Int. J. Appl. Math. Comput. Sci., 2022, Vol. 32, No. 1, 95–109
DOI: 10.34768/amcs-2022-0008

A COMPREHENSIVE STUDY OF CLUSTERING A CLASS OF 2D SHAPES

AGNIESZKA KALISZEWSKA [a], MONIKA SYGA [b,*]

[a] Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
e-mail: Agnieszka.Kaliszewska@ibspan.waw.pl

[b] Faculty of Mathematics and Information Science
Warsaw University of Technology
ul. Koszykowa 75, 00-662 Warsaw, Poland
e-mail: m.syga@mini.pw.edu.pl

The paper is concerned with clustering with respect to the shape and size of 2D contours that are boundaries of cross-sections of 3D objects of revolution. We propose a number of similarity measures based on combined disparate Procrustes analysis (PA) and dynamic time warping (DTW) distances. A motivation and the main application for this study comes from archaeology. The computational experiments performed refer to the clustering of archaeological pottery.

Keywords: shape representation, Procrustes distance, shape similarity, DTW, morphometrics, clustering, Kendall shape theory, typology of archaeological pottery.

1. Introduction

We investigate the clustering of 2D cross-sections of 3D objects of revolution with respect to shape and size. We do not tackle the problem of recognizing whether or not a given 3D object is rotationally symmetric.

Clustering is used extensively in unsupervised machine learning applications such as information retrieval and natural language understanding as well as in numerous fields such as economics, biology, medicine, physics (Leski and Kotas, 2018). However, it evades an in-depth unified framework. This is mostly due to specific challenges arising depending on the nature of data, resulting in case-specific algorithms. Some of these challenges have been presented by, e.g., Kleinberg (2002) or Palacio-Niño and Berzal (2019). There, the so-called *impossibility theorem* is proved which states that no clustering algorithm exists satisfying the following three axioms: scale invariance, richness, and consistency. For further discussion, see, e.g., the works of Cohen-Addad *et al.* (2018; 2017) or Davidson and Ravi (2007).

To obtain significant results for clustering of 2D cross-sections of objects of revolution with respect to shape and size, aside from the choice of the algorithm, the following issues are of fundamental importance: the shape concept, the data representation and the similarity measure (Wierzchoń and Kłopotek, 2015; Jain, 2010). In our investigation we address those three issues in the context of clustering of archaeological pottery.

1.1. State of the art.

The problem of automatic and semi-automatic classification of archaeological pottery has been investigated by several authors. Piccoli *et al.* (2015) proposed similarity measure based on the extraction of shape features from the silhouette through medialness. Sablatnig *et al.* (1998) make use of the 3D models of the investigated shapes in combination with the syntactic pattern recognition approach. The classification of 3D shapes and the issue of matching a fragment to its fully preserved counterpart are investigated by Maiza and Gaildart (2005), who extract the shape information based on a skeleton. Hristov and Agre (2013) present a classification, based on contour representations and representative functions. Recently, a comprehensive overview of classification and shape matching methods in the study of archaeological ceramics has been presented

*Corresponding author

by Wilczek *et al.* (2021). Procrustes analysis and generalized Procrustes analysis appear in the analysis of archaeological material as stand-alone methods, for example, in the work of Dryden (2000), where they are used to study the variability of landmark points in the study of shapes of a class of objects.

On the other hand, clustering methods in archaeological applications are scarce. One such method for shape contour clustering that uses representative functions for shape representation and the Euclidean distance to define the similarity measure is presented by Gilboa *et al.* (2004). Other contributions on the clustering of archaeological pottery include as well as those by Cao and Mumford (2002), Mumford (1991), Sharon and Mumford (2006), and the references therein. In an earlier paper (Kaliszewska and Syga, 2018) we have proposed a similarity measure based on dynamic time warping (DTW) to compare representative functions as defined by Gilboa *et al.* (2004). The performed experiments showed that DTW could provide a promising new similarity measure. This observation formed a basis for introducing the similarity measures as defined in the present paper.

In the present investigation, we use the concept of shape introduced by Kendall (1977; 1989), Procrustes analysis (PA) and DTW to propose new weighted similarity measures. Up to our knowledge, this approach has not been used yet in the context of clustering archaeological pottery. It is worth mentioning that our approach can be applied to cluster any 3D objects of revolution for which the revolution axis and the 2D section are known.

1.2. Methods. The concept of shape we adopt, and a formal basis of shape analysis in the form of the so-called Kendall shape goes back to a series of publications by Kendall (1977; 1989) (cf. Goodall, 1991). Kendall's shape analysis in turn is inspired by shape theory introduced by Karol Borsuk in the 1960s and 1970s (Borsuk and Dydak, 1980) and is motivated by applications in archaeology. For a systematic exposition of the topic, see the monograph by da Fontoura Costa and Cesar (2010).

Informally, the shape of an object X is defined as all the geometrical information that remains when location, scale and rotational effects are filtered out from the object X, or, in other words, the shape is the geometry of an object X modulo position, orientation and size. In consequence, the shape space is the space of equivalence classes defined by a given class of objects X.

Such an understanding of shape gives rise to the concept of the *Procrustes distance* or *Procrustes distance measure* (see, e.g., Goodall, 1991) between the objects X_1 and X_2 as the distance of X_1 to the equivalence class defined by X_2. The Procrustes distance is not a distance in the formal meaning of the term. Based on the Procrustes distance, Procrustes-type similarity measures between shapes are proposed.

In numerous applications, PA and Procrustes-type similarity measures are combined with other similarity measures, e.g., by Hosni *et al.* (2018), combining PA with the principal component analysis (PCA) is proposed to investigate a 3D gait recognition problem (Eguizabal *et al.*, 2019); PA is also combined with the DTW to investigate surgery task classification (Albasri *et al.*, 2019).

DTW is devised to compare time series. More about this topic can be found in the works of Aronov *et al.* (2006) and Efrat *et al.* (2007). In general, similarity measures devised with the help of DTW can be applied to any objects composed of linearly parametrized (ordered) elements (components).

In our investigations, we define similarity measures by combining Procrustes based similarity measures with DTW similarity measures. The computational experiments confirm the efficiency of the proposed approach.

2D cross-sections are represented through their contours, i.e., 2D curves that are boundaries of these 2D cross-sections. In archaeological applications, those 2D curves are generated from technical drawings. This means that their position on the plane is not random and is determined by a set of strict rules. Consequence, the problem of a possible misalignment is not as crucial as in other applications, (e.g., Sangalli *et al.*, 2010; 2012). By applying the boundary-based representation of cross-sections, the problem of clustering 3D objects of revolution is reduced to the clustering of 2D curves with a given revolution axis with respect to shape and size. The problem of shape and size analysis of 2D curves and functions is gaining increasing interest due to many important applications, such as cardiovascular analysis (Sangalli *et al.*, 2010), finance interest rates (Kanevski and Timonin, 2010), nuclear industry (Auder and Fischer, 2012). Important to note are the advances in 3D surface reconstruction from images (Kotan *et al.*, 2021).

A specific feature of the clustering of archaeological pottery is that the contours should be classified with respect to subtle differences in shapes and size. Hence, the adopted similarity measures and the resulting clusters should reflect those subtle differences as precisely as possible. The subtle differences are of great importance in the clustering of contours, as a relatively small change in the contour might have a great impact on the shape of the respective 3D object of revolution.

1.3. Aim and contribution. The main aim of the present investigation is to propose weighted similarity measures, based on PA and DTW, for the clustering of 2D curves defining boundaries of 2D cross-sections of 3D objects of revolution, with respect to shape and size.

Depending upon the particular case, the contribution of shape and size in the overall similarity measure is adjusted by a proper choice of weights.

The secondary aim is to apply these similarity measures to the automatic generation of typologies (clusters) for archaeological pottery fragments. Clustering archaeological pottery is considered a process that may depend, to some extent, upon subjective judgements. This subjective judgement can be reflected in the choice of weights in the similarity measure.

