(n) = \text{Var}\{D_{\mathbf{x}_0}^{polar}(r, \theta, \phi) |_{n\Delta r \leq r < (n+1)\Delta r, 0 \leq \theta < \pi, 0 \leq \phi < 2\pi}\}, \quad (2)$$

$$0 \leq n \leq N_r - 1,$$

The thickness of a ring is Δr , and its internal radius is $n\Delta r$. N_r is the number of the rings, thus $N_r\Delta r$ is the radius of the neighborhood. Finally, a rotation-invariant attribute vector at voxel \mathbf{x}_0 is given by

$$\mathbf{v} = [u(0), w(0), u(1), w(1), \dots, u(N_r - 1), w(N_r - 1)]^T. \quad (3)$$

It can be seen that $u(n)$ and $w(n)$ (and therefore \mathbf{v}) are rotation-invariant to D . Therefore, \mathbf{v} is rotation-invariant to the original input image I , provided its feature image D always follows the rotation of the input image I .

Wavelet-Based Attribute Vector. The AV of a voxel reflects the anatomy in the vicinity of that voxel, thus it should be calculated from the image data within a neighborhood of the voxel. The challenge here is to represent this anatomy in a most compact way. In our method the DWT [9] is used to construct the AVs.

To calculate the WAV of a voxel \mathbf{x}_0 , we examine the image characteristics within a cubic neighborhood centered on it. Denoting the image data within this neighborhood as $I_{\mathbf{x}_0}(\mathbf{x})$, the DWT is used to decompose $I_{\mathbf{x}_0}(\mathbf{x})$, thereby producing 8 component subimages (1 low-pass subimage and 7 high-pass subimages). The low-pass subimage, $I_{LLL}^{(1)}$, can be further decomposed, resulting in a second-level DWT decomposition. This procedure continues until the J th level DWT decomposition is performed.

To construct the WAV of a voxel at each DWT level j , selected high-pass subimages are combined to form a feature image $D_{\mathbf{x}_0}^{(j)}(\mathbf{x})$, ($j = 1, 2, \dots, J$) [10],

$$D_{\mathbf{x}_0}^{(j)}(\mathbf{x}) = \sqrt{|I_{LLH}^{(j)}(\mathbf{x})|^2 + |I_{LHL}^{(j)}(\mathbf{x})|^2 + |I_{HLL}^{(j)}(\mathbf{x})|^2}, \quad 1 \leq j \leq J, \quad (4)$$

where $I_{LLH}^{(j)}(\mathbf{x})$, $I_{LHL}^{(j)}(\mathbf{x})$ and $I_{HLL}^{(j)}(\mathbf{x})$ are the selected high-pass subimages at resolution level j . Experimental results show that this combination is insensitive to rotation, and is effective as long as the rotation is relatively small.

Then the RP is applied to $D_{\mathbf{x}_0}^{(j)}$, resulting in J wavelet feature vectors each formed according to Eq.(3). In addition, the feature vector of the low-pass subimage $I_{LLL}^{(j)}$ at DWT level J is also calculated using RP: the attribute vector \mathbf{v}_j at j th DWT level is calculated from the feature image $D_{\mathbf{x}_0}^{(j)}(\mathbf{x})$, where \mathbf{x}_0 is the center of $D_{\mathbf{x}_0}^{(j)}(\mathbf{x})$. The WAV of voxel \mathbf{x}_0 , \mathbf{v}_w , is then constructed by concatenating all these $J + 1$ feature vectors:

$$\mathbf{v}_w = [\lambda_1 \mathbf{v}_1^T, \lambda_2 \mathbf{v}_2^T, \dots, \lambda_T \mathbf{v}_J^T, \lambda_D \mathbf{v}_D^T]^T, \tag{5}$$

where \mathbf{v}_D is calculated from $I_{LLL}^{(j)}$ using the RP method. $\lambda_1, \dots, \lambda_T$ and λ_D are weighting coefficients. Moreover, different N_r can be used for different resolutions. This procedure for constructing the WAV is illustrated in Fig.1.

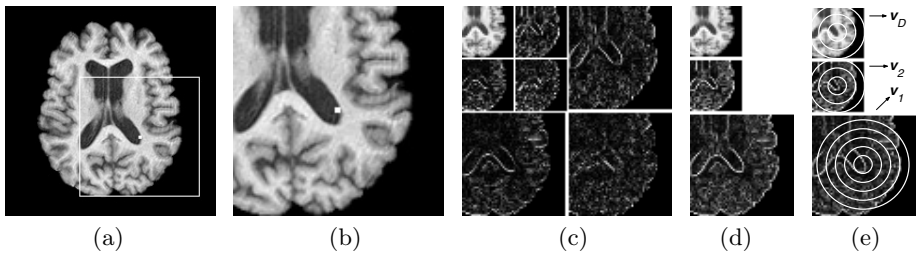


Fig. 1. Construction of the wavelet-based attribute vector. (a) The input image; (b) the image inside the sliding window; (c) J -level wavelet decomposition (here $J=2$); (d) the low-pass subimage (top) and the combined high-pass subimages; (e) constructing the WAV using RP, $\mathbf{v}_w = [\lambda_1 \mathbf{v}_1^T, \lambda_2 \mathbf{v}_2^T, \lambda_D \mathbf{v}_D^T]^T$.

