Characterising shape patterns using features derived from best-fitting ellipsoids

Amelia Gontar, Hayden Tronnolone, Benjamin J. Binder,
Murk J. Bottema

PII: S0031-3203(18)30219-X
DOI: 10.1016/j.patcog.2018.06.009
Reference: PR 6583

To appear in: *Pattern Recognition*

Received date: 16 October 2017
Revised date: 15 April 2018
Accepted date: 10 June 2018

Please cite this article as: Amelia Gontar, Hayden Tronnolone, Benjamin J. Binder, Murk J. Bottema, Characterising shape patterns using features derived from best-fitting ellipsoids, *Pattern Recognition* (2018), doi: 10.1016/j.patcog.2018.06.009

Highlights

- A method is developed to characterise highly irregular shape patterns.
- Features are defined based on best-fitting ellipsoids.
- The method is demonstrated successfully on three biomedical data sets.
- Shape patterns are learned automatically from the data.
- Landmark points are not required.

Characterising shape patterns using features derived from best-fitting ellipsoids

Amelia Gontar[a], Hayden Tronnolone[b], Benjamin J. Binder[b], Murk J. Bottema[a,*]

[a] *Flinders Mathematical Sciences Laboratory and Medical Device Research Institute, School of Computer Science, Engineering and Mathematics, Flinders University, GPO Box 2100, Adelaide SA 5001, Australia*
[b] *School of Mathematical Sciences, University of Adelaide, Adelaide, SA 5005, Australia*

Abstract

A method is developed to characterise highly irregular shape patterns, especially those appearing in biomedical settings. A collection of best-fitting ellipsoids is found using principal component analysis, and features are defined based on these ellipsoids in four different ways. The method is defined in a general setting, but is illustrated using two-dimensional images of dimorphic yeast exhibiting pseudohyphal growth, three-dimensional images of cancellous bone and three-dimensional images of marbling in beef. Classifiers successfully distinguishes between the yeast colonies with a mean classification accuracy of 0.843 (SD = 0.021), and between cancellous bone from rats in different experimental groups with a mean classification accuracy of 0.745 (SD = 0.024). A strong correlation ($R^2 = 0.797$) is found between marbling ratio and a shape feature. Key aspects of the method are that local shape patterns, including orientation, are learned automatically from the data, and the method applies to objects that irregular in shape to the point where landmark points cannot be identified between samples.

Keywords: shape analysis, dimorphic yeast, pseudohyphal growth, cancellous

*Corresponding author
Email addresses: amelia.gontar@flinders.edu.au (Amelia Gontar), hayden.tronnolone@adelaide.edu.au (Hayden Tronnolone), benjamin.binder@adelaide.edu.au (Benjamin J. Binder), murk.bottema@flinders.edu.au (Murk J. Bottema)

bone, marbling in beef

1. Introduction

1.1. Shape analysis

Shape analysis is vital in many biological applications, providing information on growth mechanisms and assisting in the detection of a variety of medical conditions [1]. In a review of shape analysis techniques, Loncaric defined the shape of an object as a binary image representing the extent of the object, and grouped shape analysis methods based on whether the method considers the shape boundary or the interior [2]. Extensive work has been done in both of these areas, and a comprehensive review has been conducted by Pavlidis [3].

Many studies have appeared in which the shape of two-dimensional (2D) irregular objects is characterised in a biological context. Elliptic Fourier analysis has been used to describe the shape of otoliths in fish [4], leaves for plant species identification [5, 6], mussel shells [7], and the human mandible [8]. Multiple resolution skeletons have been used to characterise the geometry of non-rigid objects, such as pseudopods on white blood cells [9]. Landmark methods have been used to describe the shape of the human frontal bone [10] and Old World Talpidae (mole) skulls [11]. The latter studies used projections of three-dimensional (3D) computed tomography (CT) scans onto a 2D plane, which means that not all available information was utilised.

Cootes *et al.* extended landmark methods to the idea of point distribution models [12], leading to statistical shape models (SSMs). The central idea of SSMs is to extract the mean shape and several modes of variation from a collection of training shapes using statistical methods [13]. SSMs encompass both active shape models (ASMs) and active appearance models (AAMs).

ASMs are similar to active contour models ("Snakes") [14], and for this reason are sometimes called "Smart Snakes" [12], in the sense that data from the image is used to iteratively deform the shape model in order to fit the model

to the data. However, the key step in ASMs is that constraints for allowable deformations are learned from training data [15]. ASMs have been used to model the shape of resistors, hands [12, 16], the left ventricle in echocardiograms [12, 17], prostates in magnetic resonance (MR) images, ventricles in brain MR images [17], and car brake components [18]. ASMs have also been used for classification, for example to classify plant seeds, and to recognise faces and handwritten postcodes [19]. Hill *et al.* extended ASMs to 3D using contours, and used these to segment 3D MR head images [15]. Lu *et al.* used a combination of ASMs, AAMs, and texture analysis to estimate bone fracture risk from dual-energy X-ray absorptiometry images of the human femur [20]. Although modelling the shapes is an automated process, the landmark points need to be manually located in the training images in a consistent manner. In particular, landmark points placed on one sample must correspond to equivalent landmarks on the other samples [12].

Shape context is similar to the work presented here in that a histogram of local shape descriptors is constructed [21]. This method then relies on identifying points of similarity of objects under the assumption of overall similarity. For example, silhouettes of hands in different poses may be identified by matching finger tips and the gaps between fingers.

Additional recent work on shape analysis includes a "bag of words" approach [22], a moment invariant approach for disconnected shapes [23], shape registration in the context of directional data [24], scalable methods [25], extensions of landmark methods [26], extensions of chamfer matching [27], characterising shape by modelling electric charge distributions (similar to ideas used for skeletonisation), extension of shape context [28], methods based on depicting the overall shape as a union of ellipses [29] and unsupervised learning methods [30]. All these methods apply to characterising or comparing the overall shapes of objects but are not useful for characterising local shape attributes of objects in situations where the overall shape of the object carries no significant information.

Here, the focus is on characterising the shape of objects that are irregular

in shape to the point that there are no recognisable landmarks for comparing samples. The illustrative examples presented in this paper are pseudohyphal growth exhibited by dimorphic yeast colonies [31], the structure of cancellous bone [32–34] and the structure of marbling in beef [35]. The method is fully automatic and does not require user input during training. Final shape features for characterising or classifying are learned automatically from training data.

In this study, ellipsoids are fit to local structure samples to extract various local shape characteristics. The measured properties are recorded in vectors referred to as shape primitives. The Euclidean distance between shape primitives reflects similarity in local shape patterns and clusters of shape primitives, called clustered shape primitives (CSPs), represent shape patterns that occur commonly within the object. A histogram of occurrences of these CSPs may be viewed as a quantitative descriptor of local shape patterns in the object.

The idea of fitting ellipsoids and then using CSPs to characterise shape patterns has been explored in a preliminary study [34], in which the method was trialled on the same 3D cancellous bone data that is used in the current study. In the preliminary study, a single and simplistic version of this approach was presented [34]. In the current study, various methods of recording the shape primitives are explored in depth, including whether or not to interrogate the data directly and whether to include directional information. A potential issue with scaling is addressed. In addition, the process is applied to two more problems; classifying yeast colonies based on shape and finding a connection between marbling structure and the amount of marbling in beef.

1.2. Morphology of dimorphic yeasts

Dimorphic yeasts, such as *Saccharomyces cerevisiae*, respond to nutrient deprivation by forming multicellular filaments called pseudohyphae [36]. *S. cerevisiae* is used for the production of bread, wine, and ale, making it of interest to understand and optimise growth behaviour [31]. Pseudohyphae exhibit complex shape patterns and developing methods to quantify such patterns is an open area of research.

Binder *et al.* defined a set of spatial indices to quantify the shape of the yeast colonies based on radial and angular growth, and demonstrated how the values of the indices change with time [31]. Ruusuvuori *et al.* developed a classifier that distinguished between yeast phenotypes (smooth versus fluffy) based on a large database of pre-defined features [37]. The process was automated, except for the definitions of the features, but relied on textural information and thus is not transferrable to the analysis of binary images of yeast colonies.