The contribution of the paper is as follows:

- By using the PA (Eqn. (6)) and DTW (Eqn. (5)), we propose two novel curve-based pair-wise formulas to compare two curves:
  - the direct composition similarity measure DC (cf. Eqn. (9)),
  - scale (size) factor γ^* (cf. Eqn. (6)).
- We use the direct composition similarity measure DC (Eqn. (9)), scale similarity measure γ^* (Eqn. (8)) and the Procrustes similarity measure PA (Eqn. (6)) to generate various similarity matrices with the help of weights and normalization. By introducing weights and normalizations we get the following similarity matrices for which we perform our experiments:
  1. PSM (Eqn. (15)), the Procrustes similarity matrix, emphasizes the variations in the shape of curves in data set,
  2. DCM (Eqn. (16)), the direct composition matrix, emphasizes the variations in shape and size of curves in data set,
  3. SCM (Eqn. (17)), the scale component matrix, emphasizes only the variations in the size of curves in data set,
  4. WPSM (Eqn. (18)), the weighted Procrustes and scale component matrix, emphasizes the variations in shapes and size of curves in data set, by changing the weights we can decide which feature we want to highlight in a given experiment,
  5. WNDCSM (Eqn. (19)), the weighted direct composition and scale component matrix where direct composition values are normalized. Since the direct composition values are usually much higher than the values of a scale component, we use normalization.
- We apply the proposed similarity measures to the generation of typologies of archaeological pottery. The computational experiment is based on several data sets, some of them being real-life data; larger data sets are generated automatically by augmentation procedures.

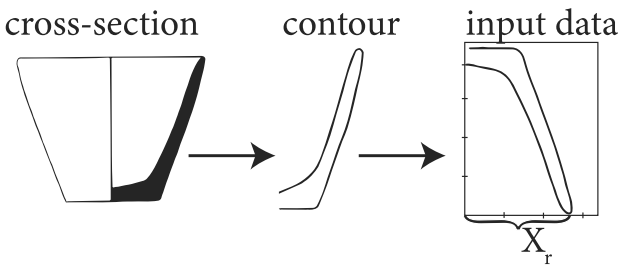

Fig. 1. Transformation of a section of rotationally symmetric object into a 2D input data curve.

1.4. Organization of the paper. In Section 2 we describe the class of investigated objects and we discuss the representations of 2D shapes (contours). In Section 3 we present basic elements of PA and DTW and we define the direct composition similarity measure DC. In Section 4 we introduce similarity measures (a)–(e) and present a general algorithm. In Section 5 we describe the archaeological objects on which we test our method. In Section 6 we describe the numerical experiments conducted and the data used. In Section 7 we discuss the obtained results. Section 8 concludes the paper.

2. Class of investigated objects

Geometrically, 3D objects of revolution is a solid obtained by rotating a plane curve around a straight line (an axis of revolution) that lies on the same plane. We investigate 3D objects of revolution obtained by rotating their 2D cross-sections (not necessarily curves) around an axis of revolution. Boundaries of 2D sections are called contours. We assume that each contour is a 2D open curve $\alpha \subset \mathbb{R}^2$. Each curve is a function $\alpha : [t_0, t_1] \to \mathbb{R}^2$ represented as

$$\alpha(t) := (x(t), y(t)), \quad t \in [t_0, t_1], \tag{1}$$

where $x : [t_0, t_1] \to \mathbb{R}$, and $y : [t_0, t_1] \to \mathbb{R}$.

We assume that

A1. the curves are contours of cross-sections of objects of revolution with the revolution axis OY and are all located in the first quadrant of the plane ($x \geq 0$, $y \geq 0$);

A2. the lowermost point (points), i.e., points with the smallest value of the coordinate y of a given curve are forced to lay on the OX axis (i.e., its y coordinate equals zero), and the point $(x_r, 0)$, where x_r is the smallest value of the x-coordinate corresponding to $y = 0$ defines the *radius r of the 3D object of revolution* (see Fig. 1),

$$r := x_r; \tag{2}$$

A3. the uppermost point (points), i.e., points with largest values of the coordinate y, are assumed to lie on the revolution axis OY.

Remark 1. Let us note that that Assumption A1 is not limiting. In case we classify objects with various revolution axes, we can always rotate (and shift) an object so as to satisfy Assumption A1. Assumptions A2 and A3 refer to the situation which is natural in the archaeological application, i.e., the locations of contours on the plane are not random and defined as in Assumptions A2 and A3. In the case where the locations of contours are random, some position normalization is needed, e.g., one can apply position normalization with which Procrustes analysis typically starts.

2.1. Shape representations. From the digitized cross-section the boundary discrete curve (contour) is extracted via standard techniques of contour extraction.

Contours are smoothed by the Savitzky–Golay filtering to retain the original shape of a profile as much as possible.

In further analysis, we consider discrete curves, i.e., a curve α is a pair of vectors,

$$\alpha := (x(i), y(i)), \quad i = 1, \ldots, k_\alpha. \tag{3}$$

Consequently, in the sequel, all data are vectors of different dimensions. When the original data are acquired in the form of bitmaps (e.g., scanned hand-made drawings), the accuracy of discretization depends on the resolution of the scanned image.

3. Similarity measures

We start by recalling PA and DTW similarity measures. For more details, see, e.g., the works of Müller (2007), Efrat *et al.* (2007) or Pizarro and Bartoli (2011).

Next, in Section 3.3, we define a new similarity measures based on PA and DTW, which is called direct composition DC.

3.1. Dynamic time warping (DTW). In view of the adopted contour representations, the compared curves X and Y are represented by sequences $X := (x_1, x_2, \ldots, x_N)$ of length $N \in \mathbb{N}$ and $Y := (y_1, y_2, \ldots, y_M)$ of length $M \in \mathbb{N}$, respectively, i.e., $X \in \mathbb{R}^N, Y \in \mathbb{R}^M$.

To make a comparison between these sentences, the DTW algorithm aligns them by applying the following procedure. First, the definition of a warping path is introduced.

Definition 1. (*Müller, 2007, Definition 4.1*) An (N, M)-warping path (or simply referred to as warping path if N and M are clear from the context) is a sequence $p = (p_1, \ldots, p_L)$ with $p_\ell = (n_\ell, m_\ell) \in \{1, \ldots, N\} \times \{1, \ldots, M\}$ for $\ell \in \{1, \ldots, L\}$ satisfying the following three conditions:

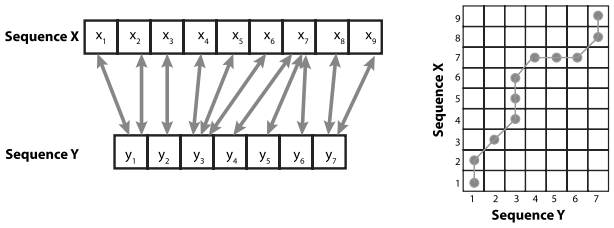

Fig. 2. Example of an alignment and a warping path p.

(i) boundary conditions: $p_1 = (1, 1)$ and $p_L = (N, M)$;

(ii) the monotonicity condition: $n_1 \leq n_2 \leq \cdots \leq n_L$ and $m_1 \leq m_2 \leq \cdots \leq m_L$;

(iii) the step size condition: $p_{\ell+1} - p_\ell \in \{(1, 0), (0, 1), (1, 1)\}$ for $\ell \in \{1, \ldots, L-1\}$.

Hence, an (N, M)-warping path $p = (p_1, \ldots, p_L)$ is defined by an alignment between two sequences $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_M)$ by assigning the element x_{n_l} of X to the element y_{m_l}, $l \in \{1, \ldots, L\}$ of Y. Figure 2 shows how an alignment is transformed into a warping path.

Now we define a cost measure which allows finding the best warping path. In the following, by F we denote the feature space, i.e., $x_n, y_m \in F$ for $n \in \{1, \ldots, N\}$, $m \in \{1, \ldots, M\}$. To compare two different elements $x, y \in F$, a local cost measure is defined to be a function $c : F \times F \to \mathbb{R}_+$. Evaluating the local cost measure for each pair of elements of the sequences X and Y, we obtain the cost matrix $C \in \mathbb{R}^{N \times M}$ defined by $C(n, m) := c(x_n, y_m)$.

The total cost $c_p(X, Y)$ of a warping path p between X and Y with respect to the local cost measure c is defined as

$$c_p(X, Y) := \sum_{\ell=1}^{L} c(x_{n_l}, y_{m_l}), \tag{4}$$

where, as in Definition 1, $p = (p_1, \ldots, p_L)$ with $p_\ell = (x_{n_\ell}, y_{m_\ell})$, $\ell \in \{1, \ldots, L\}$. An optimal warping path between X and Y is a warping path p^* having minimal total cost from among all possible warping paths. The DTW similarity measure $\mathrm{DTW}(X, Y)$ between X and Y is then defined as the total cost of p^*:

$$\begin{aligned} &\mathrm{DTW}(X, Y) \\ &:= c_{p^*}(X, Y) \\ &= \min\{c_p(X, Y) \mid p \text{ is an } (N, M)\text{-warping path}\}. \end{aligned} \tag{5}$$

The DTW similarity measure is akin to the Fréchet distance (see Aronov *et al.*, 2006; Müller, 2007). In our experiments we use the DTW function from the Matlab Signal Processing Toolbox.