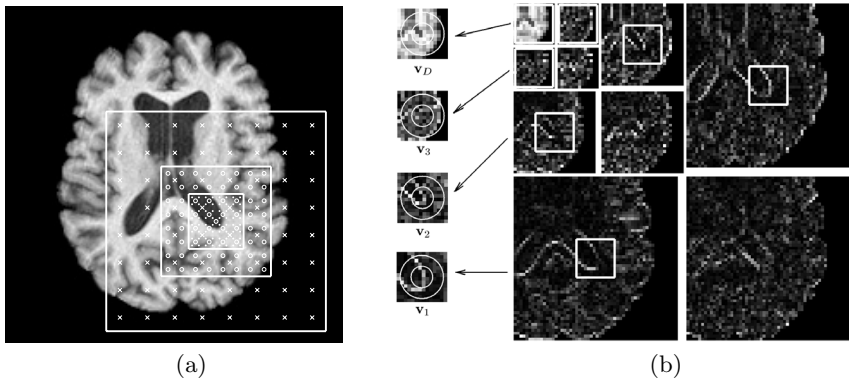


Fig. 2. Computationally efficient approximation of WAV. (a) Representing the neighborhoods to be considered in a multi-resolution fashion; (b) applying radial profiling to the wavelet transformed images.

Computationally Efficient Approximation of WAV. Unfortunately, the computational load necessary to calculate the WAV for every voxel of an image is very heavy. To overcome this problem and make the implementation practical, a multi-resolution framework is proposed for calculating WAV, which acts as an approximation of the original algorithm, and it also has the effect of feature reduction and selection. This framework is a computationally efficient calculation of certain wavelet expansions that we need to include in the AV. It is based on two facts: i) Relatively global anatomical characteristics must be determined from a relatively large window size around the voxel of interest. However, this window can be sub-sampled quite sparsely, without significantly changing the global characteristics as shown in Fig.2(a)(the “x” points); ii) High-frequency local characteristics needs to be calculated only from a small neighborhood around the voxel; (see Fig.2(b)). To implement this algorithm, first the multi-resolution images are generated using Gaussian filtering and down-sampling. Then multi-level DWTs are performed at each resolution level and the feature vectors are calculated using RP. Finally, the WAV of a voxel is constructed by combining all the feature vectors.

2.2 Correspondence Detection: Maximizing the Likelihood of AVs

The Forward Matching Problem. The forward matching problem is described as: given a particular location \mathbf{x} in the model image, how to determine the corresponding voxel \mathbf{y}^* in an individual’s image by maximizing that likelihood. We assume that the set $\Omega_v(\mathbf{x})$ of the AVs of the corresponding voxels at a particular location \mathbf{x} is characterized by a single Gaussian distribution with mean $\bar{\mathbf{v}}$ and covariance matrix Σ_v . Then, the likelihood of AV, $\mathbf{v}(\mathbf{y})$, at location \mathbf{y} in an individual’s image, belonging to $\Omega_v(\mathbf{x})$, is given by

$$f(\mathbf{v}(\mathbf{y})|\Omega_v(\mathbf{x})) = \eta \cdot \exp\left\{-\frac{1}{2}E_v\right\}, \quad (6)$$

where η is a normalization coefficient, and $E_v = (\mathbf{v}(\mathbf{y}) - \bar{\mathbf{v}})^T \Sigma_v^{-1} (\mathbf{v}(\mathbf{y}) - \bar{\mathbf{v}})$. Then, the corresponding voxel \mathbf{y}^* in the image is determined by maximizing the likelihood in Eq.(6),

$$\mathbf{y}^*(\mathbf{x}) = \arg \max_{\mathbf{y} \in N_s(\mathbf{x})} \{f(\mathbf{v}(\mathbf{y})|\Omega_v(\mathbf{x}))\}, \quad (7)$$

where $N_s(\mathbf{x})$ denotes the searching area in the subject image.