1.3. Structure of cancellous bone

A common coarse model for the structure of cancellous bone, especially in human bone, is that of rods and plates. Attributes such as trabecular number (number of rods), trabecular thickness and trabecular spacing provide useful descriptions of the structure. At a time when shape information for cancellous bone was only available via analysis of 2D slices across the bone, these attributes were derived using histomorphometry [38]. Rods and plates are fairly pronounced in load bearing bones of larger animals, including humans, but this pattern is much less apparent in rats – the pattern of trabecular structure in cancellous bone in rats is not easily parameterised according to consistent geometric attributes.

Martin and Bottema characterised the 3D structure of cancellous bone in rat tibiae by measuring the thickness of the bone in 13 directions at a collection of subsampled points [33]. The 13 directions in which thickness measurements were made were determined by the positioning of the tibia of the rat in the scanner, rather than the important directions being learned from the data itself. By instead fitting ellipsoids to the data, the local orientation may be determined independently of the laboratory coordinate frame.

1.4. Marbling in Beef

Flecks of fat in beef steaks are known as marbling. The number of flecks, their distribution, shape and size affect consumer demand [39], [40], [41]. Marbling is also of scientific interest because of the link between marbling and fat

metabolism. Previous work has focussed on characterising the shape of marbling flecks in the steaks, meaning that the flecks have been viewed as shapes in the plane. The shape of individual flecks were characterised in terms of area, eccentricity, and a measure of irregularity of shape in 2D [42]. However, when viewed in 3D, the marbling consists a of large connected highly irregular structure and, in consequence, the methods used to characterise the shape of individual marbling flecks in 2D do not apply analogously in 3D.

Here, local 3D shape patterns are used to determine if there is a connection between shape patterns of marbling in 3D and the amount of marbling.

1.5. Contribution an structure of the paper

The contribution of the paper lies in the method for constructing shape primitives and the use of clustering to determine common local shape patterns. This builds on previous work [34] in the sense that local oriented thickness patterns form the basis of local shape features. But here the orientations are determined automatically and adaptively from the data at each point. This avoids bias introduced by assuming a single preferred orientation within samples and inconsistencies of overall orientation between samples relative to a fixed laboratory reference frame.

The core of the method is described in Section 2 and is applied to shape analysis tasks on three data sets described in Section 3. The feature selection and classification strategies described in Section 4 should not be viewed as part of the proposed method. The strategies employed here were chosen to allow comparison to previous work and as conservative methods suitable for three example studies, all of which comprise small numbers of samples.

2. Representing local shape

2.1. Fitting ellipsoids

Let X denote a binary function on n-dimensional Euclidean space ($\mathbb{R}^n$), and let Ω denote the set $\{p \in \mathbb{R}^n : X(p) = 1\}$. If $n = 2$, then X is a binary image.

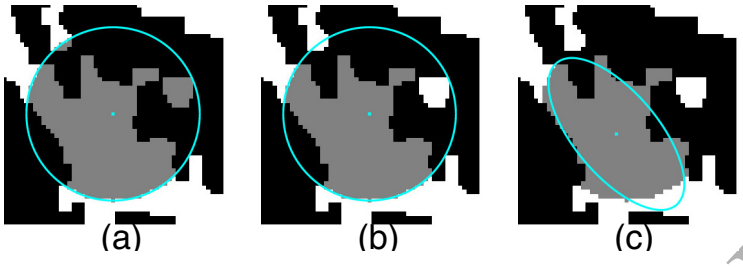


Figure 1: A demonstration of the process of fitting an ellipse in 2D using a slice taken from the cancellous bone data set as an example. The object Ω is shown in white. (a) First, a disk B_r^p (blue) is drawn at p (blue cross). Here, the region $\Omega \cap B_r^p$ (shaded) contains more than one connected component. (b) Second, the connected component Ω^p containing the largest number of occupied points (now shaded) is found. This represents the dominant local shape pattern at p. (c) Third, an ellipse E^p (now blue) is fit to Ω^p (shaded). The ellipse E^p and the connected component Ω^p have the same volume.

The set Ω is the object of interest in the image or array, and the objective is to characterise the shape patterns of Ω. For a point $p \in \Omega$, let B_r^p denote the open ball of radius r centred at p. Let Ω^p denote the largest connected component of $\Omega \cap B_r^p$ and let E^p denote the ellipsoid that best fits Ω^p in the sense that the n-dimensional volume of E^p and Ω^p are equal and that $\mathrm{vol}\,(E^p \cap \Omega^p)$ is maximal subject to the condition of equal volume (Figure 1). The best fitting ellipsoid E^p is easily found by applying principal components on the coordinates of the points in Ω^p. The directions of the n axes of the ellipsoid are given by the unit eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$, each of which has n components, $\boldsymbol{u}_i = (u_{i,1}, u_{i,2}, \ldots, u_{i,n})$. The lengths of each axis $\ell_i,\ i = 1, \ldots, n$, are given by the corresponding eigenvalues, normalised to ensure the volume of E^p is equal to the volume of Ω^p. Let $L_i,\ i = 1, \ldots, n$, be the length of the longest line segment in the direction given by $\boldsymbol{u}_i$, passing through the centroid of E^p, and lying entirely inside the object Ω (Figure 2).

Four definitions of shape primitive vectors $\boldsymbol{v}_p$ based on the constructions above will be considered.

Definition 1 (Oriented ellipsoid parameters)**.**

$$\boldsymbol{v}_p = (\ell_1 u_{1,1}, \ell_1 u_{1,2}, \ldots, \ell_1 u_{1,n}, \ell_2 u_{2,1}, \ell_2 u_{2,2}, \ldots, \ell_2 u_{2,n-1}, \ldots, \ell_n u_{n,1})\,.$$

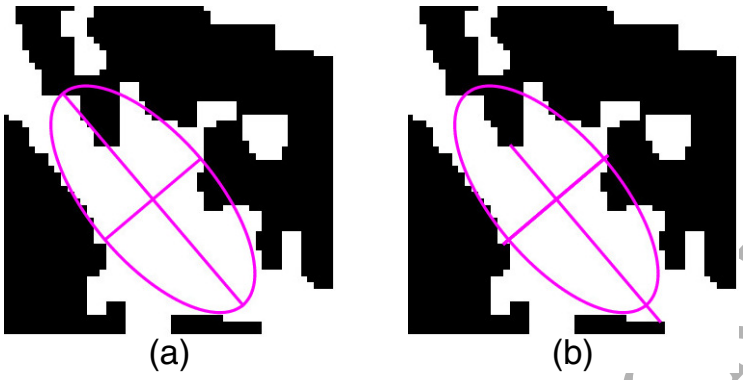


Figure 2: An illustration of the difference between the length measurements ℓ_i, $i = 1, 2$, and L_i, $i = 1, 2$, using the ellipse fit in Figure 1. (a) The lengths ℓ_1 and ℓ_2 correspond to the lengths of the major and minor axes of the ellipse, respectively. (b) The lengths L_1 and L_2 are the lengths of the longest line segments lying entirely inside Ω and passing through the centroid of the ellipse, measured in the directions of the major and minor axes, respectively. The thickness of the object, rather than the ellipse, is measured here.

Oriented ellipsoid parameter shape primitives retain all orientation and length information embodied in E^p to represent the local shape of Ω^p. Since the unit vectors are orthogonal, the ellipsoid may be specified by a total of $n(n+1)/2$ parameters instead of the full n^2 components need to specify n vectors of length n.

Definition 2 (Oriented object parameters)**.**

$$\boldsymbol{v}_p = (L_1 u_{1,1}, L_1 u_{1,2}, \ldots, L_1 u_{1,n}, L_2 u_{2,1}, L_2 u_{2,2}, \ldots, L_2 u_{2,n-1}, \ldots, L_n u_{n,1})\,.$$

Oriented object parameter shape primitives capture information about the local orientation and thickness of the segment of Ω more directly. This method has the drawback that the centroid of the ellipsoid is not guaranteed to fall inside the object Ω, in which case $L_i = 0$ for all $i = 1, \ldots, n$.

