Bayesian Classification of Local 3D Structures in Medical Images

Yiyi Miao, Mehran Kafai, and Kazunori Okada

Computer Science Department, San Francisco State University,
1600 Holloway Ave San Francisco, CA 94132 USA
{kazokada,yiyim,mkafai}@sfsu.edu

Abstract. We propose a Bayesian framework for classifying types of 3D anatomical/pathological structure (i.e., blob, surface, tube, branch) in 3D medical scans. Two probabilistic algorithms with a data-driven statistical modeling scheme are designed to provide accurate classification as well as its decision confidence. The proposed methods are quantitatively evaluated with a dataset of various structures extracted from clinical CT scans. Our experiment with 61 cases resulted in the classification rate of 93.3%, demonstrating the advantage of the proposed methods.

1 Introduction

Classification of local 3D structures is an important problem for medical image analysis, in which a local 3D anatomical/pathological structure, embedded in a dense volumetric image, is categorized into a set of archetypes, such as blob (e.g., nodule), surface (e.g., polyp), tube (e.g., vessel), and tube-branch (e.g., airway junction). Such structure type information is crucial for various medical diagnostic and planning procedures. For example, for cancer screening, computer-aided detection of lung nodules [1] and colon polyps [2] can exploit such information for differentiating the target structures from common false positives (e.g., blob-like nodule among vessel branches or wall-attached polyp among colonic folds). Other examples include tube-branch detection for bronchoscopic surgery planning [3], for breast cancer screening with ductal tree analysis [4], and for internal bleeding detection in angiography [5].

Technically, 3D structure classification is challenging because of the high variability for structure's topology, size, orientation and co-articulation. There are a number of related studies which have previously been reported in literature. The model-based approach fits a set of flexible intensity models of specific structure type to data directly [6, 7], however it tends to be inefficient in a 3D domain since the high structural variability makes the model parameter space prohibitively large. The eigen-analysis approach characterizes the structure type by analyzing eigenvalues of Hessian matrix, applied successfully to vasculature and bronchial analysis [8, 9]. Although this approach itself cannot identify a branching point as a part of vessels, hybrid methods (e.g., [10]) have successfully been applied for the branch cases as well. Despite these reported efforts and the problem's clinical significance, there are no unifying methods that can address all four archetypes of 3D local structure, to our best knowledge.

Addressing this issue, we propose two novel Bayesian probabilistic algorithms built on top of our previous work [11]. As illustrated in Fig.1, this approach transforms the 3D structure classification problem into a 2D clustering problem by a) unwrapping a 3D spherical surface, encompassing a target structure, into a 2D

Proc. MICCAI Workshop on Machine Learning in Medical Imaging, paper 03-1-8, 2010.

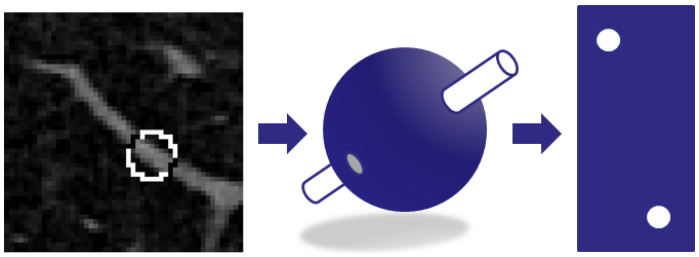


Fig. 1. Transformation of 3D structure classification to 2D clustering problem. Input 3D structure is a vessel and output is a 2D image with 2 isolated clusters.

manifold and b) counting the number of high-intensity clusters created by the target's parts protruding from the surface thus related directly to the structure type. As advantage, this approach reduces a prohibitively large number of discretized geometrical and topological variations in 3D into a mere shifts in size and location of clusters that can be readily handled by standard techniques. In this paper we propose a Bayesian version of this procedure, as well as a true Bayesian algorithm that fully exploits all available information via statistical modeling. The main contributions of the proposed methods are 1) theoretically sound Bayesian decision making for all four structure types, 2) practically useful confidence values, and 3) high-accuracy due to the new algorithm. The following describes the proposed methods and their experimental evaluation using a dataset of 61 cases extracted from clinical chest CT scans.