3.2. Procrustes analysis. Procrustes analysis (PA) allows us to perform a statistical shape analysis and to compute a similarity measure for any given pair of 2D curves represented in the discretised form (3) (Gower and Dijksterhuis, 2004; Pizarro and Bartoli, 2011).

Let us consider the set of discrete curves $\alpha := (x(j), y(j))_{j=1}^{k} \in (\mathbb{R} \times \mathbb{R})^k$ (regarded as row vectors). Let $\mathcal{T}$ be the set of similarity transformations defined on $(\mathbb{R} \times \mathbb{R})^k$. $\mathcal{T}$ is given in the form of triples $T := (\gamma, R, t)$, where:

- $\gamma \in \mathbb{R}_{++}$ represents a (uniform) scaling factor and defines the scaling transformation $\gamma : (\mathbb{R} \times \mathbb{R})^k \to (\mathbb{R} \times \mathbb{R})^k$, $\gamma(\alpha) = [\gamma(x(1), y(1)), \ldots, \gamma(x(k), y(k))]$,
- R is a 2×2 rotation matrix of the angle θ, and defines the rotation transformation $R : (\mathbb{R} \times \mathbb{R})^k \to (\mathbb{R} \times \mathbb{R})^k$,
$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$
$R(\alpha) = R\alpha$,
- $t \in (\mathbb{R} \times \mathbb{R})^k$ is a translation vector, which defines the transformation $t : (\mathbb{R} \times \mathbb{R})^k \to (\mathbb{R} \times \mathbb{R})^k$, $t(\alpha) = t + \alpha$, with components $t(j) \in \mathbb{R} \times \mathbb{R}$, $j = 1, \ldots, k$.

For the input discrete curves $\alpha_1 := (x_1(j), y_1(j)) = (z_1(j))$, $j = 1, \ldots, k$, $\alpha_2 := (x_2(j), y_2(j)) = (z_2(j))$, $i = 1, \ldots, k$, the Procrustes similarity measure $\mathrm{PA}(\alpha_1, \alpha_2)$ is defined as

$$\mathrm{PA}(\alpha_1, \alpha_2) := \inf_{T \in \mathcal{T}} \sum_{j=1}^{k} \|z_1(j) - (\gamma(R\alpha_2)(j) + t(j))\|_2^2, \quad (6)$$

where $\| \cdot \|_2$ is the Euclidean distance in $\mathbb{R}^2$. The Procrustes similarity measure is not a distance since, in general $\mathrm{PA}(\alpha_1, \alpha_2) \neq \mathrm{PA}(\alpha_2, \alpha_1)$.

Equation (6) can be interpreted in terms of equivalence classes $[\cdot]$ of some equivalence relation. Namely, a (discrete) curve $\alpha := (x(j), y(j))_{j=1}^{k} \in (\mathbb{R} \times \mathbb{R})^k$ belongs to the equivalence class defined by $\alpha_1 = (x_1(j), y_1(j))_{j=1}^{k} \in (\mathbb{R} \times \mathbb{R})^k$, $\alpha \in [\alpha_1]$ iff $(x(j), y(j)) = \gamma \cdot R(x_1(j), y_1(j)) + t(j)$, $t(j) \in \mathbb{R} \times \mathbb{R}$, $j = 1, \ldots, k$, where $\gamma > 0$ is a scaling factor, R is a 2×2 rotation matrix of the angle θ, as defined above, and $t \in \mathbb{R}^2$. In this way an equivalence relation $\equiv$ is defined as

$$\alpha_1 \equiv \alpha_2 \Leftrightarrow (x_1(j), y_1(j)) = \gamma \cdot R(x_1(j), y_1(j)) + t(j), \quad (7)$$

$j = 1, \ldots, k$ for some scaling factor $\gamma > 0$, rotation matrix R and translation vector t with components $t(j) \in \mathbb{R} \times \mathbb{R}$, $j = 1, \ldots, k$.

In our experiments we use the Procrustes function from the Matlab Statistics and Machine Learning Toolbox.

3.3. Direct composition. In this section we introduce a new similarity measure which is based on composition of PA and DTW.

Let two (discrete) curves $\alpha_1, \alpha_2 \in (\mathbb{R} \times \mathbb{R})^k$ be given. The objective function of the optimization problem (6) is strictly convex and bounded from below. By solving problem (6), we obtain the minimal value $\mathrm{PA}(\alpha_1, \alpha_2) = d$ and the curve $Z = T^*\alpha_2$, where $T^* \in \mathcal{T}$ is an optimal solution to problem (6), i.e.,

$$\begin{aligned} \mathrm{PA}(\alpha_1, \alpha_2) = d &= \inf_{T \in \mathcal{T}} \|\alpha_1 - T\alpha_2\|^2 \\ &= \|\alpha_1 - Z\|^2 = \|\alpha_1 - T^*\alpha_2\|^2, \end{aligned} \quad (8)$$

and $Z := T^*\alpha_2 = \gamma^* R^* \alpha_2 + t^*$, where $\gamma^* > 0$ is the (optimal) scaling factor and R^* is the (optimal) rotation matrix, and t^* is the (optimal) translation vector. In our experiments the curve Z is produced by the `procrustes` function from the Matlab Statistics and Machine Learning Toolbox.

With the help of the curve Z we define a new similarity measure, called the *direct composition DC* and given as

$$\mathrm{DC}(\alpha_1, \alpha_2) = \mathrm{DTW}(\alpha_1, Z). \quad (9)$$

Direct composition DC combines global shape similarity measure PA with DTW which takes into account the local shape variability as well.

The motivation for the combination of PA and DTW measures stems from our previous works (e.g., Kaliszewska and Syga, 2018), which have brought to light a necessity to introduce a similarity measure that would take into account the subtle nature of the differences between the investigated shapes. The possibility of introducing weights that account for the expert's knowledge, and bring his/her expertise into the process is also of paramount importance.

We consider PA and DTW methods complementary, as they address different characteristics of the shape. PA is, in our case, a robust tool for the alignment of shapes and the detection of the overall shape similarity. It performs very well when used to detect scale and orientation invariant similarity between significantly different shapes. On the other hand DTW is well suited to detect subtle changes in objects that could be generally defined as similar. It fails, however, to quantify the similarity when the shapes differ in attributes such as rotation, or scale.

The idea of combining PA and DTW in similarity measures related to the analysis of temporal alignment of human motion already appeared, e.g., in the work of Zhou and De la Torre (2012).

4. Proposed algorithm for clustering curves

In the present section we propose an algorithm for clustering 2D contours of cross-sections of 3D objects of revolution. We assume that the given set of contours consists of $n \geq 2$ elements. Let $i = 1, \ldots, n$ and $j = 1, \ldots, n$, denote the i-th and j-th contours, respectively. First we calculate the "distances" between two contours i and j, $i \neq j$ by using the following three formulas:

$$\mathrm{pa}(i,j) = \max\{\mathrm{PA}(i,j), \mathrm{PA}(j,i)\}, \tag{10}$$

$$\mathrm{dc}(i,j) = \max\{\mathrm{DC}(i,j), \mathrm{DC}(j,i)\}, \tag{11}$$

$$\gamma(i,j) = 1 - \min\{\gamma^*(i,j), \gamma^*(j,i)\}, \tag{12}$$

where $\gamma^*(i,j)$ is the optimal scaling factor obtained from $\mathrm{PA}(\alpha_i, \alpha_j)$ (cf. Eqn. (8)) and $\gamma^*(j,i)$ is the optimal scaling factor obtained from $\mathrm{PA}(\alpha_j, \alpha_i)$ (cf. Eqn. (8)). Whenever $i = j$, we set

$$\mathrm{pa}(i,i) = 0, \quad \mathrm{dc}(i,i) = 0, \quad \gamma(i,i) = 0.$$

The numbers calculated by the formulas (10), (11) and (12) may have very different ranges; hence, to ensure the comparability, we use the following normalization formulas for 'dc' and γ

$$\mathrm{ndc}(j) = \frac{1}{\max_{i=1,\ldots,n} \mathrm{dc}(i,j)}, \qquad j = 1, \ldots, n,$$

$$\mathrm{n}\gamma(j) = \frac{1}{\max_{i=1,\ldots,n} \gamma(i,j)}, \qquad j = 1, \ldots, n.$$

When the normalization is not needed, we set $\mathrm{ndc}(j) = \mathrm{ndc} = 1$, for all $j = 1, \ldots, n$ or $\mathrm{n}\gamma(j) = \mathrm{n}\gamma = 1$, for all $j = 1, \ldots, n$.

