This item is the archived peer-reviewed author-version of:

Product sizing with 3D anthropometry and k-medoids clustering

Reference:

Lacko Daniël, Huysmans Toon, Vleugels Jochen, De Bruyne Guido, Van Hulle Marc M., Sijbers Jan, Verwulgen Stijn.- Product sizing with 3D anthropometry and k-medoids clustering

Computer aided design - ISSN 0010-4485 - 91(2017), p. 60-74

Full text (Publisher's DOI): <https://doi.org/10.1016/J.CAD.2017.06.004>

To cite this reference: <https://hdl.handle.net/10067/1451110151162165141>

Product sizing with 3D anthropometry and k-medoids clustering

Daniël Lacko^{a,b,c,*}, Toon Huysmans^b, Jochen Vleugels^a, Guido De Bruyne^a,
Marc M. Van Hulle^c, Jan Sijbers^b, Stijn Verwulgen^a

Ambtmanstraat 1, 2000 Antwerpen, Belgium

^aDepartment of Product Development, Faculty of Design Sciences, University of Antwerp, Antwerp, Belgium

^biMinds-Vision Lab, University of Antwerp, Antwerp, Belgium

^cLaboratorium voor Neuro- en Psychofysiologie, KU Leuven, Leuven, Belgium

Abstract

Aside from anthropometric data tables, 3D shape models of the human body are becoming increasingly common and call for new product sizing methods based on 3D anthropometry. Though some shape model-based methods exist, most of them focus on mathematical clustering and do not discuss the usability of the clustering results for product design. In this paper, a new shape-model based clustering method for product sizing is presented that takes into account both shape information and usability for designers. The new method, called constrained k-medoids clustering, is applied on a shape model of 100 human heads. It is compared to a partitioning around medoid (PAM) clustering of anthropometric measurements of the same 100 heads (i.e., feature-based), as well as to PAM clustering of the shape model (i.e., shape based). Results show that both shape-based and constrained clustering perform better than feature-based clustering, with an average size-weighted variance (SWV) of $62 \times 10^3 \pm 16 \times 10^3$ and $66 \times 10^3 \pm 26 \times 10^3$ as compared to $72 \times 10^3 \pm 12 \times 10^3$, respectively. The average point-to-point distances in shape-based and constrained k-medoids were found to be similar to those of feature-based k-medoids, indicating that using 3D-anthropometry for product sizing will not have a negative impact on designer workload and/or a higher cost to implement more sizes. The results suggest that for head-based products, which require accurate shape and size fit, sizing systems should be created using either shape-based or constrained k-medoids, with the latter being slightly less accurate but more intuitive for further design and verification.

Keywords: 3D anthropometry, statistical shape model, clustering, product sizing, human head

*Corresponding author: Phone: +3232652472 Fax: +3232652245
Email address: daniel.lacko@uantwerpen.be (Daniël Lacko)

1. Introduction

Ideally, products such as clothing or headgear should fit the shape of the human body, because in many cases badly fitting products will cause discomfort or even pain [1, 2]. The best product fit can be achieved by having clothing or apparel tailored to the individual’s shape. However, custom-tailored products are expensive. On the other hand, while mass-production might be cheaper, it makes it less straight-forward to create well-fitting products. Instead of having precise measurements of the customer, clothing and wearables companies have to infer measurement data from small samples of their intended markets. In addition, globalization has resulted in the need to fit wearables to increasingly large and ethnically diverse populations. For these reasons, a number of product sizes is developed to fit several subgroups of the targeted customer population. The required number of product sizes and dimensions is determined by a sizing system based on anthropometric body measurements on a population sample. This section gives an overview of a number of methods for creating sizing systems based on traditional anthropometry and 3D anthropometry.

1.1. Product sizing based on anthropometric features

Traditionally, sizing systems are created on the basis of one or two measurements that the designers deem most relevant. The product dimensions for each product size are based on the percentage of the population that has predetermined measurement values, e.g. for a product with three sizes the 5th, 50th and 95th percentile (respectively P5, P50, P95) are commonly used [1, 3, 4]. While this method has been the industry standard for quite some time, it does not take into account the correlation between various body measurements.

To overcome this problem, multivariate statistics are used to determine the relevance of and correlation between anthropometric measurements. In multivariate sizing, principal component analysis (PCA) is performed on an anthropometric data table. PCA is a method that finds the structure of the variation in the anthropometric data by defining a new coordinate system with its axis in the directions of data variance [5]. It can be used to discover which measurements contribute most to the shape variation [6] (for a detailed description on how this is done, please see section 2.2).

Once the most important measurements have been determined, their percentile values are used to create physical or digital manikins. The required number of product sizes depends on the total percentage of the population covered by all sizes, also called the accommodation rate. Once the required number of sizes is known, products are modeled according to the manikins for each size and a design equation is determined in order to assign anthropometric measurements to different product sizes [3, 4]. Even with multivariate sizing, however, this often results in an incorrect fit [6, 7].

1.2. Product sizing based on 3D anthropometry

Over the past decade, there has been a growing set of 3D shape data collected either via medical imaging or by 3D scanners of various kinds. Because 3D data contains a complete representation of the human shape, it is a more versatile source of information than traditional anthropometric data tables, provided that the collected 3D shapes are represented in the same coordinate system and consist of corresponding points. A collection of 3D shape surfaces with corresponding orientation and points is called a shape model [8]. Apart from the fact that anthropometric measurements can still be performed on shape models digitally, a statistical analysis of shape models results in a complete and accurate representation of the local and global shape variation [7, 9].

Various research groups have suggested methods to incorporate 3D anthropometry in product sizing. For example, Wuhner et al. performed anthropometric measurements on a shape model of the human head in order to support the design of glasses or helmets [10]. They achieved this by firstly representing each head surface as a single point in a multi-dimensional space, and secondly each such head surface by two or three relevant feature dimensions. They then found the best possible way to cover most individuals in this feature space by a predetermined number of "boxes", the area or volume of which is defined by predetermined tolerances for each dimension. Manikins were created by taking the mathematical average of all head surfaces within each box. The idea was based on an optimization method for the production of electronic VLSI-chips [11, 12] and resulted in manikins with a realistic appearance, though no verification was provided that this results in better fitting product sizes.

Principal component analysis (PCA) is a well-known technique with many applications, to transform a dataset in a highly dimensional vector space into mutually orthogonal modes that describe the linear subspace generated by the dataset, see for example [13]. Luximon et al. created a shape model of the human foot from 3D scans of 70 males [6]. They performed PCA on the shape model to find the structure of variation in shape space instead of in feature (measurements) space, as in traditional multivariate sizing. They found that PC-prediction has the lowest prediction errors, but that prediction using correlated measurements still outperforms traditional measurements. This observation is similar to [14], which can be regarded as an extension of the correlation method, applied to the human head.

k-means clustering is an algorithm that divides a set of data points in a normed vector space into a partition of k clusters, based on the distance between each surface and the group means [15]. k-medoids clustering is a variant of k-means clustering where each mean is replaced by its, that is, a point in the initial dataset closest to its means. Niu et al. present a method in which the corresponding 3D surfaces are divided into uniform patches (or blocks) and are represented by vectors describing the average and standard deviation of the

inter-subject distances within each patch [16, 17]. They refer to these vectors as block-distance based vectors and create a sizing system by performing k-means clustering on block distance norms. In [18], the authors show that k-means based on principal components or on block-distances based vectors results in clusters with less internal variation than k-means performed on anthropometric features. Since the differences between PC based and block-vector based k-means clustering were small, block-distance based vectors were preferred as they yield a feasible input to product sizing.

While the methods discussed above have the potential to render mass-production design more effective by automating the creation of sizing systems, none of them utilize the full 3D shape information contained in the shape model. Instead, sizing is based either on traditional anthropometry or by simplified parameters which inevitably leave out part of the shape information, and the shape models are effectively only used at the beginning of the sizing process. Furthermore, while a number of design applications are suggested, including product sizing, it is not specified how these mathematical clustering methods should be used by product designers. As discussed in [14], statistical analysis on shape models results in a set of statistical parameters (e.g. high-dimensional vectors of principal components and principal component scores) that are not intuitive for designers, leading to a slow adoption of 3D anthropometry for product development. As long as the end-users are not considered in development and discussion of new shape modeling and clustering methods, the use of these methods in design will be limited to theoretical applications [7, 9]. Alongside the development of new clustering methods, it should be made clear how designers should interact with shape models [9].