Definition 3 (Ellipsoid shape parameters)**.**

$$\boldsymbol{v}_p = (\ell_1, \ell_2, \ldots, \ell_n)\,,$$

Ellipsoid shape parameter shape primitives retain only information about the lengths of the axes of the ellipsoids.

Definition 4 (Object shape parameters)**.**

$$\boldsymbol{v}_p = (L_1, L_2, \ldots, L_n)\,.$$

Object shape parameter shape primitives retain only information about the thickness of the object locally at p.

Regardless of which of the four previous definitions are used, $\boldsymbol{v}_p$ captures information about the shape of the corresponding connected component Ω^p in some way. For an object Ω, the collection $\{\boldsymbol{v}_p : p \in \Omega\}$ corresponds to a collection of local shape patterns occurring throughout Ω.

2.2. Clustered shape primitives

The objective is to characterise the similarities and differences in the shape of objects sampled from m classes. The classes are denoted by $G_g,\ g = 1, 2, \ldots, m$, and the full data set is $G = \bigcup G_g$.

The collection of all shape primitives in a data group G_g is

$$F_g = \{\boldsymbol{v}_p : p \in \Omega, X \in G_g\}\,,$$

and is called the representation space for group g. Shape primitives close to each other in space represent similar shape patterns. Hence, common local shape patterns occurring in the collection G_g may be identified by clustering the vectors in F_g. The resulting cluster centres are referred to as clustered shape primitives (CSPs), and may be viewed as quantitative shape descriptors learned automatically from the data. The CSPs for data group G_g are vectors C_k^g, $k = 1, 2, \ldots, K$, of the same dimension as the shape primitives $\boldsymbol{v}_p$. With m groups and K CSPs per group, there are $M = Km$ CSPs in total. For convenience, these are labelled sequentially as C_j, where $C_{(g-1)K+k} = C_k^g, g = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, K$.

The CSP map Y associated with X is obtained by replacing every point $p \in \Omega$ by the label of the CSP closest to $\boldsymbol{v}_p$. That is,

$$Y(p) = \underset{j \in \{1,2,\ldots,M\}}{\arg\min} \{\|C_j - \boldsymbol{v}_p\|\}\,,$$

where $\|\cdot\|$ is the Euclidean norm. The normalised histogram of CSP labels associated with X is

$$\boldsymbol{h}_X = (h_1, h_2, \ldots, h_M)\,, \tag{1}$$

where h_j is the proportion of points in Ω with label j. The histogram $\boldsymbol{h}_X$ represents the shape content of X and numbers h_j are viewed as features for classifying X.

In practice, the choice of the parameter r will affect the information captured by the ellipsoids. If r is very small compared to the characteristic radius of shape features in Ω, then B_r^p will often fall entirely inside the object Ω. Thus, the majority of ellipsoids fit to the object will be balls and so will fail to capture information about the orientation and thickness of the components. If r is very large compared to the characteristic radius of the shape features of Ω, then the largest component inside B_r^p will not capture local structure. For a sensible choice of r, local information is captured well (Figure 3). The parameter r may be determined empirically, or by domain knowledge of the scale of the components of interest.

3. Examples

Three data sets were used in this study: binary images (2D) of yeast colonies, blocks of cancellous bone (3D) obtained from rat tibiae and marbling in beef striploins (3D).

3.1. Yeast colony data

The yeast colony study is retrospective, based on data described in [31]. Single cells of the *S. cerevisiae* AWRI796 strain were used to initiate the growth of 10 individual yeast colonies with 50 μM ammonium sulphate nutrient concentration. Each of the colonies was imaged successively over time, with images being taken at $t = \{23, 48, 73, 87, 115, 162, 211, 233\}$ hours after initialisation of growth. The images were converted into 2D binary images using customised software [43, 44], where pixels were assigned as 'occupied by yeast cells' (Ω) or

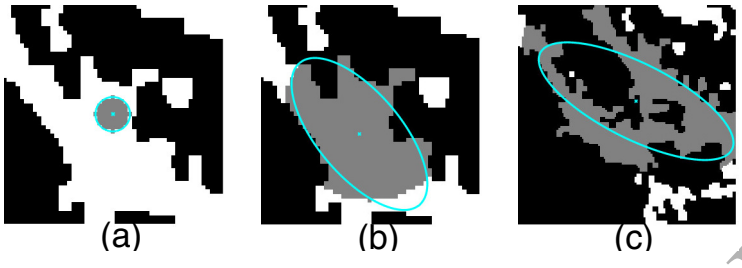


Figure 3: A demonstration of the effect of the parameter r on the information captured by the ellipse, using the same 2D example as in Figures 1 and 2. The set Ω is shown in white, the connected component Ω^p is shaded, and the fit ellipse is shown in blue with the cross representing its centroid. (a) The parameter r is very small compared to the size of the components of the object, and the resulting ellipse is a disk. (b) The choice of r appears to be sensible. (c) The value of r is very large compared with the size of the object components. The image has been zoomed out, but the starting point p has been kept consistent with that in Figure 1. The connected component Ω^p is very large, and the ellipse appears to fail to capture shape information at the local level. In this case, the centroid of the ellipse falls outside of the set Ω.

'unoccupied' (Figure 4). Depending on the size of the colony at the time of imaging, the resolution varied. For $t = 23, 48$ hours, the resolution was approximately 0.61 μm $\times$0.61 μm per pixel, and for $t = 73$ through 233 hours, the resolution was approximately 1.52 μm $\times$1.52 μm per pixel. More details may be found in Binder *et al.* [31].

The data set was separated into eight experimental groups corresponding to each of the eight timesteps. For each timestep, five of the colonies were randomly assigned to the training set, and the remaining five were assigned to the testing set.

3.2. Cancellous bone data

The data for the cancellous bone study was also collected as part of a previous study [32]. Thirty female rats were randomly assigned to three experimental groups of 10 rats each. At the beginning of the study (week 0), the sham group underwent a sham surgery, where an incision was made but no organs were removed. This was the control group. The ovx group underwent a full

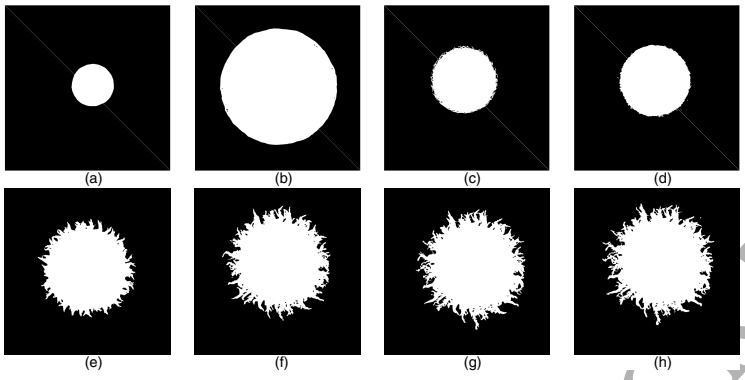


Figure 4: An illustration of the growth of colony number 5 for each of the eight time steps. (a)–(h) The colony was imaged at $t = \{23, 48, 73, 87, 115, 162, 211, 233\}$ hours, respectively. For $t = 23$ and $t = 48$ hours, the resolution is approximately 0.61 μm $\times$0.61 μm per pixel. For $t = 73, \ldots, 233$ hours, the resolution is approximately 1.52 μm $\times$1.52 μm per pixel.

ovariectomy, and the ovx+zol group underwent a full ovariectomy plus subsequent treatment with zoledronic acid, a bisphosphonate, starting two weeks after surgery (week 2). The rats in the sham and ovx groups received saline injections at the same time as the rats in the ovx+zol group received the zoledronic acid injections.

Micro-computed tomography (μ-CT) scans of the right tibia of each rat were taken at $t = \{0, 2, 4, 8, 12\}$ weeks after surgery. Since two rats died between weeks 8 and 12, the scans from week 8 were used in the current study. The resolution of the scans was 8.70 μm $\times$8.70 μm$\times$8.70 μm per voxel. For each of the 30 rats, a rectangular block of size $121 \times 121 \times 400$ voxels was manually segmented from the μ-CT reconstruction, corresponding to bone of approximate size 1 mm $\times$ 1 mm $\times$ mm. The longest edge of each block was roughly parallel to the main axis of the tibia, and the first cross-sectional slice was approximately 1.2 mm from the growth plate of the bone. After segmentation, each voxel was assigned as 'occupied by bone' (Ω) or 'unoccupied', resulting in a data set of 30 binary bone blocks (Figure 5). Further details may be found in Fazzalari *et al.* [32].