4.1. Similarity matrix. A similarity matrix is formed by calculating a similarity measure between any two contours i and j. Let $\mu, \lambda, \omega \in \mathbb{R}$ be given numbers (weights). We calculate the similarity measure (SM) as follows:

$$\begin{aligned} &\mathrm{SM}_{ij}(\mu, \lambda, \omega, \mathrm{ndc}, n\gamma) \\ &\quad := \mu \cdot \mathrm{pa}(i,j) + \lambda \cdot \mathrm{ndc}(j) \cdot \mathrm{dc}(i,j) \\ &\qquad + \omega \cdot \mathrm{n}\gamma(j) \cdot \gamma(i,j). \end{aligned} \tag{13}$$

By calculating the number SM_{ij} for every pair of contours i and j ($i = 1, \ldots, n$ and $j = 1, \ldots, n$) from the data set, we get the similarity matrix

$$\mathrm{SM}(\mu, \lambda, \omega, \mathrm{ndc}, \mathrm{n}\gamma), \tag{14}$$

which is symmetric and has zeros on the main diagonal.

This general similarity measure allows us to preform various experiments. Changing the values of numbers $\mu, \lambda, \omega, \mathrm{ndc}, \mathrm{n}\gamma$, we get the following similarity matrices:

(i) the Procrustes similarity matrix

$$\mathrm{PSM} = \mathrm{SM}(1,0,0,1,1); \tag{15}$$

(ii) the direct composition matrix

$$\mathrm{DCM} = \mathrm{SM}(0,1,0,1,1); \tag{16}$$

(iii) the scale component matrix

$$\mathrm{SCM} = \mathrm{SM}(0,0,1,1,1); \tag{17}$$

(iv) the weighted Procrustes and scale component matrix

$$\mathrm{WPSM} = \mathrm{SM}(\mu,0,\omega,1,1); \tag{18}$$

(v) the weighted direct composition and scale component matrix where direct composition values are normalized

$$\mathrm{WNDCSM} = \mathrm{SM}(0,\lambda,\omega,\mathrm{ndc},1); \tag{19}$$

(vi) the weighted direct composition and scale component matrix, where both direct composition values and scale component values are normalized,

$$\mathrm{WNDCNSM} = \mathrm{SM}(0,\lambda,\omega,\mathrm{ndc},\mathrm{n}\gamma). \tag{20}$$

For clustering we apply hierarchical clustering. We use three standard methods to measure the distance between clusters, i.e., single linkage, average linkage and weighted average linkage (see Tables 3 and 4).

4.2. Proposed algorithm. The general algorithm is presented below as Algorithm 1. By changing the values of weights μ, λ, ω and ndc, $\mathrm{n}\gamma$, we obtain different similarity measures, as described above, and consequently, different clustering results.

Algorithm 1. Curve clustering.

Step 1. Choose the values of weights μ, λ, ω and ndc, $\mathrm{n}\gamma$.

Step 2. Prepare the input data (vectors) as described in Section 2.

Step 3. Calculate the similarity between any two vectors i and j using Eqn. (13) and generate the similarity matrix via (14).

Step 4. Perform clustering on the basis of the similarity matrix (14) generated in Step 3 by standard hierarchical algorithms (Matlab toolbox) and generate the dendrogram.

5. Application: Archaeological objects

In this section, we describe the archaeological objects to which we apply our algorithm. These objects are archaeological pottery fragments. When produced on the pottery wheel, vessels are rotationally symmetric objects of revolution. We limit our attention to this kind of vessels. Consequently, a cross-section and the location of the revolution axis are enough to convey the shape of the whole vessel.

In archaeological practice, the vessels or vessel fragments are described by their cross-sections. These cross-sections are hand-made technical drawings; they are made according to a set of rules and use conventions and a visual language to maintain standards and convey as much information about the vessel as possible, without the need to supplement it with text.

Classification of pottery obtained through excavation is a form of organizing the material to conclude the investigated sets and deal with the spatio-temporal diversity of the pottery. Such classifications are traditionally performed manually and depend heavily on the expert's knowledge and experience and, as a result, are prone to being biased.

In our application, the curves, as defined in Section 2, are contours of cross-sections of vessels uncovered in the course of archaeological work. From the technical drawings, the axis of revolution of each vessel can easily be deciphered. As mentioned in Section 2, to perform the clustering (generate clusters) we standardize the location of the contours to the first quadrant of $\mathbb{R}^2$ and standardize the position of contours by considering their upside-down versions. More precisely, in the present study we consider contours (curves) for which we know the coordinate x_r, see Fig. 1 and for which the uppermost point (points) lies on the revolution axis OY to satisfy Assumptions A1–A3 of Section 2. This means that we classify vessel fragments for which the full cross-section is known. This does not mean that the full vessel is known. There exist archaeological techniques which allow reconstructing the exact full cross-section of a vessel, even in the case where the vessel is considerably damaged.

During the process of clustering, the vessels can be divided into groups based on many features. With the help of our clustering algorithm, we divide given archaeological pottery sets with respect to shape and size. We limit our attention to the clustering of those pottery fragments for which the complete cross-section can be extracted.

6. Experiment

To test Algorithm 1, we have designed and conducted a series of experiments related to the clustering of archaeological pottery. For that purpose, we have chosen 6 data sets, each with 36 to 51 elements. Set 2 is presented in Fig. 3. The profiles used for the experiments come from (Mountjoy, 1999) where real-life archaeological pottery fragments are classified by experts.

In Table 1 we summarize the features of Sets 1–6. The estimated number of clusters is given as a range, as it depends on the method of counting the clusters (clusters vs. subclusters), and on the expert's decision. The table also provides a general description of the properties for each set, listing how the elements vary in size, and how much variation in shape there is within each cluster.

The results of clustering obtained by Algorithm 1 are visualized in the form of dendrograms. The dendrogram's visual nature makes it inefficient when it comes to comparing results for large data sets.

In order to evaluate the quality of clustering, we use the cophenetic correlation coefficient to calculate a score for each clustering method and each set, see Tables 3 and 4. The cophenetic correlation coefficient is a measure of how closely the dendrogram represents pairwise dissimilarities between the objects. It was originally introduced by Sokal and Rohlf (1962) as a method for comparing dendrograms resulting from numerical taxonomic research. The cophenetic

Table 1. Description of data-sets used in the experiment.

Set no.	No. of elements	Estimated no. of clusters	Characteristics
1	41	7–9	2 size groups, little variation in shape in clusters
2	41	6–10	2 size groups, at least 2 clusters with significant variation in shape in clusters
3	45	6–8	3 size groups, some variation in shape in clusters
4	36	5–7	all similar size, 4 basic shape groups, significant variation in shape in the clusters
5	51	8–10	5 size groups, significant differences in size between groups, some variation in shape in the clusters
6	40	5–8	similar size, 2 size groups, little variation in shape in clusters.

correlation coefficient (CCF) is defined as follows:

$$c := \frac{\sum_{i<j}(Y_{ij} - y)(Z_{ij} - z)}{\sqrt{\sum_{i<j}(Y_{ij} - y)^2 \sum_{i<j}(Z_{ij} - z)^2}}, \qquad (21)$$

where Y_{ij} is the distance between objects i and j in the similarity matrix Y, and Z_{ij} is the cophenetic distance, that is the distance between two observations i and j represented in a dendrogram by the height of the link at which those two observations are first joined. Here y and z are the means of Y and Z, respectively. The closer this value is to 1, the higher the quality of the solution. For a detailed discussion of the cophenetic correlation coefficient. The reader is referred to Farris (1969).

6.1. Augmentation. The sets of archaeological pottery excavated by archaeologists are usually large and very diverse in terms of shape. Traditionally, the clustering (generating of typologies) is done manually by experts. Consequently, in practice, only a small part of the archaeological material is clustered and published, as it is the case in (Mountjoy, 1999).

