

# Gaussian Bayes Classifier for 2D Shapes in Kendall Space

Hibat Allah Rouahi<sup>1(✉)</sup>, Riadh Mtibaa<sup>1,2(✉)</sup>, and Ezzeddine Zagrouba<sup>1(✉)</sup>

<sup>1</sup> Laboratoire LIMTIC, Institut Supérieur d'Informatique,  
Université de Tunis El Manar, 2 Rue Abou Rayhane Bayrouni, 2080 Ariana, Tunisie  
hiba\_rouahi@yahoo.fr, rmtibaa@gmail.com,  
ezzeddine.zagrouba@fsm.rnu.tn

<sup>2</sup> Institut Supérieur des sciences appliquées et de technologie de Sousse (ISSAT),  
Université de Sousse, rue Khalifa Karoui Sahloul 4, 526, Sousse, Tunisie

**Abstract.** We propose a 2D-shape Gaussian Bayes classifier based upon Kendall's representations that help to quotient out the effects of non-altering shape geometric transformations. The Kendall space is a non linear space that coincides with the unit sphere modulo an isometry group. The proposed Riemannian metric is more apt in the case where shapes are different only in translation, scale and rotation. In addition to that, the manifold structure of this space renders the multivariate statistical analysis implementation unfeasible in practice. Consequently, tools such as learning and classification models are non trivial and not frequently available. To overcome these issues, we adapt the Gaussian Bayes classifier to this space. We computed the likelihood parameters through appropriate projections onto Kendall tangent space that provides a good linear approximation. In order to validate the robustness of our classifier, we proceeded to computer simulations using several benchmarks.

**Keywords:** Kendall space · Gaussian Bayes classifier · Likelihood parameters

## 1 Introduction

Objects' recognition is one of the most active areas of research with crucial applications in diverse fields. Object representation is a primordial step to systematically allocate mathematical structures to shapes in order to ease the implementation and analysis of the relevant classification algorithms. Objects may be described according to main parameters: colors, textures, shapes, movements and locations. However, shape remains a critical feature for object recognition as commonly accepted by the computer vision community. The need for consistent structures that help handle the rich repertoire of shapes has elicited the investigation of several researchers. Among the proposed representations are: continuous contour parametrization [1], medial [2], active model [3], and algebraic representations [4]. Once a particular representation is defined, a classifier

is either trained during a learning phase through probability models and pre-selected shape sets, or designed through a clustering technique. Another type of invariant shape representation has been proposed, the Fourier descriptors along with moment invariants. These are among the main techniques, yet are only based on continuous regions and contours [5]. In 1977, D.G Kendall proposed a new invariant representation where he identified a shape with the geometrical information that remains after filtering out the location, scale and rotation effects from initial configuration matrices that capture objects' shapes through landmarks. The elimination of irrelevant effects generated a non-linear space. Consequently, the application of various clustering and classification methods is not appropriate. The Riemannian metric for this space refers only to whether two shapes are identical (different only in translation, rotation and scale) or not, yet in many cases we need to measure shape similarity. An appropriate metric for shape classification should not only suit certain invariance properties but also satisfy the different input properties. This study tries to devise a robust classifier through the combination of an efficient shape representation emerging within the framework of the Kendall's theory with the Bayesian approach renowned for its rigorous theoretical foundation and optimal classification results. We based our choice of the Kendall representation upon the firm belief that its adequate implementation yields computational effectiveness. Indeed, representation-based-matrices are easy for calculate and are much less time consuming compared to those using more complex structures. Our main contribution is the adaptation of the Gaussian Bayes classifier in Kendall space and its local approximation through projections onto tangent spaces. Section 2 of this paper exposes an overview of some works related to Kendall space. The proposed method is presented in Sect. 3. Experiments are presented and discussed in Sect. 4. Finally, we conclude the work and suggest proposals in Sect. 5.