Each bone block was divided into three sub-blocks, with the first sub-block

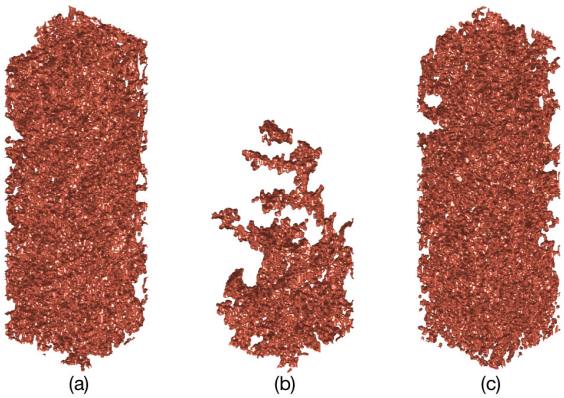


Figure 5: Examples of cancellous bone taken from rat tibiae eight weeks after the beginning of the study, one for each experimental group. The bottom of each block corresponds to the slice closest to the growth plate (slice 1). (a) The sham group corresponds to the control group. (b) The rats in the ovx group underwent an ovariectomy at the beginning of the study. (c) Rats in the ovx+zol group underwent an ovariectomy at the beginning of the study, and were subsequently treated for bone loss with zoledronic acid.

comprising the first 100 slices counting from the growth plate end, the second sub-block comprising slices 101–200, and the third sub-block comprising slices 201–300. The last 100 slices furthest from the growth plate were not used in the study because in some cases the bone was too sparse to meaningfully characterise the structure (Figure 6). A total of nine groups were used in the study, corresponding to three experimental groups and three distances from the growth plate. The reason for distinguishing between different distances from the growth plate was that the rats in the study were growing and new bone was being formed by endochondral ossification at the growth plate. Hence, different structure patterns may occur at different distances from the growth plate, in addition to differences in the structure between the three experimental groups.

For each of the three experimental groups, five of the 10 rats were randomly assigned to the training set and the remaining five rats were assigned to the testing set.

3.3. Beef marbling data

The data for the marbling in beef was originally acquired for a study on associating vitamin A intake and fat metabolism. Details of the study design the acquisition of the images have been described previously in detail [42]. Briefly, Striploins from 20 Angus steers were frozen and then sliced into 25 steaks, each 4 mm thick. The steaks were photographed at a resolution 7.5 pixels per mm. The steak images were manually segmented to retain the rib eye area and the marbling within. All further image analysis steps were performed automatically using software written by one of the authors (MJB) in Matlab. The images were converted to grayscale, filtered to remove lighting effects resulting in flattened images and then converted to binary images using a threshold (Figure 7). For each striploin, the 25 binary images were stacked to form a 3D array representing the marbling within the striploin. The marbling ratio for each striploin was taken to be the number of "on" voxels (voxels representing marbling) divided by the number of voxels comprising the full striploin.

As an example of the use of 3D shape descriptors in the present study,

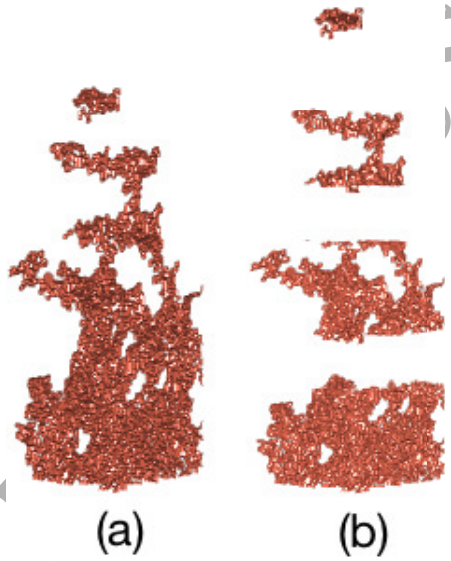
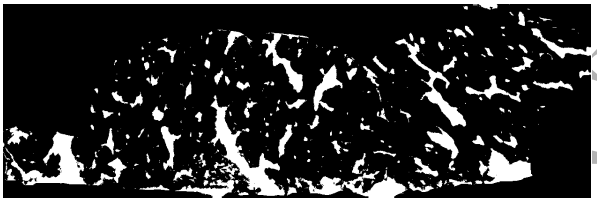


Figure 6: An illustration of splitting up a rectangular block into three sub-blocks, using a block from the ovx group as an example. The slices closest to the bottom of the image are always closest to the growth plate of the bone. (a) The original bone block. (b) From bottom to top: sub-blocks 1–4. The first three sub-blocks are used in the study.

(a)



(b)

Figure 7: Example steak image. (a) Original photograph of a steak. (b) The binary version showing the marbling.

the objective was to determine the correlation between the marbling ratio and structure of the 3D marbling as represented by local shape patterns. Thus, in this example, the shape features are used for regression rather than classification.

4. Implementation details

For all three data sets, the objective was to describe the shape characteristics of the object Ω. For each data set, the points at which the ellipsoids were initialised were subsampled from the set of points in Ω. In the case of the yeast colony data, let $\partial\Omega$ be the subset of Ω comprising boundary pixels, that is, pixels in Ω with at least one unoccupied neighbour. Let c be the centroid of the colony, found by taking the average values of the occupied x and y coordinates, and let

$$R = \arg\min_{p \in \partial\Omega} \{||c - p||\}.$$

That is, R is the radius of the largest disk B_R^c centred at c that fits entirely inside the colony. Let $\Omega_{\text{fil}} = \Omega \setminus B_R^c$, so named because this is the "filamentous" part of the colony (Figure 8). Ellipses were initialised at a subsampled set of

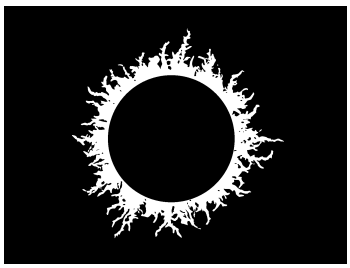


Figure 8: The subset Ω_{fil} corresponding to the filamentous part of the colony (white region). Ellipsoids were only initialised within this region. This example is colony 5 imaged at $t = 233$ hours.

points $p \in \Omega_{\text{fil}}$, since the shape of the filamentous part of the colony is of interest. In the case of the cancellous bone data and the marbling data, ellipsoids were initialised at a randomly subsampled set of points from Ω.

For all data sets, the best fitting ellipsoid E^p to Ω^p was found as described in Section 2.1. Shape primitives were constructed according to the general construction given in Definitions 1–4. For the beef marbling data, only Definition 1 was considered because of the disparity between the in-plane and between plane resolution of the data. For the oriented object parameter (Definition 2) and object shape parameter (Definition 4) shape primitives, if $L_i = 0$ for all $i = 1, \ldots, n$, then those shape primitives were omitted from the representation space. The proportion of centroids of all ellipsoids fit to all images or arrays falling outside the object Ω was recorded in order to ascertain whether an unreasonable proportion of shape primitives was being omitted.

For the yeast colony data and the cancellous bone data, CSPs were computed by clustering each representation space F_g constructed from the training images for each group g. A K-means clustering algorithm with 100 iterations and one replicate was used for clustering, with $K = 10$ clusters per class for yeast and cancellous bone data sets. The value $K = 10$ was based on preliminary experiments and previous work [45]. This resulted in $M = 80$ or $M = 90$

CSPs in total for yeast colonies and cancellous bone, respectively (Table 1). For the regression problem in the case of the marbling example, there was no previous work to guide the number of clusters. Hence, separate trails were conducted with the value of $K = 4, 6, 10, 14, 18, 22$. Normalised histograms (sum of columns equals one) of CSP labels were computed for all images or arrays in the data set.