Hence, the real-life data sets which are available consist of 30–50 elements. This size does not provide a sufficiently large sample to evaluate the performance of the algorithm. This was remedied by an automated augmentation of the data set. To ensure the comparability of the results, we used the same data-sets (see Table 1). Each element of the set was subjected to 6 different transformations, resulting in sets with 252 to 357 elements (including the original objects). For these transformations, we have chosen a warp-type transformation, as it can produce slight changes that are characteristic of our investigated elements. The warp transformations were grid-based and used a Bézier curve based grid to perform the transformations. Each of the six transformations targeted one part of the image (very generally described as top, middle, and bottom), and either bevelled the image towards the left or right edge of the image in the specific section. The transformations were chosen to be radical enough to modify the original image, to generate new groups, without distorting the image to a degree such that it does not fulfil the original requirements (Fig. 4) (see Section 2). In the same way as for the original data sets, the proposed methods have been tested on all six augmented sets, including different methods of the linkage algorithm. The results are included in Table 4.

7. Results and a discussion

The evaluation of the results is based on the dendrograms obtained for each set. The cophenetic correlation coefficient (Tables 3 and 4) is used as an additional measure, quantifying the quality of the dendrogram.

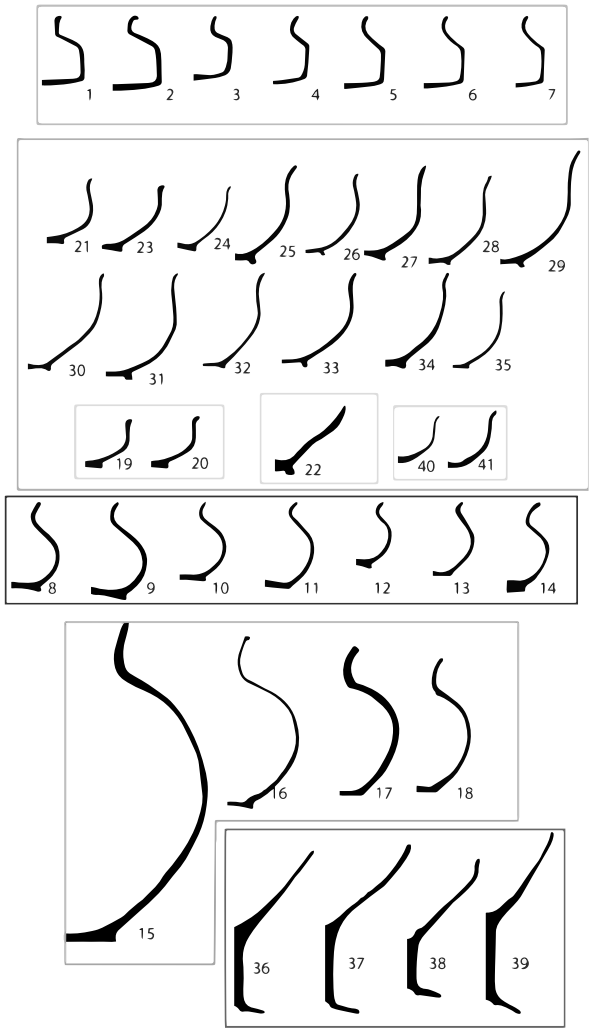

Fig. 3. Set 2.

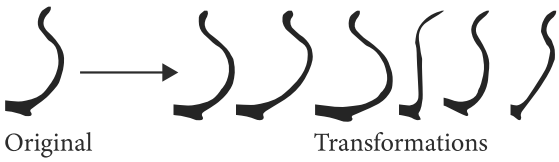

Fig. 4. Illustration of the transformations applied in the augmentation process.

A crucial point in this kind of clustering problem is the definition of a similarity measure which should reflect the similarity as it is perceived by the expert. The following properties of the proposed similarity measures can be observed based on the experiment:

1. The first type of mismatch occurs when two contours are very similar or even identical on a long portion of the curve and differ significantly on a relatively short portion of the curve. In such a case, a high

similarity, i.e., a small distance between curves on a given portion *overpowers* the similarity measure calculated for the two objects, and fails to take into account the difference between the objects. This may lead to a false-positive result, with two contours perceived as "different", and belonging to different clusters, being grouped.

2. The second type of mismatch in the dendrogram may occur when the weights are improperly selected. This means that one of the aspects of the clustering, the shape or scale, takes precedence over the other and leads to two contours being paired or separated as, e.g., their scale is very similar, despite not exhibiting similarity in shape.

3. The third type of mismatch stems from the difference between the perceived similarity (experts evaluation) and the calculated similarity measure. The in-depth discussion of this topic is beyond the scope of this paper; however, it will be demonstrated below in the discussion of the results obtained for the original data sets (Section 7.1).

Those types of mismatches do not point to a fault in the method. They are an inherent property of automated shape detection, especially of Type 1. Type 2 can be remedied by the appropriate selection of weights by the expert. These types of mismatches highlight the discrepancies between manual *handmade* classification, where the consistency of decision making is hard to maintain throughout the process and personal biases come into play, while with an automated process the same criteria are applied throughout. The same applies to Type 3.

7.1. Results obtained for the original data sets. The analysis of the dendrograms yielded the best results for the combination of the Mix method and the scale component. The exact choice of weights depends on the nature of the set. The WNDCSM similarity measure turned out to be the best choice for Sets 1, 3, 4, 5, and 6. This is due to the differences in shape coinciding with the differences in scale. That is, there are few or no contour fragments that overlap in similar shape and similar size, although belonging to different clusters. For Set 2 the best combination is weight 3/4 for Procrustes and 1/4 for the scale component. For this set, the weight of the scale component had to be reduced in favour of the shape component PA of the measure. This is due to several contours being *close*, both in terms of shape and size, and the need to put more stress on the shape component, as this is our primary goal.

The results of an expert-based evaluation of the obtained dendrograms are summarized in Table 2. The table presents the total number and percentage of the elements that are, in the expert's opinion, correctly classified. As the tool is meant to be of assistance to experts, we choose an expert-based evaluation for our results. An automated or semi-automated assessment of results based on the dendrograms, and the choice of a correct similarity threshold is itself a matter of ongoing research (e.g., Vogogias *et al.*, 2016; Siminski, 2021).

A visualization of the expert's assessment is presented in Figs. 5–8. For dendrograms that have a smaller similarity measure values between objects and groups, the difference between the single cut and the multi-level cut is not significant. For Set 2 (Figs. 5 and 6), for the single-level cut we obtain an accuracy of 80% (compared with 92,68 % with the multi-level cut—Table 2). The dendrogram has the CFF value of 0.98, which means the dendrogram structure is very good. For Set 4 (Figs. 7 and 8), the CFF is significantly lower, with 0.89. This dendrogram has higher values of similarity measures between objects. Using the multi-level cut method, we obtain the result of 91.67 %. The single-level cut method results in only 83.34 % of accuracy. The amount of correctly classified elements and the resulting percentages, as presented below, were calculated based on the expert's judgement. In calculations of the results of Table 2, the cuts of the dendrogram were made on different levels (multi-level cuts), i.e. different thresholds (cut levels) selected for splitting the dendrogram (i.e., obtaining clusters) were chosen by the expert and are based on the direct data inspection.

The expert-based evaluation is supplemented by the CCF (Table 3) calculated for each set and each method in the experiments performed. A discussion of the obtained results, summarized in Table 2 and Table 3, is presented below, in Section 7.2.

7.2. Discussion of the results. One of the results, for the WNDCSM similarity measure, obtained through the experiments is presented in Figs. 5 and 6.

Based on Table 2, we notice that the best results were obtained for the method WNDCNSM ($SM(0, 3/4, 1/4, \text{ndc}, \text{n}\gamma)$ for Sets 1, 2, 5, and 6. This can be attributed to the character of the sets, as summarized in Table 1. Sets 3 and 4 show significant variation in shape within the clusters, which might result in some under-performance of the method. However, the results are still fairly high. Combining together the results in Tables 2 and 3, we see that all proposed methods perform fairly well, as the performance is on average 85% (Table 2) for the three proposed methods, and all values for CCF are at 0.75 and above (Table 3), with most of them being above 0.9. As stated before, the values of the CCF do not quantify the clustering results, but rather the structure of the dendrogram, and in this sense, the performance of the algorithm.