Recently, Baek and Lee were the first to perform clustering based on 3D shapes and to consider the representation and interaction required by product developers [19]. In their paper, they focused on updating the anthropometric design process for shoe design. They proposed hierarchical clustering and average shapes per cluster to serve as manikins, along with a color map of how the manikins differed from the clusters. They proposed that shoe makers use anthropometric measurements, as well as the cluster manikins, to create new virtual shoe lasts for each group. However using only geometric averages as CAD-manikins for each cluster will exclude many individuals which shapes that are more extreme within this group [1]. Also, their proposed solution is a return to anthropometric sizing tables, where new customer would need to be measured for seven different foot dimensions. This increases complexity for the user compared to memorizing only one measurement (foot length) and using that as a reference for buying new shoes.

In this paper, k-medoids clustering is proposed to cluster a data set of 3D shapes of the human body that takes account of full shape information. The method is tested on a shape model of the human scalp [14]. Because clustering is performed on all 3D points instead of on intermediate parameters, the resulting

clusters more accurately represent morphological differences in the underlying population. Furthermore, a novel version of the k-medoids algorithm is provided, where clustering is still based on morphological differences between scalp surfaces, but is also constrained by one or more anthropometric features, in order to make product sizing more intuitive to designers and potentially also to end-users. Both methods are verified by comparing the within- and between-group variation of the resulting clusters to those of k-medoids clustering based on traditional anthropometric features, the hypothesis being that k-medoids clustering will perform better than traditional anthropometric clustering. Finally, in addition to comparing the clustering methods, a workflow for the application of the clustering method for product sizing is discussed, including a method to select the optimal number of product sizes.

2. Methods

In this section, the examined clustering algorithms and evaluation methods are described. The sample used for clustering consisted of 100 scalp surfaces and is briefly described in section 2.1. The independent variables for the experiment were the number of clusters and the clustering methods described in sections 2.2, 2.3 and 2.3 and number of clusters. Dependent variables were the metrics used to assess the obtained clusters described in section 2.5.

2.1. Data acquisition

The analysis below was performed on a data set of 100 MRI scans acquired using a Philips ACS III 1.5 T scanner in the sagittal acquisition plane, with a slice thickness of 1 mm, an echo time of 10 ms and a repetition time of 18 ms, which were used in previous work of the authors to develop a statistical shape model [14]. The shape model consists of an equal amount of males and females, between 20 and 40 years old.

2.2. Statistical analysis of shape variation

In a preprocessing step, each of the 100 head surfaces obtained from the MRI scans is converted into a point set of 10 000 points that are approximately uniformly distributed over the head and are at corresponding locations. Such a corresponding point set representation is obtained using a three-step, group-wise correspondence optimisation procedure [8, 14]. First, a one-to-one flattening of each surface to a common square planar region is calculated. This brings the surfaces into an initial correspondence. Second, the correspondence is improved by applying a non-rigid transformation to each of the flattenings, in a multi-resolution fashion using B-spline grids. The Minimum Description Length criterion, favouring compact shape models, is used as the optimisation objective and it is optimised with the LBFGS-algorithm. In the final step, the correspondence is used to calculate the average head shape and a set of 10 000

points is uniformly distributed over this average head. The corresponding locations in the common planar square domain are used to sample each of the 100 surfaces, finally resulting in 100 corresponding point sets.

The following anthropometric measurements were annotated on the shape model: head length (hL), face width (fW), bitragion width (bW), ear height (eH), horizontal position of the ear (hE), vertical position of the ear (vE), projected ear height (pE), circumference (cF), arc length (aL) and arc width (aW). These measurements will be used as features.

After processing, each head surface can be represented by a vector of 3D coordinates $\hat{\mathbf{x}}_i = [\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,n_p}] = [x_{i,1}, y_{i,1}, z_{i,1}, \dots, x_{i,n_p}, y_{i,n_p}, z_{i,n_p}]$ with n_p the number of vertices (in this case, $n_p = 10000$) and combined in a zero centered $3n_p \times n_s$ shape matrix

$$\mathbf{X} = [\hat{\mathbf{x}}_1^T - \bar{\mathbf{x}}^T, \dots, \hat{\mathbf{x}}_{n_s}^T - \bar{\mathbf{x}}^T], \quad (1)$$

with

$$\bar{\mathbf{x}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \hat{\mathbf{x}}_i. \quad (2)$$

PCA was then performed on shape matrix $\mathbf{X}$. The principal components $\mathbf{P}$ of shapes $\mathbf{X}$ capture all shape variation already contained in the individual shape vectors (3D coordinates of a large number of points on the scalp surfaces), which is much more accurate than if PCA were to be performed on an anthropometric measurement table and are also referred to as shape modes. The scalp shapes can be represented and parametrized by the shape mode scores $\mathbf{B}$, with 99 scores for each head instead of 30 000 points coordinates. However, because the shape modes are difficult to interpret, a mapping $\mathbf{M}$ can also be calculated between the shape mode scores and a $n_f \times n_s$ matrix of anthropometric features $\mathbf{F}$, with n_f the number of features (in this work $n_f = 10$) by multiplying the scores by the pseudoinverse $\mathbf{F}^+$ of the features (described in [14]), as in equation 3:

$$\mathbf{M} = \mathbf{B}\mathbf{F}^+ \quad (3)$$

This results in an intuitive parametrization that can be used to predict new head shapes accurately, as shown in equation 4, and is discussed in detail in a previous publication [14].

$$\mathbf{x}' = \bar{\mathbf{x}} + \mathbf{P}\mathbf{M}\mathbf{f}^T \quad (4)$$

When PCA is used for dimensionality reduction, the elements on the main diagonal of $\mathbf{S}^2$ are added until a desired percentage of total variance is reached, and all remaining shape modes are discarded [5]. While there is a number

of methods to determine which principal components are most relevant to the shape variation [5], a generally accepted method is to select those components that explain a cumulative percentage of the population [20, 21, 4, 6, 22]. A threshold between 70% and 90% is sufficient for most applications [23]. The percentage of variance explained by each shape mode was calculated by dividing each element of the diagonal of $\mathbf{S}^2$ [6] by the total variance, which is the sum of all diagonal elements λ_i of $\mathbf{S}^2$, with $i = 1, \dots, n_s$.

To determine which measurements are the most relevant for shape variation, the correlation of each anthropometric measurement with the shape modes was determined. This was done by finding the Pearson's correlation coefficient ρ between the features (or measurements) in $\mathbf{F}$ and the scores in $\mathbf{B}$ for those shape modes that add up to explain 70% of shape variation. The correlation between pc score B_i and feature F_j can be found as in equation 5:

$$\rho_{B_i, F_j} = \frac{\text{cov}(B_i, F_j)}{\sigma_{B_i} \sigma_{F_j}} \quad \forall i \in 1, \dots, n_m, j \in 1, \dots, n_f \quad (5)$$

2.3. Clustering method

As discussed in the introduction, clustering is the process of grouping data into clusters that share similar characteristics. The dissimilarity of two individuals is determined by the average difference between their characteristics, and is used to determine which cluster an individual belongs to. In clustering based on traditional anthropometry, the characteristics are anthropometric measurements and the dissimilarity between two individuals is determined by taking the average distance between their measurements. In shape model based clustering, the average Euclidean distance between all corresponding points of two individual head surfaces was used instead. In both cases, the dissimilarity is represented by a symmetric $n_s \times n_s$ -matrix $\mathbf{D}$ with main diagonal 0 and the difference between individual l and j at index (l, j) and (j, l) .

k-medoids clustering is a variant of k-means clustering that involves the distance to the cluster medoids instead of cluster means [24]. The medoid of a cluster is the individual in that cluster that is closest to the mathematical mean shape. k-medoids clustering can be applied in any metric space (X, d) . It basically comprises the following steps:

- step 1 randomly select initial cluster medoids $m_1, \dots, m_k \in X$;
- step 2 set the initial set of clusters $K_1 := \{m_1\}, \dots, K_k := \{m_k\}$;
- step 3 take $x \in X$, select $j_x \in \{1, \dots, k\}$ such that $d(x, m_{j_x})$ is minimal, i.e. $d(x, m_{j_x}) \leq d(x, m_i)$ for all $i \in \{1, \dots, k\}$;
- step 4 assign x to cluster K_{j_x} ;
- step 5 repeat step 3 and step 4 for all $x \in X$;

- step 6 For all $i \in \{1, \dots, k\}$, pick $m'_i \in K_i$ such that $\text{cost}(m'_i, K_i) := \sum_{x \in K_i} d(m'_i, x)$ is minimal.
- step 7 Repeat step 2-6 with $m_1, \dots, m_n$ replaced by $m'_1, \dots, m'_n$, until $m_1, \dots, m_n$ remains constant.
- step 8 Repeat steps 1-7 for n trials, each trial with a different selection of initial medoids $m_1, \dots, m_k \in X$ in step 1. Retain the cluster $K_1, \dots, K_n$ for which total cost in step 7 is minimal.