For the yeast colony and cancellous bone data sets, linear discriminant analysis was used as the classifier during the training and testing steps. The best combination of three features was found by exhaustive search. Each combination of features was tested individually and the combination that gave the highest classification accuracy was selected for the testing step. The K-means algorithm was repeated 30 times. For each definition of the shape primitives, the entire process was repeated five times with the parameter r set at varying lengths (Table 1).

For the beef marbling data, the choice of r was restricted by the discrepancy between the in-plane resolution (voxel side length 0.133 mm) and the between-plane resolution (voxel side length 4.0 mm). Only the value of $r = 6.78$ mm was considered. This corresponds roughly to a ball comprising three steaks in diameter. Ten striploins were randomly selected for training the linear regression and resulting regression line was used to predict the proportion of marbling for the remaining 10 striploins.

5. Results and discussion

5.1. Yeast colonies

When fitting ellipses to the yeast colony data, the largest proportion of ellipse centroids that fell outside the colony was 3.11×10^{-2} for $r = 75.6$ μm with the proportions dropping for smaller and larger values of r (Table 2). In the context of yeast colonies, the possibility of centroids of the ellipses falling outside Ω did not pose a serious problem, since this caused relatively few feature vectors to be excluded from the representation space.

Table 1: Numerical details of the classifier for each data set. # pts is the number of subsampled points at which the ellipsoids were initialised, n is the dimension, K is the number of clusters obtained using K-means clustering, M is the number of CSPs, # fts is the number of best features selected during classification, and r is the radius of the balls (in μm) used to initialise the ellipsoids. The values of r correspond to whole numbers of pixels or voxels.

Data set	**# pts**	n	K	M	**# fts**	r
Yeast colonies	10000	2	10	80	3	3.03, 15.2, 75.6, 303, 909
Cancellous bone	10000	3	10	90	3	17.4, 43.5, 131, 261, 435
Beef marbling	10000	3	10	10	1	6780

Table 2: The proportion of ellipses fit to all 80 images whose centroids fell outside the colony as a function of r (in μm) when ellipses were fit to yeast colony data.

r	**3.03**	**15.2**	**75.6**	**303**	**909**
Proportion	6.00×10^{-4}	8.27×10^{-3}	3.11×10^{-2}	7.28×10^{-4}	0

When the shape primitives were recorded using oriented ellipsoid parameters (Definition 1), the highest classification accuracy was achieved using $r = 303$ μm (Table 3). Here, information about the orientation and size of the ellipses was included in the shape primitives. Using disks B_r with $r < 303$ μm results in ellipses that capture information at the local level. Although there appears to be some filamentation at the local level, these small filaments seem to be similar in shape regardless of the development of the larger filaments. These ellipses fail to capture the shape of the larger filaments. For $r = 909$ μm the ellipses are very large and capture information about a large area of the colony, leading to information without much discriminative power (Figure 9). The ellipses fit using disks B_r, $r = 303$ μm result in a good tradeoff between picking out information at the local and the global level.

When oriented object parameters were recorded in the shape primitives (Definition 2), the best classification score was achieved when the ellipses were fit using disks of radius $r = 15.2$ μm. Here, the orientation of each ellipse was

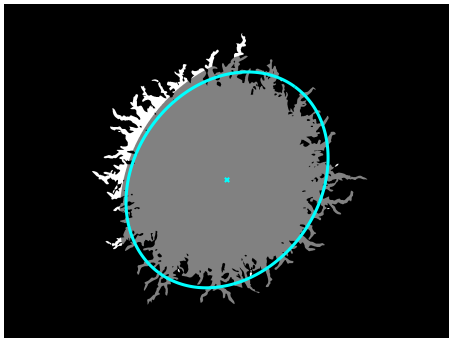


Figure 9: An illustration of an ellipse (blue) fit using a ball B_r of radius $r = 909$ μm using colony number 5 at $t = 233$ hours after initiation of growth as an example. The connected component Ω^p is shaded.

recorded, along with the length of the colony itself in the major and minor directions of the ellipse. If disks B_r with $r < 15.2$ μm are used to fit the ellipses, then the ellipses fall too close to the boundary of the colony, resulting in length measurements that do not accurately capture the shape patterns of the colony. For $r > 15.2$ μm the ellipses are large, giving centroids and directions that result in length measurements with little discriminatory power (Figure 9). Using disks with $r = 15.2$ μm results in the centroids of the ellipsoids falling far from the boundaries of the filaments, meaning that oriented object parameters would extract useful information about the length and width of the filaments (Figure 10).

When only ellipsoid shape parameters were recorded in the shape primitives (Definition 3), the highest classification accuracy was obtained with $r = 75.6$ μm. This suggests the ellipses fit the data closely, being of a comparable length and width to that of the filaments themselves. When only object shape parameters were recorded (Definition 4), the best classification accuracy occurred at $r = 3.03$ μm (Table 3). This suggests the ellipses do not fit the data closely, but their centroids correspond to a set of subsampled points at which informative object shape parameter measurements are made.

The highest classification accuracy overall ($\mu = 0.843$, $\sigma = 0.020$) was

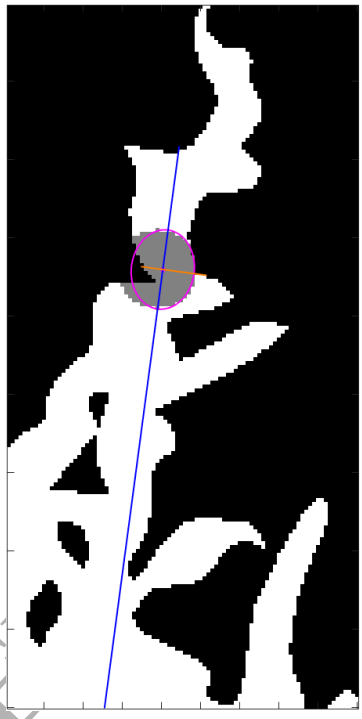


Figure 10: An illustration of an ellipse (red) fit using a ball B_r of radius $r = 15.2$ μm using one filament (white) isolated from colony number 5 at $t = 233$ hours. The largest connected component Ω^p that intersects with B_r^p is shaded in grey. The blue and orange lines show the length measurements in the major and minor directions of the ellipse, respectively. Oriented object parameters appear to adequately capture the shape of the filament at this point, whereas the thickness of the ellipse itself does not extract as much information about the shape of the colony.

Table 3: Best classification accuracy as a function of r (μm) for yeast colony data. Columns two to five correspond to the accuracy obtained using each of the four methods to record the shape primitives (Definitions 1–4, respectively). For each method, 80 CSPs were computed and 30 trials were conducted. The highest accuracy achieved by each method is shown in bold.

r	Oriented ellipsoid parameters	Oriented object parameters	Ellipsoid shape parameters	Object shape parameters
3.03	$\mu = 0.722$, $\sigma = 0.014$	$\mu = 0.819$, $\sigma = 0.025$	$\mu = 0.744$, $\sigma = 0.022$	$\mu = \mathbf{0.785}$, $\sigma = \mathbf{0.018}$
15.2	$\mu = 0.702$, $\sigma = 0.023$	$\mu = \mathbf{0.843}$, $\sigma = \mathbf{0.021}$	$\mu = 0.776$, $\sigma = 0.019$	$\mu = 0.773$, $\sigma = 0.025$
75.6	$\mu = 0.763$, $\sigma = 0.033$	$\mu = 0.808$, $\sigma = 0.020$	$\mu = \mathbf{0.840}$, $\sigma = \mathbf{0.028}$	$\mu = 0.633$, $\sigma = 0.025$
303	$\mu = \mathbf{0.827}$, $\sigma = \mathbf{0.020}$	$\mu = 0.815$, $\sigma = 0.023$	$\mu = 0.755$, $\sigma = 0.010$	$\mu = 0.576$, $\sigma = 0.021$
909	$\mu = 0.478$, $\sigma = 0.009$	$\mu = 0.485$, $\sigma = 0.013$	$\mu = 0.413$, $\sigma = 0.019$	$\mu = 0.471$, $\sigma = 0.016$

achieved by recording oriented object parameters, after fitting ellipses using disks of radius $r = 15.2$ μm (Table 3). This result suggests that interrogating the data directly, and including directional information in conjunction with this, is the most effective way of extracting information from the yeast colony data.