The results, as described above prove the efficiency

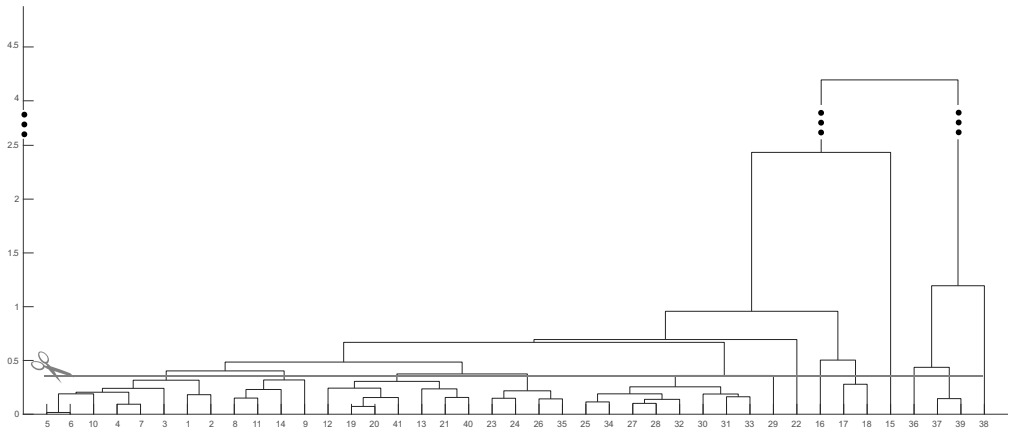

Fig. 5. Single-level cut through the dendrogram obtained for Set 2 according to formula 20. CCF $= 0.98$, correctly classified objects $= 80\%$.

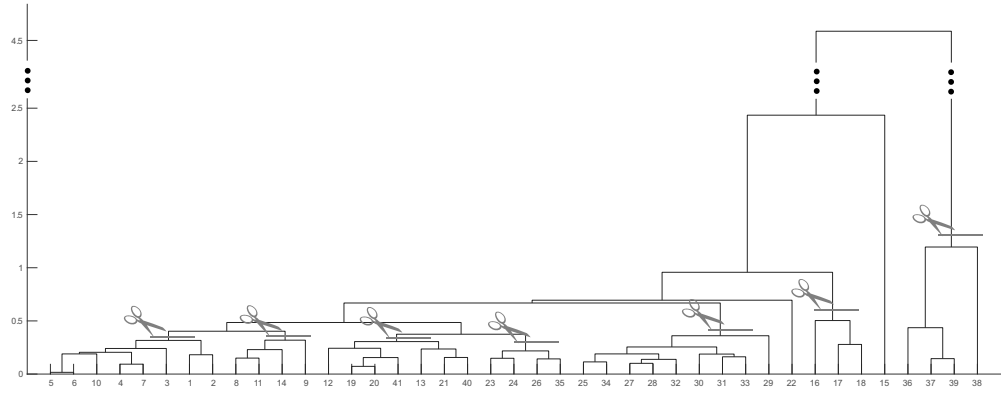

Fig. 6. Multi-level cut through the dendrogram obtained for Set 2 according to Eqn. (20). CCF $= 0.98$, correctly classified objects $= 92.68\%$.

of the WNDCSM algorithm, and the introduction of the additional DWT distance component to the similarity measure. The results point to a double function of the WPSM similarity measure and all the subsequent combinations, in the sense of Procrustes analysis being responsible for a robust detection of the "global shape" while DTW detects subtle differences between shapes, and allows taking into account of the scaling factor. Furthermore, the scaling component accounts for the difference in size, between the investigated objects. Such a composition allows for a fuller analysis of the given sets.

7.3. Results obtained for the augmented data-sets. The augmented data sets were clustered by the algorithm in Section 4.2, the same as the original data. Due to the method of augmentation (Section 6.1), each set increased in the number of shapes; however, no new shape groups were created, as the augmentation did not target size, but solely the shape. Due to the size of the resulting dendrograms, we omit here the detailed discussion. However, the satisfactory performance of the Algorithm 1 for the original sets, and the high values of the cophenetic correlation coefficient give insight into the performance of the algorithm for the augmented sets. The augmentation process will be applied in future research to increase the size of data sets to employ machine learning algorithms.

8. Conclusions

We have presented the concept and performance of the algorithm (see Section 4.2) for clustering a class of 2D objects, with application in archaeology. The performed experiment shows that the proposed algorithm provides satisfactory results. Even though the results obtained are promising, many problems related to the clustering of archaeological pottery remain open.

Conceivable extensions of the research presented

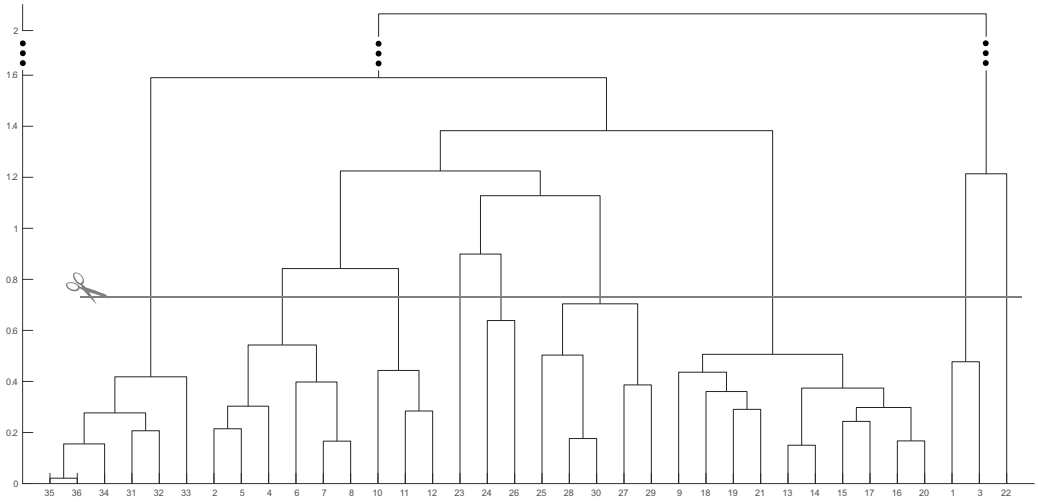

Fig. 7. Single-level cut through the dendrogram obtained for Set 4 according to Eqn. (19). CCF $= 0.98$, correctly classified objects $= 83.34\%$.

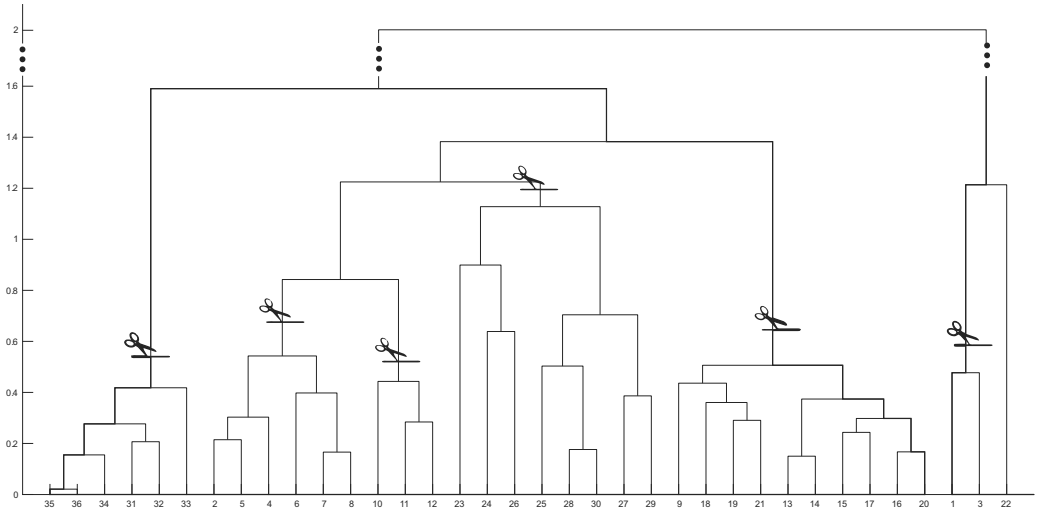

Fig. 8. Multi-level cut through the dendrogram obtained for Set 4 according to Eqn. (19). CCF $= 0.89$, correctly classified objects $= 91.67\%$.

here encompass investigating the possible advantages of using alternative shape representations and similarity measures. As part of our further research, we plan to investigate the case where only a fragment of the section is preserved. Another problem arises when the investigated pottery vessel has a handle. As far as we know this problem has not yet been addressed in the literature.

References

Albasri, S., Popescu, M. and Keller, J.M. (2019). Surgery task classification using procrustes analysis, *48th IEEE Applied Imagery Pattern Recognition Workshop, AIPR 2019, Washington, USA,* pp. 1–6.

Aronov, B., Har-Peled, S., Knauer, C., Wang, Y. and Wenk, C. (2006). Fréchet distance for curves, revisited, *in* Y. Azar and T. Erlebach (Eds), *Algorithms—ESA 2006*, Springer, Berlin, pp. 52–63.