The repetition in step 7 is called a Lloyd iteration. Note that the total cost $\sum_{i=1}^k \text{cost}(m_i, K_i)$ remains constant when the process is aborted in step 7.

The most commonly used k-medoids method is partitioning around medoids, as described in [24]. It is implemented as in algorithm 1 below, similar to the implementation presented in [25].

Algorithm 1 k-medoids clustering

Input:

$\mathbf{D}$ - symmetric $n_s \times n_s$ dissimilarity matrix
 k - number of clusters

Output:

- $\mathbf{m}$ - the set of k medoid indices
 $\mathbf{l}$ - the cluster labels: a function that assign each shape $i \in \{1, \dots, n_s\}$ to its cluster $l_i \in \mathbf{m}$
 c - total clustering cost
- 1: $\mathbf{m} \leftarrow$ random k -combinations of $\{1, \dots, n_s\}$
 - 2: $M \leftarrow \emptyset$ $\triangleright M$ tracks medoid set over Lloyd iterations
 - 3: **while** $\mathbf{m} \notin M$ **do**
 - 4: $M \leftarrow M \cup \{\mathbf{m}\}$
 - 5: $\forall i \in \{1, \dots, n_s\} : l_i \leftarrow \arg \min_{m \in \mathbf{m}} d_{i,m}$
 - 6: $\forall i \in \{1, \dots, k\} : \mathbf{m}_i \leftarrow \arg \min_{j, l_j = m_i} \sum_{i | l_i = m_i} d_{j,i}$
 - 7: $c \leftarrow \sum_{i=1}^{n_s} d_{i, l_i}$
 - 8: **end while**
-

The only required input for k-medoids clustering on the present dataset is the desired number of clusters k and a $n_s \times n_s$ dissimilarity matrix $\mathbf{D}$. A random set of k indices is initially selected from $\{1, \dots, n_s\}$ to serve as initial cluster medoids. The remaining $n_s - k$ surfaces are assigned to a cluster based on which medoid they are the most similar with, i.e. the minimum dissimilarity according to $\mathbf{D}$. The cost of this clustering is determined by adding the sums of the dissimilarities of each cluster medoid to the other surfaces within that cluster. Then in each cluster, a new medoid is selected that minimizes cost within that cluster.

It can be shown that this cluster medoid reassignment converges, and that in each Lloyd iteration, cost decreases until it remains constant. The returned clustering and cost is function of the initial medoid selection. however, the number of possible initial medoid selections is finite ($\frac{n_s!}{k!(n_s-k)!}$), so one can find

an optimal clustering by repeated execution, each time with a different initial medoid selection. If the cost of a clustering iteration is lower than the previous lowest cost, that iteration's medoid indices and labels are saved. This method thus returns the final medoid indices, the cluster labels for each surface and the total cost of clustering around these medoids.

Since $\mathbf{D}$ can be precomputed for any dataset, the same algorithm can be used for clustering by difference in anthropometric features and clustering by geometric differences. Each of the algorithms is performed 100 times for 2 to 10 clusters, and the medoid indices, cluster labels and cost of the clustering with the lowest possible cost is returned.

In this work, the k-medoids algorithm was slightly modified by constraining the clustering in the most important feature dimensions (as determined by the PC-feature correlation in section 2.2). To that end, we introduce an operator that assign to each subset K in the feature space $\mathbf{R}^{n_f}$ its axis aligned bounding box:

$$\text{box}(K) := \left\{ \mathbf{x} \in \mathbf{R}^{n_f} : \left(\min_{\mathbf{y} \in K} y_1 \right) \leq x_1 \leq \left(\max_{\mathbf{y} \in K} y_1 \right), \dots, \right. \\ \left. \left(\min_{\mathbf{y} \in K} y_{n_f} \right) \leq x_{n_f} \leq \left(\max_{\mathbf{y} \in K} y_{n_f} \right) \right\}.$$

The constrained clustering process can be applied in any metric space (X, d) that is endowed with a feature map $X \rightarrow \mathbf{R}^{n_f}$. It basically comprises the following steps:

- step 1 randomly select initial cluster medoids $m_1, \dots, m_k \in X$;
- step 2 set the initial set of clusters $K_1 := \{m_1\}, \dots, K_k := \{m_k\}$;
- step 3 Select $x \in X$ and $j \in \{1, \dots, k\}$ such that $d(x, m_j)$ is minimal, i.e. $d(x, m_j) \leq d(y, m_i)$ for all $y \in X$ and for all $i \in \{1, \dots, k\}$;
- step 4 If the intersection of $\text{box}(K_j \cup \{x\})$ with $\text{box}(K_i)$ is void for any $i \neq j$, then add x to K_j .
 Else, consider the set L of those indices l for which the intersection of $\text{box}(K_l, \{x\})$ with $\text{box}(K_i)$ is void for any $i \neq l \in \{1, \dots, k\}$.
 Then, if L is non-void, select $l' \in L$ such that $d(x, m_{l'})$ is minimal, that is $d(x, m_{l'}) \leq d(x, m_l)$ for all $l \in L$ and assign x to cluster $K_{l'}$.
 Else, if the set L is void, label x as unclusterable. Set cluster cost to infinity.
- step 5 repeat step 3 and step 4 for all $x \in X$ not yet assigned to a cluster;
- step 6 For all $i \in \{1, \dots, k\}$, pick $m'_i \in K_i$ such that $\text{cost}(m'_i, K_i) := \sum_{x \in K_i} d(m'_i, x)$ is minimal.
- step 7 Repeat step 2-6 with $m_1, \dots, m_n$ replaced by $m'_1, \dots, m'_n$, until $m_1, \dots, m_n$ remains constant.
- step 8 Repeat steps 1-7 for n trials, each trial with a different selection of initial medoids $m_1, \dots, m_k \in X$ in step 1. Retain the cluster $K_1, \dots, K_n$ for which total cost in step 7 is minimal.

This modification makes the clustering method more applicable to product design by linking it to commonly used anthropometric measurements, via a feature map that assign anthropometric measurements to a shape. It is described in algorithm 2.

Algorithm 2 constrained k-medoids clustering

Input:
 $\mathbf{f}_i$ - n_f vector with features of surface i
 $\mathbf{D}$ - symmetric $n_s \times n_s$ dissimilarity matrix
 k - number of clusters
 n - number of trials

Output:
 $\mathbf{m}$ - the set of k medoid indices
 $\mathbf{l}$ - the cluster labels
 c - total clustering cost
 U - set of indices of surfaces that could not be clustered

```

1:  $M \leftarrow \emptyset$  ▷  $M$  tracks medoid set over trials
2:  $\mathbf{m} \leftarrow$  random  $k$ -combinations of  $\{1, \dots, n_s\}$ 
3: while  $\mathbf{m} \notin M$  do
4:    $M \leftarrow M \cup \{\mathbf{m}\}$ 
5:    $\forall i : l_i \leftarrow \infty$ 
6:    $U \leftarrow \emptyset$ 
7:    $V \leftarrow \{1, \dots, n_s\}$  ▷ set of surfaces to cluster
8:   while  $V \neq \emptyset$  do
9:      $i, m \leftarrow \arg \min_{i \in V; j \in \mathbf{m}} d_{i,j}$  ▷ Surface and medoid closest together
10:     $V \leftarrow V / \{i\}$ 
11:     $\forall i \in \mathbf{m} :$ 
12:      if  $\exists j \neq i \in \mathbf{m} : \text{box}(\{\mathbf{f}_j | l_j = i\} \cup \{\mathbf{f}_i\}) \cap \text{box}(\{\mathbf{f}_j | l_j = j\}) \neq \emptyset$  then
13:         $D_{i,i} \leftarrow \infty$  ▷ track if surface  $i$  is rejected by medoid  $i$ 
14:      end if
15:       $D \leftarrow \min_{i \in \mathbf{m}} D_{i,i}$ 
16:      if  $D = \infty$  then
17:         $U \leftarrow U \cup \{i\}$  ▷ surface  $i$  cannot be clustered
18:         $l_i \leftarrow \infty$ 
19:         $d_{i,l_i} \leftarrow \infty$ 
20:      else if then
21:         $l_i \leftarrow \arg \min_{i \in \mathbf{m}} d_{i,i}$ 
22:      end if
23:    end while
24:     $c \leftarrow \sum_{i=1}^{n_s} d_{i,l_i}$  ▷ note cost is infinite if some surfaces cannot be clustered
25: end while