When there are no training samples available, Eq.(6) is approximated using the similarity between two attribute vectors:

$$f(\mathbf{v}(\mathbf{y})|\Omega_v(\mathbf{x})) \sim f(\mathbf{v}(\mathbf{y})|\{\mathbf{v}_m(\mathbf{x})\}) \sim \eta \exp\left\{-\frac{\|\mathbf{v}(\mathbf{y}) - \mathbf{v}_m(\mathbf{x})\|^2}{\epsilon}\right\}, \quad (8)$$

where $\mathbf{v}_m(\mathbf{x})$ is the attribute vector of the model (template) image at location \mathbf{x} , and ϵ is a parameter that would ideally depend on the expected range of $\|\mathbf{v} - \mathbf{v}_m\|$. When training samples are available, the PCA is used to estimate the likelihood in Eq.(6), using only the major principal components.

The Backward Matching Problem: Consistency. The backward matching problem is discussed here, with the objective being to investigate the reliability of the correspondence detection. In particular, \mathbf{y}^* is a reliable match for \mathbf{x} if \mathbf{x} is also a good match for \mathbf{y}^* , when one considers the inverse matching problem. Inverse consistency has been discussed in [11]. The backward matching procedure maximizes the likelihood of AVs within a neighborhood of \mathbf{x} in the model image domain:

$$\mathbf{x}^* = \arg \max_{\mathbf{z} \in N(\mathbf{x})} \{f(\mathbf{v}(\mathbf{y}^*) | \Omega_v(\mathbf{z}))\}. \quad (9)$$

Therefore, if voxels \mathbf{x} and \mathbf{x}^* are the same, or if the distance between them is small, the corresponding voxel pair $(\mathbf{x}, \mathbf{y}^*)$ is consistent. We consider such correspondences to be reliable, since they reflect pairs of points whose AVs are unambiguously matched.

3 Experimental Results

3.1 Experiments with WAV

In this experiment, WAV was used for correspondence detection. Fig.3 shows a representative result. Specifically, Fig.3(a) shows the model image, in which the white point P represents the selected voxel. Fig.3(b) shows the subject image, where the white point P' is the corresponding voxel found by evaluating the direct AV similarity (using Eq.(8)). Fig.3(c) is the similarity map in the subject image space, which shows the similarity between the WAV of P and the WAVs of all the voxels in the second image. In the similarity map (the intensity represents the similarity), the darker points mean the AVs of two voxels are similar, while the whiter point indicates the AVs of the voxels are quite different. Fig.3(c) shows that this pair forms a reliable correspondence.

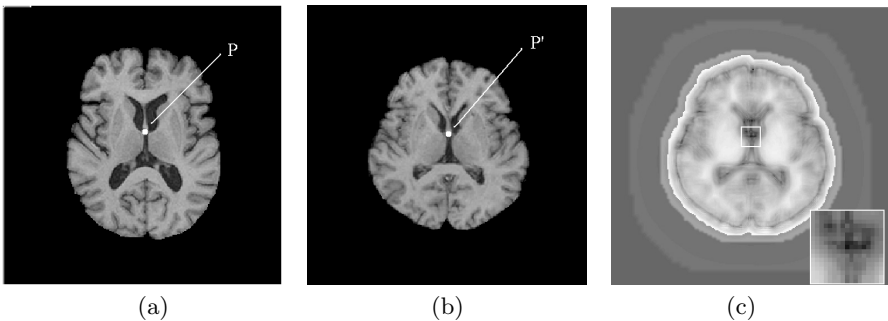


Fig. 3. Finding voxel correspondence using the wavelet-based attribute vector.

3.2 Validation and Statistical Analysis of AVs

In order to evaluate the ability of the AVs to distinctly characterize different anatomical locations, in this section we examine the AVs of the voxels located at roots of different sulci. Five sulci from the left hemisphere are used, *i.e.*, the central sulcus, the superior temporal sulcus, the superior frontal sulcus, the cingulate sulcus and the calcarine sulcus. In order to be able to display the results, we selected 2 voxels from each sulcus, *i.e.*, 10 voxels from all sulci. It is desired that these representative voxels have very different AVs, which would be an indication that our framework for calculating the AV is suitable. Table 1 gives the likelihood of 10 voxels belonging to each other's statistical models. The strong diagonal dominance of this matrix indicates that these 10 voxels have very distinctive attribute vectors. Therefore the designed AV has high specificity in determining correspondence across these 57 brain images of 19 individuals.

3.3 Comparison of Methods for Determining AV Similarity

In this experiment we present a quantitative evaluation of the correspondence determined automatically via AV similarity, by comparing it with the correspondence defined by human experts. For each brain, 18 voxels were manually marked by two experts respectively. Therefore, the correspondences of these 18 voxels among the testing images were known. Three sets of experiments were carried out to find out the corresponding voxels and their matching performance was compared with that of the experts.

- Experiment 1: Use AV similarity (see Eq.(8)).
- Experiment 2: Use maximizing likelihood (see Eq.(7)), where the AVs of the manually marked voxels were utilized as the prior knowledge.
- Experiment 3: Use the statistics of AVs for matching, and incorporate the shape constraints (using the spatial constraints that regard the 18 marked voxels as a shape, see the ASM [12] for details).