## 2 Previous Works on Kendall Space

The study of shapes dates back to D'Arcy Thompson, but the first more systematic algorithmic treatment of shape representations and metrics is due to Bookstein and Kendall. They represented shape by a collection of ordered landmark points invariant to Euclidean similarity transformations. Two objects have the same shape if they can be translated, scaled and rotated to each other so that they match exactly. This standpoint has led to the foundation of the Kendall shape theory [6, 7] which is a most popular and widely discrete used shape representation. Number of works elaborated by shape theories experts like Bookstein [8], Dryden [9], have adopted Kendall definition to get finite dimensional spaces from landmarks coordinates. The readers are referred to [10, 11] for a thorough view of the recent developments of Kendall theory. The solution to resolve the problem of non-conforming between Kendall space and the classical linear algorithm of classification was proposed by Jayasumana et al. in [12] and is to perform a mapping operation from the manifold to the Hilbert space using a kernel function. Which produces a richer representation of the data, and makes tasks

such as classification easier. However, only positive definite kernels yield a mapping to a Hilbert space and a poor choice of kernel can often result in reduced classification performance. Another idea to adapt an unsupervised learning algorithm, k-means, to Kendall space is to integrate the Procrustes mean [13]. Here we focus on adapting the Bayesian approach to Kendall space through a linear tangent space.

### 3 Our Approach of Supervised Learning in Kendall Space

#### 3.1 Landmarks Selection

At the outset we fixed the value of the number  $N_c$  of the targeted shape classes  $\omega_i$ , for  $1 \leq i \leq N_c$ , since we are working within a supervised context. Then, we generated for each  $\omega_i$  amongst the  $N_c$  classes a set  $\{X_j^{i*}\}_{1 \leq j \leq N_{ps}}$  comprising  $N_{ps}$  initial configurations of learning samples. Here, we gathered in each matrix  $X_j^{i*}$  the vector column coordinates of some fixed number of labelled landmarks, which are selected to capture the shape of the  $j$ th sample from the class  $\omega_i$ . More accurately, the selection consists of picking up  $k$  landmarks from the contours. We noticed that some of the benchmarks also provide a set of landmarks selected by experts, unless said landmarks were not valid, in that case we performed the selection task manually based on templates assigned to each class.

#### 3.2 Learning Shapes Processing

Let  $X_j^{i*} = (x_0^* \ x_1^* \ \dots \ x_{k-1}^*)$  be one of the learning samples in  $\{X_j^{i*}\}_{1 \leq j \leq N_{ps}}$  for the class  $\omega_i$  (Cf. Sect. 3.1). We started by calculating the center of  $X_j^{i*}$  to remove the translation effect from  $X_j^{i*}$ , by moving  $x_c^*$  to the origin of the coordinates system, when each  $x_\ell^*$  is substituted with  $x_\ell^* - x_c^*$ , for all  $0 \leq \ell \leq k-1$ , respectively.

$$x_c^* = \frac{1}{k} \sum_{\ell=0}^{k-1} x_\ell^* \quad (1)$$

Next, we performed a dimension reduction via a right multiplication of the last version of  $X_j^{i*}$  by the recentering orthogonal matrix  $Q$ ,

$$Q = \begin{cases} Q_{\ell 1} = \frac{1}{\sqrt{k}}, & 1 \leq \ell \leq k; \\ Q_{\ell \ell} = \frac{\ell-1}{\sqrt{\ell(\ell-1)}}, & 2 \leq \ell \leq k; \\ Q_{\ell h} = -\frac{1}{\sqrt{\ell(\ell-1)}}, & 1 \leq \ell \leq h-1, 2 \leq h \leq k; \\ Q_{\ell h} = 0, & \text{otherwise.} \end{cases} \quad (2)$$

to get an intermediate representation matrix  $\tilde{X}_j^i$  with general form  $(0 \ \tilde{x}_1 \ \tilde{x}_2 \ \dots \ \tilde{x}_{k-1})$  which we reduced naturally to

$$\tilde{X}_j^i = (\tilde{x}_1 \ \tilde{x}_2 \ \dots \ \tilde{x}_{k-1}) \quad (3)$$