5.2. Cancellous bone

When ellipsoids were fit to the cancellous bone data, a large proportion of centroids fell outside of the bone itself. The proportion increased with r, with the largest proportion being 0.340 when balls of radius $r = 435$ μm were used to fit the ellipsoids (Table 4).

When oriented ellipsoid parameters were used for the shape primitives (Definition 1), the highest classification accuracy was achieved at $r = 131$ μm (Table 5). If balls of smaller radius are used, then the subset Ω^p is small and does not contain much meaningful information about the local structure. In fact, the

Table 4: The proportion of ellipsoids fit to all 90 sub-blocks whose centroids fall outside the bone as a function of r (in μm) for cancellous bone data.

r	17.4	43.5	131	261	435
Proportion	7.28×10^{-4}	6.33×10^{-3}	0.083	0.236	0.340

balls B_r^p are likely to fall entirely inside the bone, for most points p, resulting in ellipsoids that are also balls. If balls of larger radius are used, then each Ω^p is likely to be large and complicated, and the ellipsoid properties fail to capture orientation and thickness information at the local level. Choosing $r = 435$ μm means that the diameter of each ball is almost as large as the length of the two shortest sides of the bone block, so it is not surprising that this choice of r corresponded to the lowest classification accuracy. The choice $r = 151$ μm gives the optimal tradeoff between extracting local versus global information about the trabecular structure.

When oriented object parameters were recorded (Definition 2), the equal highest classification score occurred when the ellipsoids were fit using balls of radius $r = 17.4$ and 43.5 μm. The standard deviation for $r = 17.4$ μm was slightly smaller than that for $r = 43.5$ μm (Table 5). As r increased, the proportion of centroids of the ellipsoids that fell outside the bone also increased, and the classification accuracy decreased. This suggests that, although the ellipsoids fit using balls with $r = 17.4$ and 43.5 μm may be small and capture local information only, their centroids most likely do fall close to the centre of the bone and hence result in meaningful bone thickness measurements, especially when combined with directional information.

In the case of the cancellous bone, the classification accuracy tended to improve when directional information was removed from the shape primitives. The highest classification accuracy overall ($\mu = 0.745$ $\sigma = 0.024$) occurred at $r = 131$ μm when only object shape parameters (Definition 4) were recorded (Table 5). In the context of cancellous rat bone, including directional information in the features does not increase the discriminatory power of the classifier.

Table 5: Best classification accuracy as a function of r (μm) for cancellous bone data. Columns two to five correspond to the accuracy obtained using each of the four methods to record the shape primitives (Definitions 1–4, respectively). For each method, 90 CSPs were computed and 30 trials were conducted. The highest accuracy achieved using each method is shown in bold.

r	Oriented ellipsoid parameters	Oriented object parameters	Ellipsoid shape parameters	Object shape parameters
17.4	$\mu = 0.593$, $\sigma = 0.021$	$\mu = \mathbf{0.676}$, $\sigma = \mathbf{0.016}$	$\mu = 0.633$, $\sigma = 0.025$	$\mu = 0.670$, $\sigma = 0.021$
43.5	$\mu = 0.640$, $\sigma = 0.019$	$\mu = 0.676$, $\sigma = 0.020$	$\mu = 0.670$, $\sigma = 0.014$	$\mu = 0.666$, $\sigma = 0.017$
131	$\mu = \mathbf{0.663}$, $\sigma = \mathbf{0.026}$	$\mu = 0.664$, $\sigma = 0.021$	$\mu = \mathbf{0.695}$, $\sigma = \mathbf{0.016}$	$\mu = \mathbf{0.745}$, $\sigma = \mathbf{0.024}$
261	$\mu = 0.620$, $\sigma = 0.024$	$\mu = 0.649$, $\sigma = 0.029$	$\mu = 0.670$, $\sigma = 0.024$	$\mu = 0.713$, $\sigma = 0.020$
436	$\mu = 0.613$, $\sigma = 0.023$	$\mu = 0.625$, $\sigma = 0.023$	$\mu = 0.657$, $\sigma = 0.019$	$\mu = 0.624$, $\sigma = 0.016$

However, the lengths of the bone are measured in the directions of the major, middle, and minor axes of the ellipsoid. Hence, the length measurements alone contain some directional information in the sense that the important directions are learned directly from the data, as opposed to making assumptions about the direction of the main axis of the tibia.

5.3. Marbling in beef

The shape analysis of marbling in beef differed from the shape analysis of the yeast colonies and the cancellous bone in three ways.

First, the anisotropic shape of the voxels limited the range of the radii used to define the region for fitting the ellipsoids. If the radii are measured in steak widths (4 mm), then a ball of radius $r = 1$ includes voxels from the steak containing the centre of the ball and the steak on either side (thus the diameter is 3 steak widths). A ball of radius two includes voxels from 5 steaks and

Table 6: Regression results for beef marbling data. The first row is the number of clusters. The second and third rows are the R^2 values for the regression of marbling ratio on a single shape feature for training and testing, respectively.

K	4	6	10	14	18	22
Training R^2	0.900	0.958	0.970	0.872	0.878	0.865
Testing R^2	0.786	0.759	0.797	0.871	0.832	0.455

thus corresponds to one fifth of the total extent of the striploin in the direction perpendicular to the planes of the steaks. Accordingly, a study on varying the size of the ball was not practical for this data. This illustrates that the size of the ball may well be constrained by the nature of data. The choice of the radius of the ball is not part of the method proposed here. This parameter, may be set according to prior understanding of the data as in the marbling example or by experimentation, if necessary. Even in the yeast and cancellous bone examples, simple inspection of the data (Figure 10, for example), would lead to a sensible choice for r.

Second, the marbling example was a regression study rather than a classification study. This meant that the practice of computing separate CSPs for each class could not be implemented on this data. Instead the number of clusters was viewed as a parameter to be determined by experimentation (Table 6).

Third, the experiment on correlating marbling ratio to a shape parameter was conducted without any knowledge of ground truth. There is no previous literature on the relationship between shape of marbling in 3D other properties of beef and so there is no claim that the shape parameter provides the correct correlation. If no correlation had been found, one could equally speculate that this was because the shape features did not capture pertinent shape information or because there is no correlation between marbling ratio and marbling shape. Since the correlation was found to be quite strong (Table 6), the example demonstrates that the shape features do capture pertinent shape information. The existence of the relationship between marbling ratio and shape is a discovery

of interest to meat scientists and is currently under further investigation.

The strongest correlation for the testing set (coefficient of determination $R^2 = 0.871$) was found when 14 clusters were used (Table 6) but this is somewhat misleading since, in practice, one has to apply the existing best method to new data. In this small experiment, the best training score was found with 10 clusters and if the shape feature from this trial is used, the correlation for the testing data is still substantial ($R^2 = 0.797$)

The training error reduces only very slightly if more than 14 clusters are used, but the testing error drops drastically at 22 clusters. This indicates that the increasing the number of clusters runs the risk of overfitting.

5.4. Comparison with previous work

In the current study, the definitions of the shape primitives are new. In a preliminary study [34], each ellipsoid was represented by the shape primitive

$$\boldsymbol{v}_p = (u_{1,1}, u_{1,2}, u_{1,3}, u_{2,1}, u_{2,2}, \ell_1, \ell_2, \ell_3)\,,$$

where the lengths $\ell_i,\ i = 1, 2, 3$, were measured in terms of voxel units. The elements of those shape primitives were not necessarily of equal magnitude (all values of the elements of the unit direction vectors, $u_{1,1}, u_{1,2}, u_{1,3}, u_{2,1}, u_{2,2}$, were in the interval $[-1, 1]$, whereas $\ell_i,\ i = 1, 2, 3$, could vary by up to two orders of magnitude). Here, all elements of the shape primitives were of comparable magnitude. This removes possible bias in the clustering step. In addition, all length measurements were made in μm rather than voxel units, meaning that the resolution of the data was accounted for and length measurements were standardised in some sense. Furthermore, in this study, the entire ellipsoid-based method for extracting local shape features was validated on two additional data sets, namely 2D binary images of yeast colonies and 3D binary arrays representing marbling in beef.