Auder, B. and Fischer, A. (2012). Projection-based curve clustering, *Journal of Statistical Computation and Simulation* **82**(8): 1145–1168.

Borsuk, K. and Dydak, J. (1980). What is the theory of shape?, *Bulletin of the Australian Mathematical Society* **22**(2): 161–198.

Cao, Y. and Mumford, D. (2002). Geometric structure estimation of axially symmetric pots from small fragments, *Signal Processing, Pattern Recognition, and Applications, Crete, Greece*.

Table 2. Percentage of correctly classified elements as evaluated by the expert for Sets 1–6. The number of elements in each set and the number of correctly classified elements are given in brackets. All scores have been calculated using the average distance between clusters.

Method	Set 1 [41]	Set 2 [41]	Set 3 [45]	Set 4 [36]	Set 5 [51]	Set 6 [40]
SM($\frac{1}{2}, 0, \frac{1}{2}, 1, 1$)	82.93 %[34]	78.05%[32]	75.56%[34]	75%[27]	69.23%[36]	87.5%[35]
SM($\frac{3}{4}, 0, \frac{1}{4}, 1, 1$)	85.36% [35]	90.24%[37]	**82.22%[37]**	86.11%[31]	78.43%[40]	85%[34]
SM($0, \frac{1}{2}, \frac{1}{2}$, ndc, 1)	**90.24%[37]**	80.49%[33]	77.78%[35]	86.11%[31]	78.43%[40]	82.5%[33]
SM($0, \frac{3}{4}, \frac{1}{4}$, ndc, 1)	87.8%[36]	90.24%[37]	77.78%[35]	**91.67%[33]**	78.43%[40]	87.5%[35]
SM($0, \frac{1}{2}, \frac{1}{2}$, ndc, n$\gamma$)	87.8%[36]	75.61%[31]	75.56%[34]	88.89%[32]	78.43%[40]	**92.5%[37]**
SM($0, \frac{3}{4}, \frac{1}{4}$, ndc, n$\gamma$)	**90.24%[37]**	**92.68%[38]**	80%[36]	86.11%[31]	**82.35%[42]**	**92.5%[37]**

Cohen-Addad, V., Kanade, V. and Mallmann-Trenn, F. (2018). Clustering redemption—Beyond the impossibility of Kleinberg's axioms, *Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montreal, Canada*, pp. 8526–8535.

Cohen-Addad, V., Kanade, V., Mallmann-Trenn, F. and Mathieu, C. (2017). Hierarchical clustering: Objective functions and algorithms, *Proceedings of the 2018 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), New Orleans, USA*, pp. 378–397.

Davidson, I. and Ravi, S.S. (2007). Intractability and clustering with constraints, *Proceedings of the 24th International Conference on Machine Learning, Corvalis, USA*, pp. 201–208.

Dryden, I.L. (2000). Statistical shape analysis in archaeology, *Spatial Statistics in Archaeology, Chieti, Italy*.

Efrat, A., Fan, Q. and Venkatasubramanian, S. (2007). Curve matching, time warping, and light fields: New algorithms for computing similarity between curves, *Journal of Mathematical Imaging and Vision* **27**(3): 203–216.

Eguizabal, A., Schreier, P.J. and Schmidt, J. (2019). Procrustes registration of two-dimensional statistical shape models without correspondences, *CoRR* abs/1911.11431.

Farris, J. (1969). On the cophenetic correlation coefficient, *Systematic Zoology* **18**(3): 279–285.

da Fontoura Costa, L. and Cesar, R.M. (2010). *Shape Analysis and Classification: Theory and Practice*, CRC Press, Boca Raton.

Gilboa, A., Karasik, A., Sharon, I. and Smilansky, U. (2004). Towards computerized typology and classification of ceramics, *Journal of Archaeological Science* **31**(6): 681–694.

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape, *Journal of the Royal Statistical Society B: Methodological* **53**(2): 285–339.

Gower, J.C. and Dijksterhuis, G.B. (2004). *Procrustes Problems*, Oxford University Press, Oxford.

Hosni, N., Drira, H., Chaieb, F. and Amor, B.B. (2018). 3D Gait recognition based on functional PCA on Kendall's shape space, *2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China*, pp. 2130–2135.

Hristov, V. and Agre, G. (2013). A software system for classification of archaeological artefacts represented by 2D plans, *Cybernetics and Information Technologies* **13**(2): 82–96.

Jain, A.K. (2010). Data clustering: 50 Years beyond k-means, *Pattern Recognition Letters* **31**(8): 651–666.

Kaliszewska, A. and Syga, M. (2018). On representative functions method for clustering of 2D contours with application to pottery fragments typology, *Control and Cybernetics* **47**(1): 85–108.

Kanevski, M. and Timonin, V. (2010). Machine learning analysis and modeling of interest rate curves, *ESANN 2010: European Symposium on Artificial Neural Networks—Computational Intelligence and Machine Learning, Bruges, Belgium*, pp. 47–52.

Kendall, D.G. (1977). The diffusion of shape, *Advances in Applied Probability* **9**(3): 428–430.

Kendall, D.G. (1989). A survey of the statistical theory of shape, *Statistical Science* **4**(2): 87–99.

Kleinberg, J. (2002). An impossibility theorem for clustering, *Proceedings of the 15th International Conference on Neural Information Processing Systems, NIPS'02, Vancouver, Canada*, p. 463–470.

Kotan, M., Öz, C. and Kahraman, A. (2021). A linearization-based hybrid approach for 3D reconstruction of objects in a single image, *International Journal of Applied Mathematics and Computer Science* **31**(3): 501–513, DOI: 10.34768/amcs-2021-0034.

Leski, J.M. and Kotas, M.P. (2018). Linguistically defined clustering of data, *International Journal of Applied Mathematics and Computer Science* **28**(3): 545–557, DOI: 10.2478/amcs-2018-0042.

Maiza, C. and Gaildart, V. (2005). Automatic classification of archaeological potsherds, *8th International Conference on Computer Graphics and Artificial Intelligence, 3IA'2005, Limoges, France*, pp. 11–12.

Müller, M. (2007). *Information Retrieval for Music and Motion*, Springer, Berlin/Heidelberg.

Mountjoy, P.A. (1999). *Regional Mycenaean Decorated Pottery*, Deutsches Archäologisches Institut, Berlin.

Table 3. Cophenetic correlation coefficient calculated for six investigated data-sets.

Method		Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Arithmetic mean
SM$(1,0,0,1,1)$								
	single	0.8564	0.9952	0.9396	0.6340	0.5906	0.7495	0.7942
	average	**0.9958**	0.9221	0.9194	0.7930	0.8141	0.8546	0.8832
	weighted	0.9189	**0.9956**	0.9047	0.7838	0.7692	0.8456	0.8696
SM$(0,1,0,1,1)$								
	single	0.8549	0.9872	0.9492	0.8325	0.8165	0.7894	0.8716
	average	0.9082	0.9892	0.9735	0.8446	0.8966	0.8328	0.9075
	weighted	0.9038	0.9886	0.9725	0.8426	0.8598	0.8172	0.8974
SM$(0,0,1,1,1)$								
	single	0.9760	0.8939	0.9964	0.8973	0.9155	0.9056	0.9308
	average	0.9792	0.9379	**0.9974**	0.9423	**0.9293**	0.9170	0.9505
	weighted	0.9774	0.8730	0.9961	0.9404	0.9284	0.8993	0.9358
SM$(\frac{1}{2},0,\frac{1}{2},1,1)$								
	single	0.9645	0.9747	0.9874	0.8816	0.9100	0.9467	0.9441
	average	0.9705	0.9815	0.9881	**0.9450**	0.9256	**0.9542**	**0.9608**
	weighted	0.9681	0.9803	0.9880	0.9414	0.9248	0.9532	0.9593
SM$(\frac{3}{4},0,\frac{1}{4},1,1)$								
	single	0.8963	0.9917	0.9446	0.8986	0.8915	0.9150	0.9229
	average	0.9287	0.9933	0.9450	0.9298	0.9068	0.9292	0.9388
	weighted	0.9268	0.9933	0.9428	0.9386	0.9061	0.9262	0.9390
SM$(0,\frac{1}{2},\frac{1}{2},\mathrm{ndc},1)$								
	single	0.9441	0.9670	0.9704	0.8023	0.8034	0.9110	0.8997
	average	0.9611	0.9583	0.9717	0.8862	0.8991	0.9199	0.9327
	weighted	0.9588	0.9724	0.9708	0.8763	0.8954	0.9188	0.9321
SM$(0,\frac{3}{4},\frac{1}{4},\mathrm{ndc},1)$								
	single	0.9042	0.9806	0.9353	0.8189	0.7574	0.8456	0.8737
	average	0.9309	0.9845	0.9368	0.8279	0.8739	0.8641	0.9030
	weighted	0.9285	0.9828	0.9360	0.8463	0.8707	0.8561	0.9034
SM$(0,\frac{1}{2},\frac{1}{2},\mathrm{ndc},\mathrm{n}\gamma)$								
	single	0.9513	0.9589	0.9726	0.8147	0.8543	0.9248	0.9128
	average	0.9650	0.9583	0.9738	0.9030	0.9125	0.9372	0.9416
	weighted	0.9626	0.9535	0.9734	0.9029	0.9047	0.9355	0.9388
SM$(0,\frac{3}{4},\frac{1}{4},\mathrm{ndc},\mathrm{n}\gamma)$								
	single	0.9116	0.9784	0.9372	0.8177	0.7518	0.8610	0.8763
	average	0.9343	0.9836	0.9387	0.8509	0.8855	0.8762	0.9115
	weighted	0.9361	0.9817	0.9378	0.8365	0.8828	0.8746	0.9083