2 Bayesian Framework

Let Y, C, R be discrete random variables representing three key factors to be examined in our methods. $Y : \mathbb{R}^3 \rightarrow \mathbb{R}_+$ stands for a volume of interest (VOI) containing a target structure. This VOI, as a cubic bounding box ($s \times s \times s$), is extracted semi-automatically from an input 3D medical image by adopting a robust Gaussian fitting method proposed in [12]. $C \in \{C_j | j = 0, .., 3\}$ denotes class labels for the four 3D structure types. Throughout this paper, we let C_0, C_1, C_2, and C_3 indicate blob (nodule), surface (wall-attached nodule), tube (vessel) and tube-branch (vessel-branch), respectively. $R \in \{R_k | k = 0, .., K-1\}$ represents the radius of a 3D bounding sphere encompassing the target as shown in Fig.1. R can take a value from a set of monotonically increasing radii that is bounded by $\frac{s}{2}$. In this study, we set $K = 15$. The intensity distribution on a bounding sphere at R_k in 3D domain can be represented in a 2D unwrapped image by 1) transforming each voxel's Cartesian coordinate (x, y, z) to a spherical coordinate (θ, ϕ, ρ) where $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, and $\rho \in \mathbb{R}_+$ and 2) creating a 2D image with bicubic interpolation on a regular integer lattice at $\rho = R_k$.

2.1 Bayesian Classification Framework for the Original Algorithm

The original algorithm of [11] has two successive steps: 1) estimation of the best radius and 2) structure classification given the best radius. Both steps can be formulated as maximum-a-posteriori (MAP) estimations with conditional probability distributions $P(R|Y)$ and $P(C|Y, R)$, respectively. In order to model the posterior $P(C|Y, R)$ from data statistics, we derive an equivalent form using

the Bayes rule together with an assumption that the class label C and the radius R are independent.

$$k^* = argmax_k P(R_k \mid Y) \tag{1}$$
$$j^* = argmax_j P(C_j \mid Y, R_{best}) \tag{2}$$
$$= argmax_j P(Y \mid C_j, R_{best}) P(C_j) \tag{3}$$

where $R_{best} = R_{k^*}$ denotes the best radius estimate given the input VOI Y, the Bayes optimal classification (under the 0-1 loss) is given by $C_{best} = C_{j^*}$, and $P(C_{best}|Y, R_{best})$ denotes *confidence* of the decision. A 3D bounding sphere defined by R_{best} is unwrapped then used as an input to the step of (3). We call this probabilistic classification framework *MAP-I*.

2.2 True Bayesian Framework

Note that the estimation of the bounding sphere's radius in (1) is not central to the goal of our classification problem. Thus another approach is to treat R as a hidden variable and formulate a single-step estimator that maximizes a marginal distribution $P(C|Y)$. Similar to MAP-I, with Bayes and sum rules, we derive an expression with distributions that can be estimated from data more easily. By assuming the independence of C and R, we have

$$j^* = argmax_j P(C_j \mid Y) \tag{4}$$
$$= argmax_j \sum_{k=0}^{K-1} P(Y \mid C_j, R_k) P(R_k) P(C_j) \tag{5}$$

where $C_{best} = C_{j^*}$ denotes the Bayes optimal classification and $P(C_{best}|Y)$ its decision confidence. We call this new MAP estimation framework *MAP-II*.

Intuitively, this framework removes the extra radius estimation step in MAP-I by considering all possible radii of the bounding sphere. This approach can be more accurate than MAP-I by avoiding inevitable errors from the extra estimation step, as well as by utilizing the maximum amount of observation, available within the given input data, directly toward classification decision making.

3 Statistical Modeling

This section describes our statistical modeling methods for estimating, from image appearance Y, a pair of distributions $P(R)$ and $P(Y|C, R)$ appeared in (3) and (5). For this study, we model the class label prior unbiasedly by $P(C) = 1/4$.