The best classification scheme for cancellous rat bone proposed in this study ($\mu = 0.745,\ \sigma = 0.024,\ 30$ trials) significantly outperformed the classification schemes using shape features alone ($\mu = 0.607,\ \sigma = 0.018,\ 10$ trials, $p < 0.001$),

and standard attributes alone (classification accuracy of 0.467, $p < 0.001$) in Martin and Bottema [33], as well as the classification scheme proposed using ellipsoids ($\mu = 0.699$, $\sigma = 0.014$, 30 trials, $p < 0.001$) in the preliminary study [34]. The p-values for the first and third comparisons correspond to a two-sample t-test. The p-value for the second test is the value of the cumulative distribution function of the normal distribution with mean 0.745 and standard deviation 0.024 at $x = 0.467$.

A key attribute of the method proposed here is that no information is used regarding the overall shape of samples or large scale shape similarity between samples. The method does not attempt to classify or characterise overall shape as to most other methods for shape analysis. The full shape of the object is not of interest because, in all the three example studies, the overall shape caries little or no information (yeast colonies are all essentially round, the cancellous bone samples are rectangular blocks, the extent of the marbling is determined by the striploin itself). Instead, shape is characterised in terms of frequencies of local shape patterns only. Some other methods use frequencies of local shape but do so as a vehicle for characterising the global shape. Shape context, for example, uses histograms of local shape features but also a matching step to align overall shape [21]. Indeed, the final goal of most shape analysis is to understand the overall shape for problems in image understanding, image retrieval, handwriting recognition, etc. In most of these problems, the focus is on objects that are well understood in terms of within-class and between-class attributes of overall shape (distinguishing butterflies from leaves or airplanes, etc.) Naturally, most algorithms focus on these large scale features.

A consequence is that these other methods cannot be applied to situations where the global shape is immaterial. Equally, the method proposed here is not expected to compete well with classical methods for shape analysis on problems where overall shape is the objective (or plays a major role) because overall shape is not considered. The precludes meaningful benchmarking against many methods.

6. Conclusion

For the yeast colonies and cancellous bone data sets, the highest classification accuracy was achieved using either oriented object parameters or object shape parameters, suggesting that interrogating the data directly is advantageous. For the yeast data, recording oriented object parameters resulted in the highest classification accuracy, suggesting that information about the thickness and orientation of the colony is important towards the classification. For the cancellous bone data, the highest classification score was achieved using object shape parameters (Definition 4).

For the beef marbling data, computing oriented object parameters was not practical due to the discrepancy between in-plane and between-plane resolution. This example shows that, while computing oriented object parameters may be preferred, oriented ellipsoid parameters are able to deliver good results when the former is not practical.

The results from three diverse data sets demonstrate that the shape primitives defined in Section 2 together with the clustering step capture local shape patterns well enough to distinguish class membership or in the case of beef marbling, establish correlations. Local thickness values are computed in only three directions according to the orientations of the three axis of the ellipsoids in the 3D data sets and two directions in the case of the yeast colonies. In [33], 13 orientations were used but were fixed over each sample and set according to an external frame of reference. The fact that, in the present study, better results were obtained using less interrogation (three orientations instead of 13) demonstrates the advantage of defining orientations locally, adaptively and independent of an external reference frame. The method applies in situations where the overall shape of the object is not important but local shape patterns carry useful information.

References

[1] F. J. Rohlf, Morphometrics, Annu. Rev. Ecol. Syst. 21 (1990) 299–316.

[2] S. Loncaric, A survey of shape analysis techniques, Pattern Recogn. 31 (8) (1998) 983–1001.

[3] T. Pavlidis, Algorithms for shape analysis of contours and waveforms, TPAMI 2 (4) (1980) 301–312.

[4] J. P. Keating, D. Brophy, R. A. Officer, E. Mullins, Otolith shape analysis of blue whiting suggests a complex stock structure at their spawning grounds in the Northeast Atlantic, Fish. Res. 157 (2014) 1–6.

[5] J. Camargo Neto, G. E. Meyer, D. D. Jones, A. K. Samal, Plant species identification using Elliptic Fourier leaf shape analysis, Comput. Electron. Agr. 50 (2006) 121–134.

[6] T. McLellan, J. A. Endler, The relative success of some methods for measuring and describing the shape of complex objects, Syst. Biol. 47 (2) (1998) 264–281.

[7] D. J. Innes, J. A. Bates, Morphological variation of *Mytilus edulis* and *Mytilus trossulus* in eastern Newfoundland, Mar. Biol. 133 (1999) 691–699.

[8] S. Y. Y. Chen, P. E. Lestrel, W. J. S. Kerr, J. H. McColl, Describing shape changes in the human mandible using elliptical Fourier functions, Eur. J. Orthodont. 22 (2000) 201–216.

[9] A. R. Dill, M. D. Levine, P. B. Noble, Multiple resolution skeletons, TPAMI 9 (4) (1987) 495–504.

[10] F. Bookstein, K. Schäfer, H. Prossinger, H. Seidler, M. Fieder, C. Stringer, G. W. Weber, J.-L. Arsuaga, D. E. Slice, F. J. Rohlf, W. Recheis, A. J. Mariam, L. F. Marcus, Comparing frontal cranial profiles in archaic and modern *Homo* by morphometric analysis, Anat. Rec. B New Anat. 257 (1999) 217–224.

[11] F. J. Rohlf, A. Loy, M. Corti, Morphometric analysis of Old World Talpidae (Mammalia, Insectivora) using partial-warp scores, Syst. Biol. 45 (3) (1996) 344–362.

[12] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active Shape Models – Their training and application, Comput. Vis. Image Und. 61 (1) (1995) 38–59.

[13] T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: A review, Med. Image Anal. 13 (2009) 543–563.

[14] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, Int. J. Comput. Vision 1 (1988) 321–331.

[15] A. Hill, A. Thornham, C. J. Taylor, Model-based interpretation of 3D medical images, in: Proceedings of the British Machine Vision Conference, BMVA Press, 1993, pp. 339–348. doi:10.5244/C.7.34.

[16] T. F. Cootes, C. J. Taylor, Active Shape Models - 'Smart Snakes', in: Proceedings of the British Machine Vision Conference, BMVA Press, 1992, pp. 266–275. doi:10.5244/C.6.28.

[17] T. F. Cootes, A. Hill, C. J. Taylor, J. Haslam, The use of active shape models for locating structures in medical images, in: B. H.H., G. A.F. (Eds.), Information Processing in Medical Imaging, Lecture Notes in Computer Science 687, Springer, Berlin, Heidelberg, IPMI 1993, pp. 28–37.

[18] T. F. Cootes, C. J. Taylor, A. Lanitis, D. H. Cooper, J. Graham, Building and using flexible models incorporating grey-level information, in: Proceedings of the Fourth International Conference on Computer Vision, 1993, pp. 242–246.

[19] A. Lanitis, C. J. Taylor, T. F. Cootes, A generic system for classifying variable objects using flexible template matching, in: Proceedings of the British Machine Vision Conference, BMVA Press, 1993, pp. 329–338. doi:10.5244/C.7.33.

[20] R.-S. Lu, E. Dennison, H. Denison, C. Cooper, M. Taylor, M. J. Bottema, Texture analysis based on Gabor filters improves the estimate of

bone fracture risk from DXA images, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualizationdoi:10.1080/21681163.2016.1271726.

[21] G. Mori, J. Malik, Estimating human body configurations using shape context matching, in: A. Heyden, G. Sparr, M. Nielsen, P. Johansen (Eds.), Computer Vision — ECCV 2002, Springer Berlin Heidelberg, Berlin, Heidelberg, 2002, pp. 666–680.