After that, we eliminated the scaling effect through normalization to get

$$X_j^i = \frac{1}{\sqrt{\text{tr}(\tilde{X}_j^i(\tilde{X}_j^i)^t)}} \tag{4}$$

The set  $\{X_j^i\}_{1 \leq j \leq N_{ps}}$  coincides with points on the  $2(k-1)-1$  dimensional unit sphere, denoted by  $S_2^k$ . We notice that any matrix  $TX_j^i$ , for  $T \in \mathbf{SO}(2)$ , has the same shape as  $X_j^i$ . So, from now on,  $X_j^i$  is treated as a pre-shape representation and the sought shape denoted by  $\pi(X_j^i)$  identifies an equivalence class modulo the left action of rotations  $T$  in  $\mathbf{SO}(2)$  on the pre-shape  $X_j^i$ ; the space of all possible shapes is the  $2(k-1)-2$  dimensional Kendall space, which is the quotient space.

$$\Sigma_2^k = S_2^k / \mathbf{SO}(2) \tag{5}$$

The pseudo-singular values decomposition helps to write any pre-shape  $X_j^i$  as a three-factors product,

$$X_j^i = U(\Lambda 0)V \tag{6}$$

where  $U \in \mathbf{SO}(2)$ ,  $V \in \mathbf{SO}(k-1)$ ,  $0$  is the null matrix of dimensions  $2 \times (k-3)$ , and  $\Lambda$  is the  $2 \times 2$  diagonal matrix  $\text{diag}\{\lambda_1, \lambda_2\}$  such that  $\lambda_1 \geq |\lambda_2|$ ,  $\lambda_1^2 + \lambda_2^2 = 1$ , and  $\lambda_2 \geq 0$  unless  $k = 3$ . This decomposition provides a systematic way to decide whether or not any couple of learning pre-shapes belong to the same equivalence class. In order to quotient out the left acting orthogonal matrices to obtain the learning shapes in  $\Sigma_2^k$  (5), we calculated  $U_{X_{j_1}^i}(\Lambda_{X_{j_1}^i} 0)V_{X_{j_1}^i}$  and  $U_{X_{j_2}^i}(\Lambda_{X_{j_2}^i} 0)V_{X_{j_2}^i}$  of each couple of pre-shapes  $X_{j_1}^i$  and  $X_{j_2}^i$  in  $\{X_j^i\}_{1 \leq j \leq N_{ps}}$  (6). Then, we decided that  $\pi(X_{j_1}^i)$  and  $\pi(X_{j_2}^i)$  are identical if and only if  $\Lambda_{X_{j_1}^i} = \Lambda_{X_{j_2}^i}$  and both first rows of  $V_{X_{j_1}^i}$  and  $V_{X_{j_2}^i}$  are exactly the same, the remaining  $k-3$  rows of  $V_{X_{j_1}^i}$  and  $V_{X_{j_2}^i}$  do not matter since they are multiplied by the null matrix of dimensions  $2 \times (k-3)$  appearing in  $(\Lambda_{X_{j_1}^i} 0)$  and  $(\Lambda_{X_{j_2}^i} 0)$ , respectively. Naturally, we did not care about the left acting orthogonal matrices  $U_{X_{j_1}^i}$  and  $U_{X_{j_2}^i}$  because they do not affect shapes. This way, we succeeded in regrouping the  $N_{ps}$  learning pre-shapes of  $\{X_j^i\}_{1 \leq j \leq N_{ps}}$  into  $N_s^i$  learning equivalence classes. Concretely, for all  $1 \leq p \leq N_s^i$ , the  $p$ th equivalence class is represented by any candidate denoted by  $\pi(X_p^i)$ , and qualified as the learning shape.