```

The main difference with algorithm 1 is that in this case, a matrix $\mathbf{F}$ contain-

ing features to constrain the clustering is also required as input. $\mathbf{F}$ in this case has a number of features in its rows (e.g. head length and circumference) and the scalp surfaces corresponding to the columns of $\mathbf{X}$, $\mathbf{D}$ and $\mathbf{B}$ in its columns. The clustering is again based on the dissimilarity between the medoids and the other surfaces. However, in this case the minimum and maximum feature boundaries of each cluster are also determined dynamically by considering boxes as each cluster grows. Before assigning any surface to the cluster with the closest nearest medoid, a check is made to see if assigning the surface to this cluster would result the cluster to overlap in all feature dimensions with any other cluster. If all features overlap with any of the other clusters, the surface is not assigned to the cluster but is instead grouped with the next closest medoid - again provided that this does not cause all features to overlap. As long as there is at least one feature dimension for which the clusters can be clearly separated, the overlap test passes and the surface can safely be assigned to the current cluster. Just as before, the algorithm is performed 100 times for 2 to 10 clusters, and the medoid indices, cluster labels and cost of the clustering with the lowest possible cost is returned.

We have opted to set cluster cost at infinity as soon as there exists one data point that can not be assigned to a cluster. However, this was never noted during the numerous execution of the algorithm. Despite of that, an indication that it can occur that a point can not be assigned to any cluster, is provided by figure 1 below.

Figure 1: An abstract visual example in which any point in the middle of $K1$, $K2$, $K3$, and $K4$ can not be assigned to any cluster.

The fact that such configuration were not detected, can be due to the fact that we used constraints that are highly correlated to the shape, such that constrained clustering and shape based k -medoids clustering yields partitions that are highly similar.

By execution, we have also noted that a re-assignment of cluster medoids that minimizes within cluster cost, increases total cost in a subsequent iteration. This

indicates that algorithm 1 converges in our setting. The formal convergence is however not investigated in this setting.

Note that adding constraints to the clustering will not be as optimal as unconstrained clustering, and will result in clusters (product sizes) with higher within-cluster shape variability. Depending on the number and nature of the constraining features, this will have an effect on the runtime on the clustering algorithm and -as there is more shape variation to take into account- will make it more difficult for the designer to create a product size that fits each cluster. However, if the clusters can be distinguished by a distinct set of anthropometric features, this enables the creation of a sizing table to assign any individual to the correct cluster (i.e. product size) without performing a 3D scan or trial-and-error fitting of several sizes. Therefore, the benefits to the user of the products might be higher than the inconvenience of more geometrically heterogeneous clusters. For further discussion on this, see section 4.6.

2.4. Evaluating clustering quality using cluster validity indices

A number of metrics exist to evaluate the quality of k-means or k-medoids clustering. These metrics, referred to as cluster validity indices, are generally defined in terms of intra- and inter-cluster variance. For example, the Ray-Turi index is the ratio of the squared distances between all points in a cluster and their barycenter (usually the mean) to the minimum of the squared distances between all barycenters [26]. As such, it is a relative measure of dissimilarity.

For the application to 3D clustering methods, the medoids were used as barycenters and the squared Euclidean distances between the 3D-coordinates of each individual shape were used as a distance metric (see section 2.3). Niu et al., however, observed that the Ray-Turi index has a sensitivity problem and propose a modified version where the numerator is the average instead of the minimum inter-cluster distance [17]. In this work, their suggestion is adopted and the modified Ray-Turi index, which will be referred to as *CVI* (cluster validity index), is calculated as in equation 6, with k the number of clusters, n_s^c the number of surfaces in cluster c and m^c the medoid for cluster c :

$$\begin{aligned} intra &= \frac{1}{N} \sum_{c=1}^k \sum_{i=1}^{n_s^c} \|\dot{\mathbf{x}}_i^c - \mathbf{m}^c\|^2 \\ inter &= \frac{2}{k(k-1)} \sum_{c=1}^{k-1} \sum_{c'=c+1}^k \|\mathbf{m}^c - \mathbf{m}^{c'}\|^2 \\ CVI &= \frac{intra}{inter} \end{aligned} \tag{6}$$

In addition, the size-weighted variance (*SWV*) was calculated as a measure of absolute dissimilarity with equation 7:

$$SWV = \frac{1}{n_s} \left(\sum_{c=1}^k n_s^c \times \frac{1}{n_p} \sum_{i=1}^{n_p} \text{Var}(\mathbf{x}_i^c) \right) \quad (7)$$

The reason for selecting the Ray-Turi index and SWV as metrics is twofold. Firstly, while most cluster metrics have been used primarily for image processing purposes, these metrics were used in previous research regarding clustering for product design [17]. Secondly, a number of relative and absolute clustering metrics from [27] were calculated alongside the CVI and SWV and were found to give similar results for the clustering algorithms in this work. Because the minor differences in performance or precision were not relevant for distinguishing between clustering methods and because a full comparison between metrics was considered out of scope, other metrics were not included in the present study.

2.5. Product-related clustering metrics

When using mathematical clustering for product development, each cluster represents a group in the population that required a separate product size. While cluster validity indices provide a good overall measure of clustering quality, they are not directly usable in product design or sizing systems, since they represent geometric shape variation in a manner that is not clearly applicable to subject-product distances. Therefore, this section introduces a number of product-related metrics. Each of these metrics has some general rules of thumb, but their specific relevance and thresholds will vary on a product-to-product basis. In order to provide a tangible example, a prototype EEG (electroencephalography) headset that is currently under development at the University of Antwerp will be used, shown in figure 2. EEG is the measurement of electrical activity in the brain, either by electrodes implanted in the brain or electrodes on the scalp. One major challenge in designing EEG headsets such as the Epoc+ by Emotiv [28], is getting the electrodes in the correct position and maintaining contact with the skin throughout measurement. For the UAntwerp prototype, the goal is to create a headset in a number of product sizes without any adjustable or movable parts, thereby eliminating positioning errors by the user as much as possible. While the device is still under development at the time of writing, the following product-related metrics have been considered.

- **Number of clusters.** The number of product sizes is generally a trade-off between cost and complexity: the fewer sizes are used, the faster and cheaper the product can be produced. However, with each product size accommodating a larger set of shapes, the within-cluster variability becomes larger, which makes it more difficult to design a fitting product.

For example, the first version of the EEG headset was printed in three sizes due to budget and time constraints. Initial tests revealed that while the positioning and stability of the device were good and the overall fit was acceptable, two frontal electrodes didn't make contact with the head

Figure 2: One product size for a prototype EEG headset currently under development at UAntwerp.

of a number of head subjects in all three clusters. This indicates that there is one or more head geometry type (e.g. elongated heads) which are now spread over the other clusters. In this case, increasing the number of clusters would result in an extra product size specifically suited for people with this head type.

- **Population percentage covered by each cluster and by all clusters combined.** The accommodation rate or population covered by a set of product sizes should be as large as possible. Furthermore, each cluster should ideally contain equivalent portions of the population, because it is economically undesirable to design for outliers and to produce a specific product size for an extremely small portion of the population. In practice, designers will often opt to design for a specific part of the population, such as those people between the 5th and the 95th percentile. However, the feature-based, shape-based and constrained clustering algorithms described above assign each surface in the dataset to a cluster and will thus result in a total accommodation rate of 100%, theoretically providing a solution for all members of a population. When comparing these kinds of clustering methods, only the proportion of the population represented by each cluster should be compared, and methods that provide few clusters containing outliers will be favoured. (Of course, these methods can always

be adapted so as to only accommodate a certain population percentage.)

To use the example of the EEG headset again, it might be that increasing the number of clusters reveals that the extra head type discussed above only represents a very small percentage of the populations, e.g. the accommodation rates for the four clusters are 50%, 20%, 25%, and %5 (the latter being the new head type). Instead of developing and producing a separate cluster, the designers could choose to find another solution, such as using extra padding or a more flexible material. Alternatively, they could decide not to take this part of the population into account at all.