Mumford, D. (1991). Mathematical theories of shape: Do they model perception?, *Geometric Methods in Computer Vision, San Diego, USA*, pp. 2–10.

Palacio-Niño, J. and Berzal, F. (2019). Evaluation metrics for unsupervised learning algorithms, *CoRR* abs/1905.05667.

Piccoli, C., Aparajeya, P., Papadopoulos, G.T., Bintliff, J., Leymarie, F., Bes, P., Van der Enden, M., P.J. and Daras, P. (2015). Towards the automatic classification of pottery sherds: Two complementary approaches, *in* A. Traviglia (Ed.), *Across Space and Time*, Amsterdam University Press, Amsterdam, pp. 463–474.

Pizarro, D. and Bartoli, A. (2011). Global optimization for optimal generalized procrustes analysis, *CVPR'11: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA*, pp. 2409–2415.

Sablatnig, R., Menard, C. and Kropatsch, W. (1998). Classification of archaeological fragments using a description language, *European Association for Signal Processing (EUSIPCO), Rhodes, Greece*, Vol. 2, pp. 1097–1100.

Sangalli, L.M., Secchi, P., Vantini, S. and Vitelli, V. (2010). Classification of functional data: Unsupervised curve clustering when curves are misaligned, *2010 JSM Proceedings, Vancouver, Canada*, pp. 4034–4047.

Sangalli, L.M., Secchi, P., Vantini, S. and Vitelli, V. (2012). Joint

Table 4. Cophenetic correlation coefficient results for the augmented data sets.

Method		Set 1a	Set 2a	Set 3a	Set 4a	Set 5a	Set 6a	Arithmetic mean
SM$(1,0,0,1,1)$								
	single	0.6780	0.9788	0.7984	0.5813	0.6211	0.1760	0.6389
	average	0.8777	0.9836	0.8968	0.7844	0.8088	0.7933	0.8574
	weighted	0.8530	0.9808	0.8583	0.8042	0.7633	0.7536	0.8355
SM$(0,1,0,1,1)$								
	single	0.8520	0.9736	0.9454	0.7676	0.8182	0.6921	0.8415
	average	0.8972	0.9771	0.9584	0.7852	0.8599	0.7548	0.8721
	weighted	0.9104	0.9617	0.9750	0.6916	0.8505	0.7464	0.8559
SM$(0,0,1,1,1)$								
	single	0.9592	0.8019	0.7192	0.6006	0.8912	0.8213	0.7989
	average	**0.9731**	0.8770	0.9613	0.8664	0.9131	**0.8602**	**0.9085**
	weighted	0.9687	0.8517	0.9566	**0.8742**	0.7725	0.7979	0.8703
SM$(\frac{1}{2},0,\frac{1}{2},1,1)$								
	single	0.9439	0.9619	**0.9874**	0.7321	0.8515	0.6190	0.8493
	average	0.9573	0.9713	0.9091	0.8688	**0.9009**	0.7758	0.8972
	weighted	0.9485	0.9513	0.8162	0.6914	0.8706	0.7385	0.8361
SM$(\frac{3}{4},0,\frac{1}{4},1,1)$								
	single	0.8457	0.9803	0.7350	0.7211	0.8139	0.3865	0.7471
	average	0.8850	**0.9825**	0.8749	0.8276	0.8581	0.7608	0.8648
	weighted	0.7902	0.9804	0.8155	0.8077	0.8472	0.7680	0.8348
SM$(0,\frac{1}{2},\frac{1}{2},\mathrm{ndc},1)$								
	single	0.9428	0.9539	0.9046	0.7411	0.7916	0.7523	0.8477
	average	0.9537	0.9550	0.9737	0.7923	0.8566	0.8314	0.8938
	weighted	0.9571	0.9351	0.9685	0.7862	0.8402	0.7957	0.8805
SM$(0,\frac{3}{4},\frac{1}{4},\mathrm{ndc},1)$								
	single	0.9082	0.9701	0.9562	0.7725	0.7410	0.7258	0.8456
	average	0.9396	0.9720	0.9735	0.8261	0.8388	0.7858	0.8893
	weighted	0.9386	0.9660	0.9720	0.7718	0.8590	0.7646	0.8787
SM$(0,\frac{1}{2},\frac{1}{2},\mathrm{ndc},\mathrm{n}\gamma)$								
	single	0.9474	0.9459	0.9008	0.7191	0.8154	0.7832	0.8520
	average	0.9574	0.9556	0.9739	0.8208	0.8867	0.8519	0.9077
	weighted	0.9536	0.9338	0.9651	0.7628	0.8519	0.7893	0.8761
SM$(0,\frac{3}{4},\frac{1}{4},\mathrm{ndc},\mathrm{n}\gamma)$								
	single	0.9104	0.9686	0.9552	0.7625	0.7558	0.7232	0.8460
	average	0.9431	0.9711	0.9737	0.7773	0.8550	0.8012	0.8869
	weighted	0.9279	0.9631	0.9721	0.7560	0.8670	0.7957	0.8803

clustering and alignment of functional data: An application to vascular geometries, *in.* A. Di Ciaccio *et al.* (Eds.), *Advanced Statistical Methods for the Analysis of Large Data-Sets*, Springer, Berlin/Heidelberg, pp. 34–43.

Sharon, E. and Mumford, D. (2006). 2D-shape analysis using conformal mapping, *International Journal of Computer Vision* **70**(1): 55–75.

Siminski, K. (2021). An outlier-robust neuro-fuzzy system for classification and regression, *International Journal of Applied Mathematics and Computer Science* **31**(2): 303–319, DOI:10.34768/amcs-2021-0021

Sokal, R.R. and Rohlf, J.F. (1962). The comparison of dendrograms by objective methods, *Taxon* **11**(2): 33–40.

Vogogias, A., Kennedy, J., Archambault, D., Smith, V.A. and Currant, H. (2016). MLCut: Exploring multi-level cuts in dendrograms for biological data, *in* C. Turkay and T.R. Wan (Eds), *Computer Graphics and Visual Computing*, Eurographics Association, Geneve.

Wierzchoń, S.T. and Kłopotek, M.A. (2015). *Algorithms of Cluster Analysis*, Information Technologies: Research and Their Interdisciplinary Applications 3, Polish Academy of Sciences, Warsaw.

Wilczek, J., Monna, F., Navarro, N. and Chateau-Smith, C. (2021). A computer tool to identify best matches for pottery fragments, *Journal of Archaeological Science: Reports* **37**: 102891.

Zhou, F. and De la Torre, F. (2012). Generalized time warping for multi-modal alignment of human motion, *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA*, pp. 1282–1289.

Agnieszka Kaliszewska holds an MA from the University of Warsaw, Institute of Archaeology, and Universita degli Studii di Catania, and has studied graphics at the Warsaw School of Information Technology. She is currently employed at the Systems Research Institute, Polish Academy of Sciences. Her research interests include the application of mathematics in the field of archaeology, the theory of classification in humanities and its application to computerized methods, and the use of 3D data in the preservation of cultural heritage.

Monika Syga earned her PhD degree in mathematics from the Warsaw University of Technology in 2016. Since then, she has been working as an assistant professor at the Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland. Her research interests focus on generalized convexity theory and its application in optimization.

Received: 29 June 2021
Revised: 18 October 2021
Accepted: 28 November 2021