### 3.3 Likelihood for Parameters Inference

We describe here, our approach of learning procedure to estimate the expectation vector  $\mu_i$  as well as the covariance matrix  $\Sigma_i$  of the Gaussian likelihood. We assume that the learning samples in each set  $\{\pi(X_p^i)\}_{1 \leq p \leq N_s^i}$  are independent and identically distributed according to the likelihood of each  $\omega_i$ , respectively. In a first step, we calculated the mean shape  $\pi(\hat{\nu}_i)$  of the set of learning shapes

$\{\pi(X_p^i)\}_{1 \leq p \leq N_s^i}$  which we used later as a reference shape where we approximated  $\Sigma_2^k$  locally by its tangent space [14]. Specifically, we looked for  $\pi(\hat{\nu}_i)$  as a solution of the minimization problem

$$\arg \inf_{\pi(\nu) \in \Sigma_2^k} \sum_{p=1}^{N_s^i} d_F^2(\pi(X_p^i), \pi(\nu)) \quad (7)$$

which involves the procrustes distance between  $\pi(X_p^i)$  and  $\pi(\nu)$

$$d_F^2(\pi(X_p^i), \pi(\nu)) = \sin(\rho(\pi(X_p^i), \pi(\nu))) \quad (8)$$

Here,  $\rho$  is the distance function defined for  $\Sigma_2^k$  as

$$\rho(\pi(X_p^i), \pi(\nu)) = \arccos(\lambda_1 + \lambda_2) \quad (9)$$

with the pseudo-singular values  $\lambda_1$  and  $\lambda_2$  of  $X_p^i \nu^t$  for arbitrary pre-shapes  $X_p^i$  and  $\nu$  of  $\pi(X_p^i)$  and  $\pi(\nu)$ , respectively. In a second step, we mapped each learning shape  $\pi(X_p^i)$  in  $\{\pi(X_p^i)\}_{1 \leq p \leq N_s^i}$  onto its projection  $\bar{\pi}_i(X_p^i)$  computed as

$$\bar{\pi}_i(X_p^i) = \left( I_m - \pi(\hat{\nu}_i) \pi(\hat{\nu}_i)^t \right) \pi(X_p^i), \quad (10)$$

In practice, we computed  $\pi(\hat{\nu}_i)$  and  $\bar{\pi}_i(X_p^i)$  for all  $1 \leq i \leq N_c$  and  $1 \leq p \leq N_s^i$  using the generalised procrustes analysis [15]. We emphasize here on two important features of the last function: first, we get always the same mean shape  $\pi(\hat{\nu}_i)$  even if we use left rotated versions of the actual shapes  $\pi(X_p^i)$ , for all  $1 \leq i \leq N_c$ , respectively. Second, all the left rotated versions of the shape  $\pi(X_p^i)$  project always to the same point of the tangent space to  $\Sigma_2^k$  at  $\pi(\hat{\nu}_i)$ . In a last step, we reshaped each one of the matrices  $\bar{\pi}_i(X_p^i)$  onto a row vector  $\bar{\pi}_i^v(X_p^i)$ . Then, we profited from the maximum likelihood method to get likelihood parameters

$$\mu_i = \frac{1}{N_s^i} \sum_{p=1}^{N_s^i} \bar{\pi}_i^v(X_p^i), \quad (11)$$

$$\Sigma_i = \frac{1}{N_s^i} \sum_{p=1}^{N_s^i} (\bar{\pi}_i^v(X_p^i) - \mu_i)^t (\bar{\pi}_i^v(X_p^i) - \mu_i), \quad (12)$$

Both parameters dimensions are  $1 \times 2(k-1)$  and  $2(k-1) \times 2(k-1)$ , respectively.