- **Point-to-point distances for each cluster.** The point-to-point distance is defined as the Euclidean distance between corresponding points of two surfaces in the dataset. To determine and visualize point-to-point distances within clusters, the minimum and maximum cluster envelopes are created by calculating the standard deviation of each point in a cluster in the normal direction, then moving the medoid point in the normal direction by three times that standard deviation. The normal direction at each vertex point in this case is the normalized average of the surface normals of the three faces that contain the vertex [29]. This results in (aesthetically unrealistic) head shapes that show the minimum and maximum location of each point within the cluster and thus the physical boundaries for products designed for that size.

Figure 3: Detailed view of electrode holder and compressed sponge of the prototype EEG headset. Minimum and maximum surface points for the cluster are shown as blue stars.

Choosing a clustering with appropriate point-to-point distances depends heavily on the desired flexibility or adjustability of the product, as well as

on the contact points of the product with the user’s head. In completely fixed products, the desired point-to-point distances will depend on the maximum allowable errors in the product shape. On the other hand, in flexible or adjustable products the optimal point-to-point distances is determined by the degree of product flexibility/adjustability. If the allowable range of point-to-point distances is sufficiently large, a good clustering can decrease the required flexibility or even eliminate the need for flexibility altogether, although in the latter case a larger range of product sizes needs to be designed and manufactured.

In the case of the EEG headset, the most important contact points are between the electrodes and the user’s scalp. Each electrode has a removable and compressible sponge in front of it, that is soaked in a saline solution before use in order to ensure good electrical contact. A detailed view of the CAD design with the sponge in compresses state and the maximum and minimum surface points in the cluster is visualized in figure 3. The sponges are 10 mm thick and can be compressed to 1 mm. This means that for this prototype, each product size can accommodate 9 mm of variation within its respective cluster. Thus, when selecting clustering methods or determining the number of cluster, the within-cluster point-to-point variations should be no more than 9 mm.

As became obvious from the discussion above, even for a single product, the choice of clustering metrics (relevance and threshold values) requires design experience, a good understanding of the target market and familiarity with the user. Therefore, it is difficult (if not impossible) to provide a general set of guidelines. However, when the target product is known, these metrics offer are a good solution to compare clustering methods and guide product sizing.

3. Results

In this section, first, the clustering results are presented in section 3.1. Next, constrained clustering is explored in more detail, with section 3.2 showing example headforms constructed from the resulting clusters and showing within-cluster point-to-point differences when the number of clusters is varied. Section 3.3 shows which features should be selected as clustering constraints, based on their individual contribution to the scalp shape variation. Finally, section 3.4 shows the effect of varying the number of constraints on the composition and shape variation of the clusters.

3.1. Comparison of clustering methods using cluster validity indices

Table 1 shows a summery of the three different k-medoids based clustering methods. Feature-based clustering is k-medoids clustering where the average difference between the 10 anthropometric features of the scalp shapes was used as a dissimilarity metric (i.e. traditional anthropometry-based clustering). Shape-based clustering is k-medoids with the Euclidean distances between scalp shapes

as a measure of dissimilarity (as described in 2.3, and with no constraints added). Finally, constrained clustering is k-medoids constrained by a set of anthropometric features, e.g. by head circumference in table 1. The average CVI was 0.31 ± 0.04 for feature-based clustering, 0.21 ± 0.08 for shape-based clustering and 0.25 ± 0.03 for constrained clustering. The average SWV was 71642 ± 11728 , 62144 ± 16396 and 66445 ± 2651 , respectively, and the average point-to-point distances were 8.94 ± 1.91 mm, 8.31 ± 2.38 mm and 8.44 ± 0.77 mm, respectively. As expected, 100% of the population is covered by each of the clustering methods.

method	# clusters	comb. % pop.	CVI	SWV	mean avg pt-pt (mm)
feature-based	3	100 %	0.41	87630	11.06
	5	100 %	0.35	65716	8.67
	7	100 %	0.25	69176	8.14
shape-based	3	100 %	0.30	79779	10.73
	5	100 %	0.22	59106	8.47
	7	100 %	0.17	53923	7.38
constrained	3	100 %	0.31	78979	10.58
	5	100 %	0.23	63151	8.67
	7	100 %	0.23	59392	7.61

Table 1: Summary of clustering results for three different methods and three different numbers of clusters.

In figure 4, the clustering results for the three methods are visualized in 2D for $k = 3$ and using circumference and arc length as constraints. Note that in all three methods, all clusters are mutually disjoint and cover the entire space of models, meaning that an individual can never belong to two or more product sizes at the same time. However, when using sizing tables (visualized by the black boxes in figure 4), there is more ambiguity as to which cluster an individual should belong. Using constrained clustering, with features from the sizing table as constraints, mitigates this problem and clearly separates all three clusters in at least one feature dimension.

Figure 5 shows the response of the cluster validity indices for an increasing number of clusters, as well as for the three different clustering methods. While this figure confirms the results from table 1, it is also clear that feature-based clustering has a less stable progression as the number of clusters is increased. Although the CVI should not necessarily be perfectly linear at small intervals of clusters, the CVI value between two adjacent data points (numbers of clusters) should not differ too much and a higher number of clusters should usually show a lower CVI score (due to lower intra-cluster and higher inter-cluster distances). Figure 5 (a) shows that CVI fluctuates less in shape-based or constrained clustering (with the exception of the peak at 8 clusters for constrained-clustering), suggesting that the latter two methods are more robust.

Figure 4: Clustering results in 2D for the three different clustering methods.

Figure 5: Ray-Turi cluster validity index (a) and size-weighted variance (b) for the three different clustering methods using different numbers of clusters.

3.2. Point-to-point distances and minimum-maximum envelopes in k -medoids constrained clustering

Table 2 shows the average of the minimum, maximum and average point-to-point distances for different numbers of clusters in constrained clustering using the circumference.

# clusters	cluster	% pop.	avg pt-pt	std pt-pt	min pt-pt	max pt-pt	cF min.	cF max.
2	1	51	12.98	3.46	6.06	31.15	563	626
	2	49	12.81	3.41	7.06	34.77	511	561
3	1	44	12.35	3.19	5.79	31.11	566	626
	2	18	9.69	2.95	3.77	31.10	511	541
	3	38	9.71	3.09	4.24	24.41	542	565
4	1	27	9.94	3.33	3.83	23.57	566	591
	2	17	8.71	2.74	3.31	18.78	591	626
	3	18	9.69	2.95	3.77	31.10	511	541
	4	38	9.71	3.09	4.24	24.41	542	565
5	1	18	7.70	2.59	2.71	20.39	538	549
	2	12	8.37	3.06	2.48	25.64	511	536
	3	17	8.71	2.74	3.31	18.78	591	626
	4	28	9.04	2.99	3.87	24.46	551	567
	5	25	9.53	3.48	3.32	23.37	571	591
6	1	20	7.80	2.57	3.30	20.60	538	551
	2	24	8.32	3.02	2.58	24.45	552	567
	3	2	1.93	1.54	0.00	16.28	551	551
	4	12	8.37	3.06	2.48	25.64	511	536
	5	17	8.71	2.74	3.31	18.78	591	626
	6	25	9.53	3.48	3.32	23.37	571	591
7	1	17	8.71	2.74	3.31	18.78	591	626
	2	12	6.37	2.43	1.89	20.16	541	549
	3	6	7.20	3.28	0.92	22.63	511	524
	4	11	7.14	2.70	2.27	24.46	549	555
	5	18	7.52	2.59	2.58	20.58	555	567
	6	25	9.53	3.48	3.32	23.37	571	591
	7	11	6.80	2.20	2.27	17.83	528	541
8	1	17	8.71	2.74	3.31	18.78	591	626
	2	6	7.20	3.28	0.92	22.63	511	524
	3	25	9.53	3.48	3.32	23.37	571	591
	4	12	7.01	2.13	2.27	17.76	528	541
	5	11	7.14	2.70	2.27	24.46	549	555
	6	5	4.71	2.25	0.11	15.68	542	546
	7	18	7.52	2.59	2.58	20.58	555	567
	8	6	5.11	2.35	0.95	19.92	547	549
9	1	6	7.20	3.28	0.92	22.63	511	524
	2	8	6.52	2.20	2.08	18.87	606	626
	3	11	7.14	2.70	2.27	24.46	549	555
	4	11	6.80	2.20	2.27	17.83	528	541
	5	9	6.46	2.62	1.42	17.74	591	605
	6	18	7.52	2.59	2.58	20.58	555	567
	7	12	6.37	2.43	1.89	20.16	541	549
	8	1	0.00	0.00	0.00	0.00	591	591
	9	24	8.99	3.40	2.97	23.25	571	590
10	1	6	7.20	3.28	0.92	22.63	511	524
	2	1	0.00	0.00	0.00	0.00	555	555
	3	6	5.32	2.39	0.38	24.53	552	555
	4	10	7.12	2.87	1.54	17.99	591	605
	5	17	7.36	2.55	2.58	20.71	556	567
	6	2	1.93	1.54	0.00	16.28	551	551
	7	14	6.64	2.38	2.08	20.60	542	551
	8	24	8.99	3.40	2.97	23.25	571	590
	9	8	6.52	2.20	2.08	18.87	606	626
	10	12	7.01	2.13	2.27	17.76	528	541

Table 2: Population percentages, point-to-point differences in mm, and measurements for each cluster in constrained clustering (by circumference).
Note: clusters are not sorted according to size in this table.