3.1 Radius Prior $P(R)$

The prior distribution of radius $P(R)$ is modeled in a bootstrap fashion from data Y. Our method first unwraps 3D bounding spheres of various radii with one degree interval into a set of 15 2D images $I = \{I_R | R = R_0, .., R_{K-1}\}$ of size $(n = 180) \times (m = 360)$ and then treats each image as a 2D probability distribution $\widetilde{I}_{R_k}$ by normalizing the intensity distribution such that $\sum_x \sum_y \widetilde{I}_{R_k}(x, y) = 1$. Shannon entropy $E(k)$ is then computed from each 2D distribution $\widetilde{I}_{R_k}$. $P(R = R_k)$ is finally modeled as an exponential function of $E(k)$ as follows,

$$P(R = R_k) = e^{-E(k)} \tag{6}$$

The smaller entropy $E(k)$ indicates clearer appearance of clusters representing the protruding target parts. Thus $P(R = R_k)$ encodes how it is easy to differentiate high intensity clusters from the background for a given R_k. Solving (1) with the above-modeled $P(R = R_k)$ yields the optimal 2D unwrapped image $I^* = M_{R_{best}}$ extracted from Y at the MAP-I radius R_{best}.

3.2 Data Likelihood $P(Y|C, R)$

We model the conditional distribution of image appearance Y over class labels C_j and radii R_k with three conditionally independent image appearance features (M, S, F). The likelihood function is then expressed by a product of following three distributions,

$$P(Y|C_j, R_k) = P(M|C_j, R_k)P(S|C_j, R_k)P(F|C_j, R_k) \tag{7}$$

where M denotes a feature based on residual error of an intensity distribution model fitted to Y; S represents another feature representing size estimation error with respect to pre-learned class-specific structure size; and F indicates a feature indicating the consistency of classification results over varying internal parameters of clustering analysis.

Clustering Analysis In order to extract image features described above, the 2D unwrapped image I_R is first subjected to a clustering analysis. The shape of clusters in I can be irregular because our target structures themselves often have irregular shapes. Addressing this issue, we adopt non-parametric mean shift (MS) clustering [13]. This method also is appealing to our application since the number of clusters are estimated from data rather than given *a priori*.

The image I_R is first subjected to thresholding with a value predefined by a pilot study. We then randomly sample 1000 data points from the regions of high intensity: $\mathbb{S} = \{\mathbf{x}_0, \cdots, \mathbf{x}_{999}\}$ where $\mathbf{x} = (x, y)$ and $0 \leq x \leq \pi$ and $0 \leq y \leq 2\pi$ $\forall \mathbf{x} \in \mathbb{S}$. This clustering analysis yields the following two outcomes: 1) *number of clusters* N in I_R and 2) *cluster membership* $(\mathcal{X}_0, .., \mathcal{X}_i, .., \mathcal{X}_{N-1})$ denoting a set of sample points that belong to i-th cluster and representing a partition of $\mathbb{S}$.

This clustering analysis is repeated over varying kernel bandwidth h, yielding *different* numbers of clusters for the same input. We use a normal kernel and a set of ten bandwidths uniformly sampled from a range of [0.3, 0.5] radians, $\mathcal{H} = \{h_0, \cdots, h_9\}$. Then we denote the numbers of clusters for the bandwidth-varying analyses by $N_{\mathcal{H}} = \{N_{h0}, .., N_{h9}\}$ and sets of cluster memberships by $\mathcal{Y}_{\mathcal{H}} = \{\mathcal{Y}_0 = \{\mathcal{X}_{0,0}, .., \mathcal{X}_{0,N_{h1}-1}\}, .., \mathcal{Y}_9 = \{\mathcal{X}_{9,0}, .., \mathcal{X}_{9,N_{h9}-1}\}\}$.

Residual Error of Intensity Model Fitting (M) A Gaussian mixture model (GMM) with N_j components is used to model the intensity distribution of each 2D unwrapped image I_R. For the four structure types of blob, surface, tube, and tube-branch, we set $N_0 = 0$, $N_1 = 1$, $N_2 = 2$, $N_3 = 3$, respectively. For blobs, we use a uniform distribution because no cluster is expected in the image,

$$g_{X|j}(\mathbf{x}) = \begin{cases} \frac{1}{n \times m} & j = 0 \\ \frac{1}{N_j} \sum_{i=0}^{N_j - 1} f_{\mathcal{X}_i}(\mathbf{x}), & j = 1, 2, 3 \end{cases} \tag{8}$$

This GMM model is fitted to I_R for each structure type j by estimating each component Gaussian $f_{\mathcal{X}_i}$ from corresponding sample set $\mathcal{X}_i$ by $f_{\mathcal{X}_i}(\mathbf{x}) = |2\pi\Sigma_i|^{-\frac{1}{2}}$