### 3.4 Generalization

In the present subsection we detail how we conducted the classification task of general shapes. After obtaining the likelihood phase (Cf. Sect. 3.3), we used the

values of the Gaussian likelihood  $\mu_i$  and  $\Sigma_i$  of each class  $\omega_i$  to compute the posteriori probabilities

$$p(\omega_i|\pi(X)) = \frac{p(\bar{\pi}_i^v(X)|\omega_i)P(\omega_i)}{p(\bar{\pi}_i^v(X))} \quad (13)$$

where

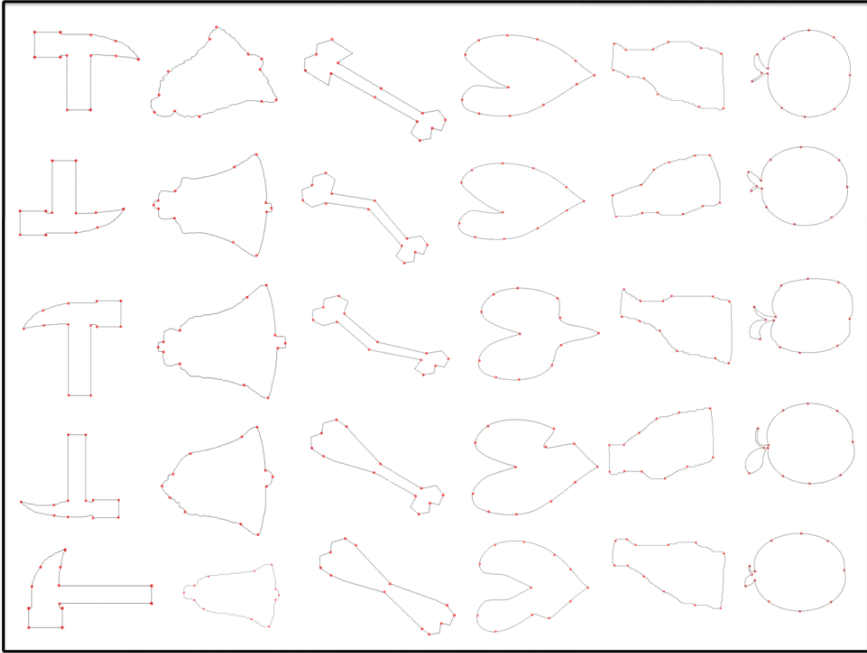
$$p(\bar{\pi}_i^v(X)|\omega_i) = \frac{1}{(2\pi)^{(k-1)}\sqrt{\det(\Sigma_i)}} \exp\left(-\frac{1}{2}(\bar{\pi}_i^v(X) - \mu_i)\Sigma_i^{-1}(\bar{\pi}_i^v(X) - \mu_i)^t\right), \quad (14)$$

$P(\omega_i)$  is an a priori probability of  $\omega_i$ , and  $p(\bar{\pi}_i^v(X))$  is the evidence term. Finally, the maximum amongst the values of  $p(\omega_i|\pi(X))$  indicates the class of  $\pi(X)$ . Here again, we draw the attention of the readers that if we represent the shape  $\pi(X)$  by any  $TX$  where  $T \in \mathbf{SO}(2)$ , then we still have exactly the same likelihood value in (14), because all such  $TX$  matrices project always through (10) to the same  $\bar{\pi}_i^v(X)$  in the tangent space to  $\Sigma_2^k$  at  $\pi(\hat{\nu}_i)$ ; that is to remind that our a posteriori probability distribution does depend on shapes rather than pre-shapes.

## 4 Experiments and Results Analysis

We ran several experiments for the purpose of evaluating the behavior of our classifier on  $\Sigma_2^k$  (see Sect. 3 above). We used mainly four 2D shape benchmarks namely, MPEG-7, Swedish leaves [14], great apes data, and T2 mouse vertebrae data [15]. Our classifier faced an important challenge since the handled shapes correspond to domestic objects, leaves, skulls or even mouse vertebrae. The MPEG-7 offers 70 classes and 20 images per class, yet in our experiments we only used 60 classes. Figure 1 contains some annotated samples from MPEG-7. We assigned 10 images per class to the learning phase, and the rest for generalization. We will now demonstrate the benefits of our Classifier on the problem of leaf identification. We employed the Swedish leaves dataset, which contains 15 different classes with 75 images per class. We used 35 images for training and the rest for testing. Some annotated samples are shown in Fig. 2. The apes database contains the skulls of 167 specimens of great apes, with different species and both sexes: chimpanzee (26 females and 28 males), gorilla (30 females and 29 males) and orangutan (30 females and 30 males). We used 15 samples of each class for the learning phase, the rest were used for the test. Finally, the T2 mouse vertebrae database contains the second thoracic vertebra of three groups of mice: large (23 samples), small (23 samples) and control group (30 samples). We used 13 samples for each group for the learning stage and the rest for generalization.