The medoids for three clusters are shown in figure 6. The minimum and maximum envelopes for three clusters are shown in figure 7.

Figure 6: Side, front and top view of medoid surfaces for three clusters, using constrained clustering with circumference as constraint.

3.3. Selecting features for sizing and constraints

Table 3 shows the squared eigenvalues of the first ten principal components of the shape model. The first principal component already explains 71 % of shape variation, so the only features that need to be considered are those that are strongly correlated with this PC.

PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10
0.71	0.09	0.04	0.02	0.02	0.01	0.01	0.01	0.01	0.01

Table 3: Percentage of shape variation explained by each shape-space principal component.

Table 4 shows Pearson’s correlation coefficients between the anthropometric measurements and the first 10 principal components.

The most important features for variation seem to be circumference, head length and arc length, all of which have a correlation of $>0.75\%$ on the first principal component.

3.4. Varying the number of constraints in constrained clustering

The effects of varying the number of constraints for a low number of clusters can be seen in table 5 for a clustering with three clusters. There seems to be

Figure 7: Side, front and top view of minimum (red) and maximum (white, transparent) envelopes for three clusters, using constrained clustering with circumference as constraint.

no difference for within-cluster accommodation rates when more features are added, suggesting that only one distinguishing dimension is required.

If a higher number of clusters is desired, adding more constraints allows for a clustering that offers better shape-based separation. For example, in table 6, the cluster metrics for constrained clustering in 6 clusters with only circumference as constraint are compared against using both circumference and head length as constraint. SWV and clustering cost are both lower for the latter type of constrained clustering. (CVI is higher, but as will be discussed in section 4.3, it is better suited to compare different numbers of clusters than to compare clustering methods.)

However, this comes at an expense: a more complex sizing table will be required when more constraints are used, as demonstrated in figure 8.

4. Discussion

The aim of this article was to compare the application of several mathematical clustering methods to product sizing. While the methods and results sections focus mostly on the clustering algorithms, this section aims to define a clustering method that is maximally useful for designers. Chapter 4.2 discusses the results of the comparison, and chapters 4.3 to 4.6 describes how the proposed metrics and results should be incorporated into product design and development.

Feature	hL	fW	bW	eH	hE	vE	pE	cF	aL	aW
PC 1	0.87	0.57	0.83	0.40	0.68	0.75	0.42	0.97	0.80	0.74
PC 2	0.43	0.22	0.39	0.08	0.42	0.20	0.08	0.12	0.12	0.34
PC 3	0.06	0.17	0.15	0.05	0.01	0.43	0.02	0.03	0.30	0.33
PC 4	0.05	0.21	0.01	0.01	0.17	0.02	0.02	0	0.19	0.17
PC 5	0.0	0.08	0.06	0.30	0.21	0.20	0.31	0.11	0.05	0.13
PC 6	0.04	0.20	0.09	0.09	0.06	0.09	0.08	0.02	-2.53	0.13
PC 7	0	0.16	0	0.05	0.26	0.08	0.05	0.01	0.02	0.01
PC 8	0.03	0.04	0.07	0.10	0.05	0.13	0.11	0	0.01	0.06
PC 9	0.01	0.06	0.08	0.09	0.08	0.06	0.09	0.01	0.01	0
PC 10	0.01	0.01	0	0.15	0.02	0.07	0.15	0.01	0.06	0.03

Table 4: Pearson correlation coefficients between principal component scores and anthropometric measurements. (0 is no correlation, 1 indicates strong positive correlation, and -1 indicates strong negative correlation.)

constraint	# clusters	cluster	% pop.	avg pt-pt	std pt-pt	min pt-pt	max pt-pt
cF	3	1	18 %	9.69	2.95	3.77	31.10
	3	2	38 %	9.71	3.09	4.24	24.41
	3	3	44 %	12.35	3.19	5.79	31.11
cF + hL	3	1	18 %	9.69	2.95	3.77	31.10
	3	2	38 %	9.71	3.09	4.24	24.41
	3	3	44 %	12.35	3.19	5.79	31.11
cF + hL + bW	3	1	18 %	9.69	2.95	3.77	31.10
	3	2	38 %	9.71	3.09	4.24	24.41
	3	3	44 %	12.35	3.19	5.79	31.11

constraint	cF min.	cF max.	hL min.	hL max.	bW min.	bW max.
cF	511	541	n.a.	n.a.	n.a.	n.a.
	542	565	n.a.	n.a.	n.a.	n.a.
	566	626	n.a.	n.a.	n.a.	n.a.
cF + hL	511	541	169	194	n.a.	n.a.
	542	565	183	202	n.a.	n.a.
	565	626	193	222	n.a.	n.a.
cF + hL + bW	511	541	169	194	131	149
	542	565	183	202	136	153
	565	626	193	222	141	167

Table 5: Point-to-point distances in mm and measurements ranges in mm for constrained clustering in three clusters using circumference, a combination of circumference and head length, and a combination of circumference, head length, bitrignon with.

4.1. The effect of the statistical shape modeling characteristics

Several properties of the statistical shape model used for shape-based or constrained clustering might have an effect on the results. Firstly, because shape-based clustering is completely based on geometric differences between the surfaces, the sampling method and resolution for the original input scans are extremely important. In this work, the surfaces were sampled from MRI images with a resolution of $1 \times 1 \times 1$ mm, using a uniform grid based on the triangle area distortion when the original surfaces are projected into a planar rectangle, as discussed in [8]. Each surface in the dataset is represented by 10000 points and all further statistical analysis is based on these points. Increasing the resolution to 100000 points did not have a large effect on previous statistical analyses with this shape model, so the current resolution was deemed sufficient for the clustering of human scalp shapes. However, for more complex shapes

constraint	CVI	SWV	cost
cF	0.21	62173	210
cF + hL	0.28	61267	207

Table 6: Clustering metrics for constrained clustering using circumference only as a constraint, versus using head length and clustering.

Figure 8: Graphical representation in 2D feature space of constrained clustering with circumference only as a constraint (a) and with both head length and circumference (b).

with many large surface angles or high local shape variation, it would be more appropriate to use higher resolutions, and to experiment with different sampling methods.

Secondly, using more principal components might result in different clustering results for constrained clustering, since it would require a different set of anthropometric features. For example, if a threshold of 80% explained variation was chosen (see section 2.2), PC2 would have to be included in constrained clustering. From table 4, the most important feature for PC2 is head length. Therefore, circumference and head length would have to be used as constraints. However, from table 5, it appears that adding head length does not affect the clustering results in this case. This is due to the fact that the circumference provides a clear separation between clusters each time, as is further discussed in section 4.4. For more complex body shapes where the shape variation is spread out over more principal components, more PCs and more features might be required. Note that this has no further implications for the purely shape-based clustering, as the underlying shape model contains all principal components.

Finally, principal component analysis is a linear method, which assumes the underlying data to be Gaussian distributed. A Shapiro-Wilk test confirmed that this was the case for the anthropometric features, but not for the shape space, although PC1 -which the constrained clustering in this work was based on- was

linearly distributed with W 0.99 and p-value 0.69 (> 0.05). However, a comparison of different dimensionality reduction methods for the underlying model were considered out of scope. Following general practice ([21, 19]) and for the sake of simplicity, other dimensionality reduction methods were not considered. However, non-linear methods might provide significantly different clustering results, especially for more complex shapes.

However, the presented statistical shape model suffices to compare clustering methods for the scalp shape and to demonstrate the product sizing workflow, which is the main topic of this work.