$\exp[-\frac{1}{2}(\mathbf{x}-\mu_{\mathbf{i}})^T \Sigma_i^{-1}(\mathbf{x}-\mu_{\mathbf{i}})]$, where Σ_i is a 2×2 covariance matrix of $\mathcal{X}_i$ and μ_i is the corresponding mean. For each class assumption j, we choose the median among all bandwidths that resulted in j clusters. Due to the phase-wraparound nature of the (θ, ϕ)-domain., we adopt directional statistics [14] as follows,

$$\mathbf{x}-\mu_{\mathbf{i}} = \begin{cases} -\frac{\pi}{2} + modulo((\mathbf{x}_x - \mu_{\mathbf{i}_x}), \pi) \\ -\pi + modulo((\mathbf{x}_y - \mu_{\mathbf{i}_y}), 2\pi) \end{cases} \tag{9}$$

A residual error $err_M(j)$ of fitting this intensity model is defined as,

$$err_M(j) = \sum_{x=0}^{n-1}\sum_{y=0}^{m-1} \left| g_{X|j}(x,y) - \widetilde{I}(x,y) \right| \tag{10}$$

where $\widetilde{I}$ denotes an intensity-normalized image.

Finally a conditional distribution $P(M|C_j, R_k)$ is modeled by the following exponential function of the derived residual error $err_M(j)$,

$$P(M \mid C_j, R_k) = \begin{cases} e^{-2} & \text{if } j \notin N_{\mathcal{H}} \\ e^{-err_M(j)} & \text{else} \end{cases} \tag{11}$$

When the MS clustering does not converge to the specific number of clusters for all evaluated bandwidth $\mathcal{H}$, we set the corresponding probability with the model fitting error value of 2 that is the error's upper-bound.

Error of Target Size Estimation (S) The size of a target structure offers another useful evidence for predicting the structure type. We estimate the target size by using the one-click Gaussian fitting method described in Sec. 2. This method fits a 3D ellipsoid to the input VOI's high intensity region robustly. We use the minimum radius of the fitted ellipsoid S_I as our size estimate. The size estimation error $err_S(j)$ is then defined as the normalized difference between this estimate and the pre-learned class-specific structure size S_j,

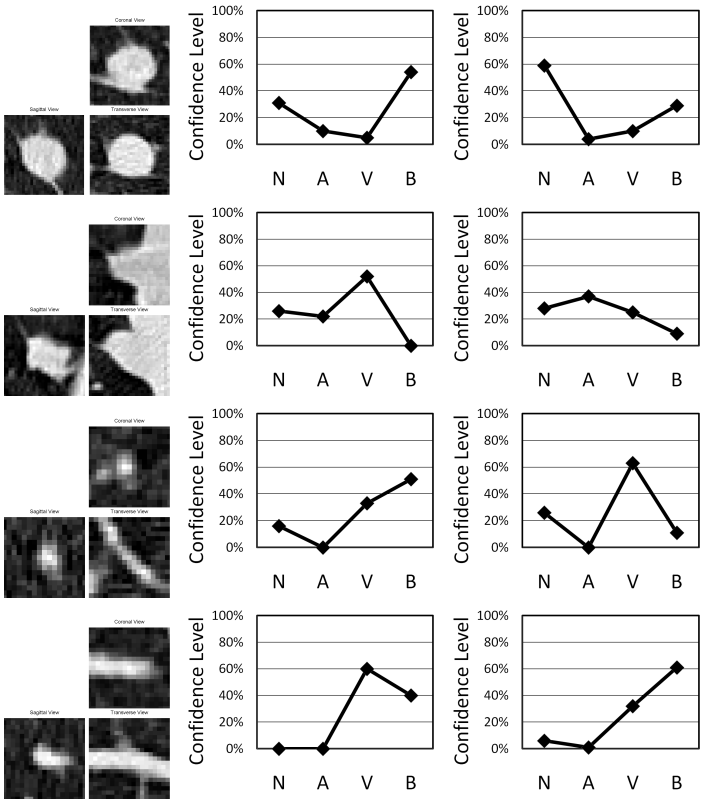
$$err_S(j) = \frac{|S_I - S_j|}{S_j} \tag{12}$$