In Table 1 we summarize the outcomes of the experiments on the aforementioned benchmarks, where the numbers of landmarks are 12 and 27 for MPEG-7 and Swedish leaves respectively, which we selected since they are not available as shown in the table. For the case of apes and T2 mouse vertebrae, the landmarks

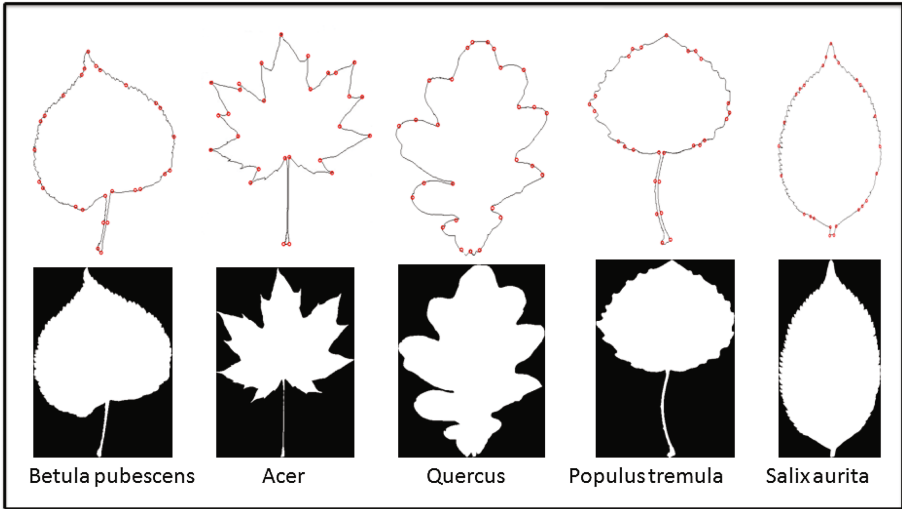


**Fig. 1.** Samples of hammer, bell, bone, heart, bottle, and apple shapes from MPEG-7. The red dots indicate the selected landmarks for each shape sample. (Color figure online)

have been provided by an expert. We end the current subsection by comparing the results of our classifier in  $\Sigma_2^k$  to those of the Gaussian Bayes classifier in Euclidean space and to the kernel SVM classifiers in Hilbert space. It should be noted that the result of the last classifier is derived from this paper [12]. We gathered the results of generalization of all these classifiers in Table 2 where it is clear that our classifier outperforms the classical one, proof of its robustness.

We expect that the improvement in classification results comes from the nonexistence of redundancy within the learning sets  $\{\pi(X_p^i)\}_{1 \leq p \leq N_s^i}$  used by our classifier in Kendall space, compared to the sets  $\{X_j^{i*}\}_{1 \leq j \leq N_{ps}}$  used by the classical classifier. Here, the nonexistence of the redundancy stems straightforwardly from the systematic elimination of translation, scale, and rotation effects during the construction of learning shapes. Besides, the complexity of our classifier in Kendall space amounts to an order determined mainly by the

$2(k-1) \sum_{i=1}^{N_c} N_s^i$  coordinates values involved in learning phase, which should be much smaller than the  $2kN_cN_{ps}$  coordinates values involved in the learning phase of the classical classifier since  $N_s^i \ll N_{ps}$ , for all  $1 \leq i \leq N_c$ , when the initial learning data sets get larger.