Table 1. Likelihoods of 10 different voxels belonging to each other's statistical models.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
Statistical Model of P_1	0.192	0.001	0.001	0.000	0.000	0.016	0.006	0.003	0.001	0.002
Statistical Model of P_2	0.000	0.167	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000
Statistical Model of P_3	0.001	0.000	0.278	0.003	0.007	0.028	0.033	0.001	0.000	0.033
Statistical Model of P_4	0.000	0.000	0.000	0.334	0.000	0.000	0.000	0.000	0.000	0.000
Statistical Model of P_5	0.015	0.001	0.139	0.000	0.408	0.064	0.000	0.018	0.000	0.035
Statistical Model of P_6	0.029	0.021	0.002	0.000	0.004	0.216	0.003	0.014	0.000	0.034
Statistical Model of P_7	0.002	0.005	0.002	0.000	0.000	0.001	0.170	0.000	0.001	0.000
Statistical Model of P_8	0.004	0.003	0.000	0.000	0.000	0.012	0.000	0.175	0.000	0.003
Statistical Model of P_9	0.000	0.000	0.000	0.000	0.001	0.000	0.006	0.000	0.181	0.000
Statistical Model of P_{10}	0.002	0.000	0.000	0.000	0.000	0.020	0.000	0.000	0.001	0.166

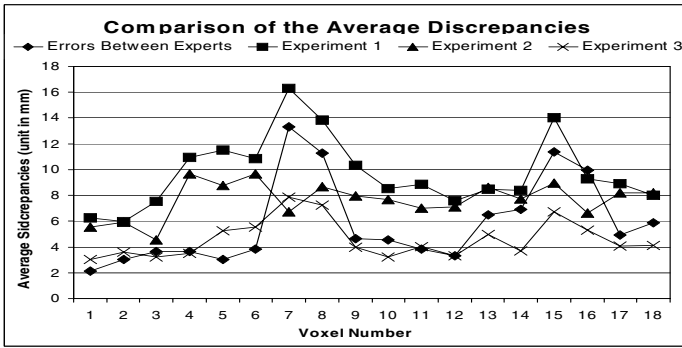


Fig. 4. The average discrepancies.

The average discrepancies between the corresponding voxels found by our algorithm and the voxels marked by the experts for the three experiments are shown in Fig.4. It is clear that among the three experiments, the results for Experiment 1 have the largest discrepancies, while Experiment 3 obtains the best results. For Experiment 2, the results obtained are comparatively close to those of the experts. In conclusion, the designed WAV performs at least as well as the experts. Moreover, among these three AV matching schemes, the method incorporating the shape variation constraint performs the best.

4 Conclusion

This paper proposed an automatic correspondence detection algorithm for 3D MR images using wavelet-based attribute vectors. First, the radial profiling method that serves as a basic rotation-invariant feature extraction tool to construct the attribute vectors is proposed. Then the wavelet-based AV is designed. Correspondence detection is achieved by AV matching. Three AV matching schemes are presented: i) direct AV matching, ii) matching through maximizing the likelihood of AVs, and iii) matching by incorporating shape constraints. Experimental results indicate that the proposed algorithm performs at least as well as experts, and thus it has potential for use in automatic image registration by being incorporated into image warping algorithms.

References

1. J. Duncan and N. Ayache, "Medical image analysis: Progress over two decades and the challenges ahead," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 85–106, 2000.
2. A. Toga and P. Thompson, "The role of image registration in brain mapping," *Image and Vision Computing*, vol. 19, pp. 3–24, 2001.

3. M. Kamber, R. Shinghal, D. Collins, G. Francis, and A. Evans, "Model-based 3d segmentation of multiple sclerosis lesions in magnetic resonance brain images," *IEEE Transactions on Medical Imaging*, vol. 14, pp. 442–453, 1995.
4. D. Rueckert, L. Sonoda, C. Hayes, C. Hill, M. Leach, and D.J.Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.
5. P. Viola and W. Wells, "Alignment by maximization of mutual information," in *5th International Conference on Computer Vision, ICCV'95*, pp. 16–23, 1995.
6. J. Liu, B. Vemuri, and J. Marroquin, "Local frequency representations for robust multimodal image registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 462–469, 2002.
7. L. Staib and J. Duncan, "Boundary finding with parametrically deformable models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1061–1075, 1992.
8. C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Computer Vision and Image Understanding*, vol. 66, pp. 207–222, 5 1997.
9. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, 7 1989.
10. D. Tsai and C. Chiang, "Rotation-invariant pattern matching using wavelet decomposition," *Pattern Recognition Letters*, vol. 23, pp. 191–201, 2002.
11. D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1421–1439, 2002.
12. T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models – their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.