For each structure type, we pre-learned the class-specific size by averaging the minimum radii in our pilot experimental data set for the chest CT case, resulting 8.6, 8.4, 3.5, and 4 in pixels for $j = 0, 1, 2, 3$, respectively. Finally, we model the data likelihood due to the structure size as follows,

$$P(S \mid C_j, R_k) = e^{-err_S(j)} \tag{13}$$

Consistency of Cluster Number Estimation (F) The MS clustering with varying bandwidth $\mathcal{H}$ described in Sec 3.2 results in 10 different results. A histogram f_j of the estimated cluster numbers $N_{\mathcal{H}}$ indicates how consistent the MS clustering predicts certain structure type. We define the data likelihood due to this cluster number consistency as follows,

$$P(F \mid C_j, R_k) = \begin{cases} \frac{1-\sum_{j=1}^{3} P(F|C_j,R_k)\cdot P(S|C_j,R_k)}{P(S|C_0,R_k)} & j = 0 \\ \frac{f_j}{10} & j = 1, 2 \\ \sum_{j\geq 3} \frac{f_j}{10} & j = 3 \end{cases} \tag{14}$$



(a) 3D input (b) MAP-I result (c) MAP-II result

Fig. 2. Illustrative examples of MAP-I and -II. Row 1: nodule; Row 2: wall-attached nodule; Row 3: vessel; Row 4: vessel branch. Left column: MPR views. N: nodule, A: wall-attached nodule, V: vessel, B: vessel-branch.

For the branch case ($j = 3$), the normalized counts for $j \geq 3$ are accumulated because the number of clusters can be larger than or equal to 3 in this case. Note also the nodule case ($j = 0$) cannot directly be computed because the histogram count for this case is always zero ($f_{j=0} = 0$) as the MS clustering detects at least one mode. To address this issue, we introduce a benign constraint that the total probability of a product of conditional distributions for F and S sums up to one, $\sum_{j=0}^{3} P(F|C_j, R_k)P(S|C_j, R_k) = 1$. The special case $P(F \mid C_0, R)$ can then be derived as shown above in (14).

4 Experimental Results

Data used in this study is multi-slice high-resolution 3D CT scans of human chest diagnosed with lung cancer. The data consists of 512×512 axial slices with 12-bit intensity quantization. A total of 61 distinctive cases were manually collected using the method described in Sec. 2, which contain 15 nodule, 15 wall-attached nodule, 15 vessel, and 16 vessel-branch cases.

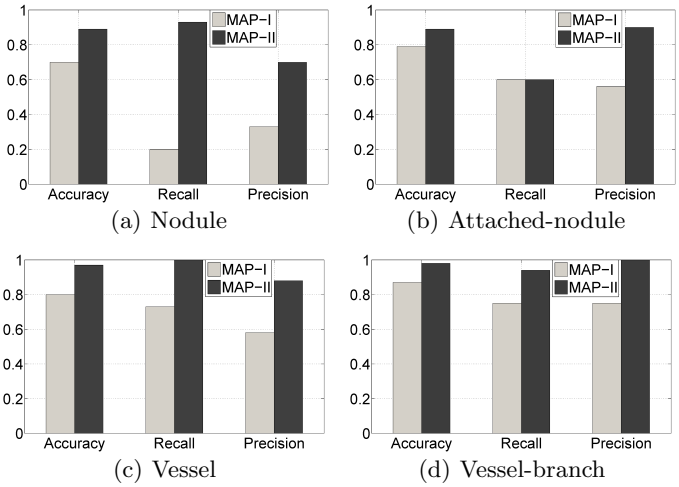


Fig. 3. Quantitative performance comparison of MAP-I and MAP-II.

Fig. 2 compares two proposed algorithms in four illustrative examples where MAP-II outperforms MAP-I for handling complex data structures with shape and size variations. In each row, classification results are displayed in confidence level plots over the four lung structure types for both frameworks. In row 1, MAP-II succeeds in classifying a nodule while MAP-I falsely classifies it as a branch. In row 2, MAP-I mistakenly classified the target as a vessel due to the highly irregular shape of the wall-attached nodule. In row 3, MAP-I failed due to the inclusion of non-target structures located nearby, while MAP-II showed its robustness against such non-targets. In row 4, MAP-I failed to capture small branches thus misclassified the target as a vessel, however, MAP-II avoids this misclassification by considering more observation available.