**Fig. 2.** Swedish leaves dataset samples.

**Table 1.** Computer simulations summary

Benchmark	Landmarks	Availability	Success rate
MPEG-7	12	No	97.16%
Swedish leaves	30	No	91.16%
Apes	8	Yes	98.70%
T2 vertebrae	6	Yes	97.29%

**Table 2.** Comparison between our classifier and two other supervised classifiers

Benchmark	Gaus-Bay in Ken	Gaus-Bay in Eucl	SVM in Hilb
MPEG-7	97.16%	86.33%	96.57%
Swedish leaves	91.16%	90.37%	91.47%
Apes	98.70%	90.90%	—
T2 vertebrae	97.29%	78.37%	—

## 5 Conclusion and Future Works

In the present paper we proposed an approach of supervised learning for shape classification in  $\Sigma_2^k$ . The supervised learning has concerned the inference of the expectation vectors and covariance matrices of the Gaussian distributions. The inference here has been realized in the context of maximum likelihood, using the available learning shapes. We detailed the procedure that helped us to construct a learning shape set from initial configurations. Then, we used the Riemannian



structure to specify the mean shape where the  $\Sigma_2^k$  is approximated by its tangent space. Consequently, we succeeded in establishing a likelihood density function on  $\Sigma_2^k$ . The results of the experiments has confirmed the robustness of our model where the success rates are outstanding. The major drawback of our approach comes from the landmarks selection which is at best a semi-automatic process where the selection remains subjective because the human supervision is necessary. In our future work, we propose to use a deformable templates for automatic detection of landmarks. We will also further the analysis of the abilities of our classifier through considering noisy or corrupt samples.

## References

1. Srivastava, A., Liu, X., Mio, W., Klassen, E.: A computational geometric approach to shape analysis in images. In: *Advances in Neural Information Processing Systems*, pp. 1579–1586 (2003)
2. Siddiqi, K., Pizer, S. (eds.): *Medial Representations: Mathematics, Algorithms and Applications*, vol. 37. Springer Science & Business Media, Netherlands (2008)
3. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. *Comput. Vis. Image Underst.* **61**(1), 38–59 (1995)
4. Subrahmonia, J., Cooper, D.B., Keren, D.: Practical reliable Bayesian recognition of 2D and 3D objects using implicit polynomials and algebraic invariants. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(5), 505–519 (1996)
5. Zhang, J., Zhang, X., Krim, H., Walter, G.G.: Object representation and recognition in shape spaces. *Pattern Recogn.* **36**(5), 1143–1154 (2003)
6. Kendall, D.G., Barden, D., Carne, T.K., Le, H.: *Shape and Shape Theory*, vol. 500. Wiley, Chichester (1999)
7. Kendall, D.G.: Shape manifolds, procrustean metrics and complex projective spaces. *Bull. Lond. Math. Soc.* **16**(2), 81–121 (1984)
8. Bookstein, F.L.: Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Image Anal.* **1**(3), 225–243 (1997)
9. Dryden, I.L., Mardia, K.V.: *Statistical Shape Analysis*, vol. 4. Wiley, Chichester (1998)
10. Amor, B.B., Su, J., Srivastava, A.: Action recognition using rate-invariant analysis of skeletal shape trajectories. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(1), 1–13 (2016)
11. Giebel, S.M., Schiltz, J., Schenk, J.P.: Statistical shape analysis for the classification of renal tumors affecting children. *Pak. J. Statist.* **29**(1), 129–138 (2013)
12. Jayasumana, S., Salzmann, M., Li, H., Harandi, M.: A framework for shape analysis via Hilbert space embedding. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1249–1256 (2013)
13. Vinu, G., Sim, A., Alemany, S.: The k-means algorithm for 3D shapes with an application to apparel design. *Adv. Data Anal. Classif.* **10**(1), 103–132 (2014)
14. Soderkvist, O.: *Computer vision classification of leaves from swedish trees*. Masters thesis, Linköping University (2001)
15. Dryden, I.L.: *Shapes package*. R Foundation for Statistical Computing, Vienna, Austria, Contributed package (2015). <https://www.maths.nottingham.ac.uk/personal/ild/shapes>