4.2. Performance of clustering methods

From table 1 and figure 5, it appears that shape-based clustering performs better than the other two methods in terms of cluster validity indices, although the difference between shape-based and constrained clustering is small. As expected, feature-based clustering scores worse than the other two methods for all metrics. This is because there is much more information to base the clustering on: instead of a limited set of measurements the clustering can take into account the dissimilarity between tens of thousands of points, i.e. the full geometric shape of the head. Since the implementation of both methods is equally simple, if a statistical shape model is available and close product fit is important for functionality or comfort, designers should always use shape-based or constrained clustering.

A surprising result is that the average point-to-point distances between clustering methods do not differ significantly. This could be explained by the observation that the spacing between the minimum and maximum envelopes does not change a lot between clustering methods. In shape-based clustering, however, the point-to-point differences can be explained due to shape and size differences in a cluster, whereas in feature-based clustering the overall size is the most important factor (shape changes only for a small set of features).

The population percentages are also similar for the different clustering methods. The current method of evaluating the accommodation rate is therefore not suitable to distinguish between cluster methods. Even so, population percentages could be relevant for product design and should be reevaluated in future studies.

4.3. Interpretation of cluster metrics

The cluster metrics chosen to evaluate the clustering methods were the Ray-Turi index (CVI) and the size-weighted variances (SWV). As mentioned in [18], the former is a good relative measure of the shape variability in each cluster, while the latter is a good absolute measure for the overall variability resulting from the clustering method. This can clearly be seen in figure 5, where SWV

shows a gradual -almost linear- progression, whereas CVI shows more local fluctuations. As such, SWV is better suited to compare clustering methods to each other, whereas CVI is a good measure to determine the clustering quality for changes in the number of clusters.

The point-to-point differences show an expected decrease as the number of clusters increases and the clusters become smaller. Since there is hardly any difference between clustering methods, it is better not to use them for clustering comparison.

However, knowing the point-to-point differences can help designers to select the optimal number of product sizes, as discussed in section 4.5. Furthermore, dividing the point-to-point distances by the standard deviation within a cluster, or by the number of members, might be a more suitable measure of clustering quality because it would give less weight to cluster with less individuals or with higher inter-cluster variability. Another way to perform a more product-based cluster evaluation might be to only select a predefined region of interests or points of interest when calculating point-to-point differences. Alternatively, using only the dissimilarity between these points might provide a more product-specific clustering result, at the expense of losing overall accuracy. These options should be explored in future research.

4.4. Selecting features for sizing and constraints

In this paper, circumference, arc length and head length were chosen as most significant features because they have the highest loading on the first principal component, see section 3.3. However, arc length and head length are highly correlated. Furthermore, head length is more difficult to measure than arc length. Therefore, in real-life scenarios it might be more efficient to choose a different measurement. The next best feature is the vertical position of the ear, but since this is even more difficult to measure than the head length, it will not make a good sizing system. When using three constraining features, arc width would be the next best choice, with a 63% loading on PC1.

Based on the results in table 5, using multiple constraining features should be avoided when using a small number of clusters, as in this case it has no effect on the clustering, but complicates product sizing. Furthermore, the most intuitive sizing tables are those with only one feature, e.g. as for shoe sizes. However, multiple constraints are likely to have a bigger impact for higher numbers of clusters (as shown in section 3.4), or when sufficiently distinctive constraint can not be found, e.g. when multiple features have the same loading on the relevant principal components. Further research into the effect of using different numbers of clusters is required to verify this.

4.5. Selecting the optimal number of sizes and creating representative manikins

In theory, it would be possible to automatically select the optimal number of clusters. However, several cluster metric algorithms tend to get stuck at the extremes of the cluster ranges, which in most cases is not desirable for product development. For example, the Ray-Turi index will generally favor a small number of clusters, because in this case the inter-cluster distances are large while the intra-cluster distances are small, resulting in a small Ray-Turi index [26]. On the other hand, the average point-to-point distances in a cluster will continue to decrease as the number of clusters grows. Ray and Turi advised to circumvent this problem by not considering anything below four clusters. Though this might be feasible for their intended application (color clustering in 2D images), there are many products that are limited to three sizes (e.g. small, medium, large). Thus, the Ray-Turi index will not be ideal for fully automated product sizing.

In practice, a product designer will usually need to set some boundaries for the right number of sizes. A combination of the point-to-point distance plot and the CVI plot can be used for this purpose. The designer should first set the upper boundary for the number of sizes, depending on the manufacturing constraints (i.e., the maximal number of sizes that is economically feasible). Then, they set the lower limit by looking at the point-to-point distances. As discussed in section 4.3, point-to-point distances will decrease as the number of clusters increases. The designer should determine the maximum admissible point-to-point distance based on material properties and prior experience (i.e., if the material is flexible, larger within-cluster point-to-point differences can be accommodated). The maximum point-to-point distance will determine the minimum possible number of product sizes. Finally, the optimal number of product clusters is the one with the minimum CVI within this range. A visualization of this method is shown in figure 9.

Figure 9: Workflow for product sizing using clustering based on 3D anthropometry. The presented workflow can be used with shape-based clustering and -after anthropometric feature constraints have been selected- with constrained clustering.

Once the optimal number of product sizes is determined, a number of representative head forms can be created by using 3D information from the statistical shape model. For example, the minimum and maximum envelopes shown in figure 7, section 3.2, can be used as manikins: the contact points of the product should fit within these two envelopes. Alternatively, the shape model can also be used to predict realistic minimum and maximum shapes within a cluster by using the constraining feature (in this case, circumference) as a parameter (see [14]). Minimum and maximum shapes for three clusters are shown in figure 10. Another option is to use the average or medoid head forms as manikins and use markers such as colors or arrows to show the required offset at that point. For example, figure 11 shows the average head forms for three clusters with a color map indicating the offset.

Figure 10: Side, front and top view of minimum (red) and maximum (white, transparent) shape for three clusters, using constrained clustering with circumference as constraint.

Various visualization options exist within 3D anthropometry, and the choice heavily depends on the type of product and the preferences of each designer. Therefore, the best solution might be to develop a tool that gives designers the flexibility to select their own visualization options. Further research is needed in order to find out which kind of manikins are preferred, if any, as well as the

Figure 11: Side, front and top view of medoid shapes for three clusters, using constrained clustering with circumference as constraint, with color map indicating the average point-point distances.

specifications of the required software interfaces.

4.6. *Creating intuitive sizing tables*

Using constrained k-medoids clustering results in groups that have a very different shape and can be clearly separated for one or a number of measurements. From section 3, it is evident that constrained clustering performs only slightly worse than pure shape-based clustering. The biggest advantage is that once the product has been created in different sizes, new individuals can easily be assigned to their product size by taking a small number of measurements. Because the cluster mannequin represents the underlying shapes better than in traditional sizing systems, it is expected that no fitting is required, which is especially advantageous for products distributed through online shopping.

In fact, table 5 indicates that for most types of headgear, only the circumference is required: each cluster has exclusive ranges for the circumference while there is some overlap in the other measurements. The other measurements have no impact on the clustering methods. Therefore, a very simple size table can be made to assign a product size to a new individual. See figure 12 for an example.

product size	circumference (mm)
A	511 - 541
B	542 - 565
C	566 - 626

Table 7: Sizing table for three product sizes using constrained clustering based on head circumference.

Figure 12: Size manikins for sizing table 7 (cluster medoids with colors showing the point-point distances).

An advantage of using a one-dimensional size table is that this is faster and more intuitive, because people are used to finding fitting clothing items based on single measurements, e.g. hats (head circumference) or shoes (foot length). The only reason not to use constrained clustering is when the accuracy of the clustering is more important than the usability of the method. This might occur with the design of products containing biosensors, such as the EEG headset discussed in section 2.5, but this needs to be evaluated case by case.

4.7. The disadvantages of shape-based clustering

As has been demonstrated in the previous sections, clustering based on 3D shape surfaces instead of anthropometric data table has a great potential to deliver better fitting products and to optimize a range of product sizes. However, shape-based clustering also has a number of disadvantages when it comes to feature-based clustering. While 3D scanners are becoming cheaper and statistical shape modeling is becoming larger, the financial cost and the time investment in creating a 3D shape database are still significant and requires several trade offs. For example, if the entire body is 3D-scanned, the local resolution of various body parts will be low, resulting in less representative shape models. If, instead, a specific body part of interest is scanned, the 3D model will be

better, but the entire process will have to be repeated for each body part and will thus be much more expensive. Furthermore, the number of individuals that should be included in the dataset in order to properly represent the underlying population is not clear. There are models with less than 100 and models with more than 1000 individuals, but as of yet there is no industry standard to ensure representative models. Finally, in contrast to feature-based clustering (and traditional anthropometry), there are currently few publicly available 3D datasets and their suitability for product design has not yet been established.