Fig. 3 summarizes our quantitative performance evaluation with confusion matrices. For each type, the results are categorized into true positive (TP), false positive (FP), true negative (TN) and false negative (FN) cases. We then compute the standard performance measures of accuracy, recall and precision. These results indicate that MAP-II framework outperforms the MAP-I in all the statistics described above. On average, the success rate of MAP-I is 79.0% and that of MAP-II is 93.3%, resulting in 14% of increase in accuracy.

5 Conclusion

This paper proposed two novel Bayesian algorithms for classifying local 3D structures in medical scans. Our Bayesian approach unifies the classification task of various structure types into a theoretically sound framework with high accuracy and useful confidence information. Our experimental results demonstrated high accuracy (93.3%) and suggests a potential improvement for general chest CAD applications by using our method as a post-filter to remove false positive cases. Note that the overall design is not restricted to this chest CAD application and would be useful in various 3D medical images in general. As future work, we plan to extend our experimental study to much larger dataset including other applications and modalities such as internal bleeding detection with

various angiographic data. Furthermore, our study suggests that an inclusion of more evidence, shown in (7) in our Bayesian decision making, leads to better overall accuracy. To this end, we may consider introducing more measurements such as structural orientation, cluster size and structure density.

References

1. Armato, S.G.I., Giger, M.L., Moran, C.J., Blackburn, J.T., Doi, K., MacMahon, H.: Computerized detection of pulmonary nodules on CT scans. RadioGraphics **19** (1999) 130–131
2. Summers, R., Beaulieu, C., Pusanik, L., Malley, J., Jeffrey, J.R.B., Glazer, D., Napel, S.: Automated polyp detector for CT colonography: feasibility study. Radiology **216** (2000) 284–290
3. Kiraly, A.P., Helferty, J.P., Hoffman, E.A., McLennan, G., Higgins, W.E.: Three-dimensional path planning for virtual bronchoscopy. IEEE Trans. Medical Imaging **23**(11) (2004) 1365–1379
4. Nuzhnaya, T., Barnathan, M., Ling, H., Megalooikonomou, V., Bakic, P., Maidment, A.: Probabilistic branching node detection using adaboost and hybrid local features. In: Proc. IEEE Int. Symp. on Biomedical Imaging. (2010)
5. Yoon, W., Jeong, Y., Shin, S., Lim, H., Song, S., Jang, N., Kim, J., Kang, H.: Acute massive gastrointestinal bleeding: detection and localization with arterial phase multi-detector row helical CT. Radiology **239**(1) (2006) 160–167
6. DeCarlo, D., Metaxas, D.: Shape evolution with structural and topological changes using blending. IEEE Trans. Pattern Anal. and Machine Intell. **20**(11) (1998) 1186–1205
7. Sato, Y., Westin, C., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikinis, R.: Tissue classification based on 3D local intensity structures for volume rendering. IEEE Trans. Visualization and Computer Graphics **6** (2000)
8. Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., Kikinis, R.: Segmentation and visualization of curvilinear structures in medical images. Medical Image Analysis **2** (1998) 143–168
9. Frangi, A.F., Niessen, W.J., Hoogeveen, R.M., van Walsum, T., Viergever, M.A.: Model-based quantitation of 3D magnetic resonance angiographic images. IEEE Trans. Medical Imaging **18** (1999) 946–956
10. Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Trousset, Y.: Model based detection of tubular structures in 3d images. Computer Vision and Image Understanding **80** (2000) 130–171
11. Bahlmann, C., Li, X., Okada, K.: Local pulmonary structure classification for computer-aided nodule detection. In: Proc. SPIE Conf. Medical Imaging. Volume 6144., International Society for Optical Engineering (2006) 1–5
12. Okada, K., Comaniciu, D., Krishnan, A.: Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT. IEEE Trans. Medical Imaging **24(3)** (2005) 409–423
13. Comaniciu, D., Meer, P.: Mean shift: A Robust Approach Toward Feature Space Analysis. IEEE Trans. Pattern Anal. and Machine Intell. **24**(5) (2002) 603–619
14. Mardia, K.V., Jupp, P.E.: Directional Statistics. Wiley (1999)