[22] X. Wang, B. Feng, X. Bai, W. Liu, L. J. Latecki, Bag of contour fragments for robust shape classification, Pattern Recognition 47 (6) (2014) 2116 – 2125. doi:https://doi.org/10.1016/j.patcog.2013.12.008.
URL http://www.sciencedirect.com/science/article/pii/S0031320313005426

[23] J. uni, P. L. Rosin, V. Ili, Disconnectedness: A new moment invariant for multi-component shapes, Pattern Recognition 78 (2018) 91 – 102. doi:https://doi.org/10.1016/j.patcog.2018.01.010.
URL http://www.sciencedirect.com/science/article/pii/S0031320318300116

[24] M. Grogan, R. Dahyot, Shape registration with directional data, Pattern Recognition 79 (2018) 452 – 466. doi:https://doi.org/10.1016/j.patcog.2018.02.021.
URL http://www.sciencedirect.com/science/article/pii/S003132031830075X

[25] M. Huang, J. Lin, N. Chen, W. An, W. Zhu, Reversed sketch: A scalable and comparable shape representation, Pattern Recognition 80 (2018) 168 – 182. doi:https://doi.org/10.1016/j.patcog.2018.03.001.
URL http://www.sciencedirect.com/science/article/pii/S0031320318300852

[26] R. Guler, S. Tari, G. Unal, Landmarks inside the shape: Shape matching using image descriptors, Pattern Recognition 49 (2016) 79 – 88.

doi:https://doi.org/10.1016/j.patcog.2015.07.013.
URL http://www.sciencedirect.com/science/article/pii/S0031320315002733

[27] Q. Yu, H. Wei, C. Yang, Local part chamfer matching for shape-based object detection, Pattern Recognition 65 (2017) 82 – 96. doi:https://doi.org/10.1016/j.patcog.2016.11.020.
URL http://www.sciencedirect.com/science/article/pii/S0031320316303788

[28] V. Premachandran, R. Kakarala, Perceptually motivated shape context which uses shape interiors, Pattern Recognition 46 (8) (2013) 2092 – 2102. doi:https://doi.org/10.1016/j.patcog.2013.01.030.
URL http://www.sciencedirect.com/science/article/pii/S0031320313000691

[29] C. Panagiotakis, A. Argyros, Parameter-free modelling of 2d shapes with ellipses, Pattern Recognition 53 (2016) 259 – 275. doi:https://doi.org/10.1016/j.patcog.2015.11.004.
URL http://www.sciencedirect.com/science/article/pii/S0031320315004240

[30] Y. Cai, A. Lum, A. Mercado, M. Landis, J. Warrington, S. Li, Unsupervised shape discovery using synchronized spectral networks, Pattern Recognition 69 (2017) 39 – 51. doi:https://doi.org/10.1016/j.patcog.2017.03.032.
URL http://www.sciencedirect.com/science/article/pii/S0031320317301425

[31] B. J. Binder, J. F. Sundstrom, J. M. Gardner, V. Jiranek, S. G. Oliver, Quantifying two-dimensional filamentous and invasive growth spatial patterns in yeast colonies, PLOS Comput. Biol. 11 (2) (2015) e1004070. doi:10.1371/journal.pcbi.1004070.

[32] N. L. Fazzalari, B. L. Martin, K. J. Reynolds, T. M. Cleek, A. Badiei, M. J. Bottema, A model for the change of cancellous bone volume and structure over time, Math. Biosci. 240 (2012) 132–140.

[33] B. L. Martin, M. J. Bottema, Textons for 3D binary data with applications to classifying cancellous bone, in: Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, 2015, pp. 724–729.

[34] A. Gontar, S. Williams, M. J. Bottema, Characterising 3D structure of cancellous bone, in: Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, Gold Coast, Australia, 2016, pp. 365–371.

[35] Z. Kruk, M. Bottema, L. Reyes-Veliz, R. Forder, W. Pitchford, C. Bottema, Vitamin a and marbling attributes: Intramuscular fat hyperplasia effects in cattle, Meat Science 137 (2018) 139 – 146. doi:https://doi.org/10.1016/j.meatsci.2017.11.024.
URL http://www.sciencedirect.com/science/article/pii/S030917401731152X

[36] C. J. Gimeno, P. O. Ljungdahl, C. A. Styles, G. R. Fink, Unipolar cell divisions in the yeast S. cerevisiae lead to filamentous growth: Regulation by starvation and RAS, Cell 68 (1992) 1077–1090.

[37] P. Ruusuvuori, J. Lin, A. C. Scott, Z. Tan, S. Sorsa, A. Kallio, M. Nykter, O. Yli-Harja, I. Shmulevich, A. M. Dudley, Quantitative analysis of colony morphology in yeast, BioTechniques 56 (1) (2014) 18. doi:10.2144/000114123.

[38] A. M. Parfitt, M. K. Drezner, F. H. Glorieux, J. A. Kanis, H. Malluche, P. J. Meunier, S. M. Ott, R. R. Recker, Bone histomorphometry: Standardization of nomenclature, symbols, and units, J. Bone Miner. Res. 2 (6) (1987) 595–610.

[39] A. Oka, Y. Maruo, T. Miki, T. Yamasaki, T. Saito, Influence of vitamin A on the quality of beef from the Tajima strain of Japanese Black cattle, Meat Sci. 48 (1/2) (1998) 159–167.

[40] A. F. Egan, D. M. Ferguson, J. M. Thompson, Consumer sensory requirements for beef and their implications for the Australian beef industry, Aust. J. Exp. Agr. 41 (2001) 855–859.

[41] K. G. Grunert, L. Bredahl, K. Brunsø, Consumer perception of meat quality and implications for product development in the meat sector—a review, Meat Sci. 66 (2004) 259–272.

[42] Z. Kruk, M. Bottema, L. Reyes-Veliz, R. Forder, W. Pitchford, C. Bottema, Vitamin a and marbling attributes: Intramuscular fat hyperplasia effects in cattle, Meat Science 137 (2018) 139 – 146. doi:https://doi.org/10.1016/j.meatsci.2017.11.024.
URL http://www.sciencedirect.com/science/article/pii/S030917401731152X

[43] B. J. Binder, M. J. Simpson, Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions, Phys. Rev. E 88 (2013) 022705. doi:10.1103/PhysRevE.88.022705.

[44] D. J. G. Agnew, J. E. F. Green, T. M. Brown, M. J. Simpson, B. J. Binder, Distinguishing between mechanisms of cell aggregation using pair-correlation functions, J. Theor. Biol. 352 (2014) 16–23.

[45] B. L. Martin, M. J. Bottema, Textons for 3d binary data with applications to classifying cancellous bone, in: 2015 International Conference on Digital Image Comuting: Techniques and Applications (DICTA), IEEE, 2015. doi:10.1109/DICTA.2015.7371312.

Funding

Benjamin J. Binder and Hayden Tronnolone are supported by Australian Research Council Discovery Project DP160102644.

Author biographies

Amelia Gontar received a B.Sc. in Mathematics from Flinders University in 2012. She is now doing her PhD in the School of Computer Science, Engineering and Mathematics at Flinders University. Her research interests include shape analysis, especially in a biomedical imaging context.

Hayden Tronnolone completed a PhD in applied mathematics in 2016, which investigated very viscous fluid flows arising during the fabrication of microstructured optical fibres. He currently works as postdoctoral research associate in the School of Mathematical Sciences at the University of Adelaide. His current research considers the spatial patterns observed during the growth of yeast colonies. This involves a combination of image analysis, modelling (both continuous and discrete), and the analysis of experimental data.

Benjamin Binder was awarded his PhD in Applied Mathematics from the University of East Anglia, UK, 2005. The research involved numerical and analytical studies on non-linear free-surface flows. Shortly after completing his PhD he accepted a Research Associate position at the University of Adelaide, researching the chaotic dynamical systems approach to fluid mixing and its application to designing mixing devices. Benjamin briefly left Adelaide in 2007, joining the University of Melbourne as a Research Fellow. There he developed both discrete and continuous mathematical models for the neural crest cell invasion in the embryonic gut. He returned to the University of Adelaide in 2009, having been appointed a continuing position as a Lecturer in Applied Mathematics. His current research interests include (i) quantifying and modelling biological spatial patterns and (ii) predicting channel bed topography in free-surface flows.

Murk Bottema received a B.A. degree in Physics from the University of Colorado, in 1976 and Ph.D. in Mathematics from Oregon State University in 1986. His research interests include medical image analysis and modelling of medical and biological systems.