Therefore, 3D anthropometry and shape-based clustering are not necessarily appropriate for each type of product. Currently, for most commercial products, it would still be advisable to use traditional anthropometry in combination with previous experience and/or fitting experiments with population samples. However, products which require a correct fit to multiple points of the human body and which are now mostly provided through high cost customized solutions, would benefit from using 3D anthropometry. For this type of products, shape-based clustering will provide new opportunities for mass-customization which will benefit all parties involved: users, designers and manufacturers.

4.8. Future work

While this work indicates that shape-based and constrained clustering are feasible methods for product sizing, the verification was purely theoretical. Further steps are required before the proposed workflow can be implemented in product design. Apart from repeating the described analyses using different data sets and parameters, a number of practical aspects should be resolved:

- The parameters of the underlying shape model should be varied, such as the sampling method, resolution, and non-linear dimensionality reduction methods (see section 4.1). The effect of this on the clustering should be evaluated.
- The optimal input parameters for the clustering algorithms should be determined: the number of trials, the range of clusters, and the dissimilarity metric.
- User research should be performed with designers in order to discover whether the proposed clustering method is indeed sufficiently intuitive, and whether the clustering metrics are sufficiently informative. If necessary, the methods should be compared again using a different set of metrics.
- The clustering for the prototype EEG headsets described in section 2.5 should be repeated with a larger allowable number of product sizes, and the effect of this on product fit should be evaluated.

- The clustering method described in this article should be used to create a set of prototypes for other products requiring a close fit to the human head (i.e. sports helmets). The physical fit of these prototypes should be verified on a representative sample of the intended population.
- The usability of the workflow presented in section 4.5 should be verified. For example, by giving a set of ergonomic product specifications for a relatively simple product (e.g. a hat or earphones) to two groups of experienced designers, with one group using traditional anthropometry (or feature-based clustering) and the other using 3D anthropometry (either shape-based or constrained clustering) and comparing the two methods in terms of time, effectiveness, learning curve ease of use.

5. Conclusion

In this paper, a method is presented to incorporate 3D anthropometric data in product sizing by applying k-medoids clustering on a statistical shape model of the human head. When compared to traditional sizing based on the head circumference, k-medoids clustering resulted in superior clusters based on both the Ray-Turi index and size-weighted variance. In order to make the method more easily applicable in product design, the k-medoids algorithm was modified by adding the constraint that clusters should not overlap in one or more key dimensions. This constrained clustering was found to perform almost as good as traditional feature-based clustering, with an average SWV of 66445 ± 2651 compared to 62144 ± 16396 . While shape-based k-medoids clustering was slightly better, constrained clustering supports the creation of sizing tables and is more intuitive to use.

Furthermore, a method was presented to select the optimal number of product sizes, given manufacturing constraints, within-cluster point-to-point distances and the Ray-Turi cluster validity index as a measure of sizing quality. An example of creating sizing tables and digital manikins from constrained clustering was also provided.

The results show that shape-based and constrained clustering are feasible methods for creating sizing systems, and that they are preferable to traditional feature-based sizing for head products requiring an accurate fit, e.g. EEG equipment. It is expected that constrained clustering based on 3D anthropometry will result in better fitting and more comfortable headgear.

References

- [1] J. Molenbroek, *Op maat gemaakt*, Delftse Universitaire Pers, Delft, 1994.
- [2] P. Voskamp, P. van Scheijndel, K. Peereboom (Eds.), *Handboek Ergonomie 2010*, Arbo, Kluwer, Alphen aan den Rijn, 2010.

- [3] S. Pheasant, C. M. Haslegrave, *Bodyspace - Anthropometry, Ergonomics and the Design of Work*, third edition Edition, CRC Press, Taylor & Francis Group, United States of America, 2006.
- [4] J. Chang, K. Jung, J. Hwang, Y. Kang, S. Lee, A. Freivalds, Determination of Bicycle Handle Diameters considering Hand Anthropometric Data and User Satisfaction, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54 (20) (2010) 1790–1793. doi:10.1177/154193121005402010.
- [5] H. Abdi, L. J. Williams, Principal component analysis, *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (4) (2010) 433–459. doi:10.1002/wics.101.
- [6] A. Luximon, Y. Zhang, Y. Luximon, M. Xiao, Sizing and grading for wearable products, *Computer-Aided Design* 44 (1) (2012) 77–84. doi:10.1016/j.cad.2011.07.004.
- [7] Y. Luximon, R. Ball, L. Justice, The 3d Chinese head and face modeling, *Computer-Aided Design* 44 (1) (2012) 40–47. doi:10.1016/j.cad.2011.01.011.
- [8] T. Huysmans, J. Sijbers, V. Brigitte, Automatic construction of correspondences for tubular surfaces, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32 (4) (2010) 636–651.
- [9] S.-Y. Baek, K. Lee, Parametric human body shape modeling framework for human-centered product design, *Computer-Aided Design* 44 (1) (2012) 56–67. doi:10.1016/j.cad.2010.12.006.
- [10] S. Wuhrer, C. Shu, P. Bose, Automatically creating design models from 3d anthropometry data, *Journal of Computing and Information Science in Engineering* 12 (4) (2012) 041007.
- [11] D. S. Hochbaum, W. Maass, Approximation schemes for covering and packing problems in image processing and VLSI, *Journal of the ACM (JACM)* 32 (1) (1985) 130–136.
- [12] D. S. Hochbaum, W. Maass, Fast Approximation Algorithms for a Non-convex Covering Problem, *Journal of Algorithms* 8 (1987) 305–323.
- [13] A. Makiewicz, W. Ratajczak, Principal components analysis (PCA), *Computers and Geosciences* 19 (3) (1993) 303–342.
- [14] D. Lacko, T. Huysmans, P. M. Parizel, G. De Bruyne, S. Verwulgen, M. M. Van Hulle, J. Sijbers, Evaluation of an anthropometric shape model of the human scalp, *Applied Ergonomics* 48 (2015) 70–85. doi:10.1016/j.apergo.2014.11.008.

- [15] J. Hartigan, M. A. Wong, Algorithm AS 136: A k-means clustering algorithm, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 28 (1) (1979) 100–108.
- [16] J. Niu, Z. Li, S. Xu, Block Division for 3d Head Shape Clustering, in: *Digital Human Modeling*, Springer, 2009, pp. 64–71.
- [17] J. Niu, Z. Li, G. Salvendy, Multi-resolution description of three-dimensional anthropometric data for design simplification, *Applied Ergonomics* 40 (4) (2009) 807–810. doi:10.1016/j.apergo.2008.05.005.
- [18] J. Niu, Z. Li, G. Salvendy, Multi-resolution shape description and clustering of three-dimensional head data, *Ergonomics* 52 (2) (2009) 251–269. doi:10.1080/00140130802334561.
- [19] S.-Y. Baek, K. Lee, Statistical foot-shape analysis for mass-customisation of footwear, *International Journal of Computer Aided Engineering and Technology* 8 (1-2) (2016) 80–98.
- [20] P. Hammond, T. J. Hutton, J. E. Allanson, B. Buxton, L. E. Campbell, J. Clayton-Smith, D. Donnai, A. Karmiloff-Smith, K. Metcalfe, K. C. Murphy, others, Discriminating power of localized three-dimensional facial morphology, *The American Journal of Human Genetics* 77 (6) (2005) 999–1010.
URL <http://www.sciencedirect.com/science/article/pii/S0002929707633849>
- [21] C.-H. Chu, Y.-T. Tsai, C. C. Wang, T.-H. Kwok, Exemplar-based statistical model for semantic parametric design of human body, *Computers in Industry* 61 (6) (2010) 541–549.
- [22] Developing and implementing parametric human body shape models in ergonomics software.
- [23] L. Jolliffe, *Principal Component Analysis*, 2nd Edition, Springer, 2002.
- [24] H.-S. Park, C.-H. Jun, A simple and fast algorithm for K-medoids clustering, *Expert Systems with Applications* 36 (2) (2009) 3336–3341. doi:10.1016/j.eswa.2008.01.039.
- [25] C. Bauckhage, *Numpy/scipy Recipes for Data Science: k-Medoids Clustering*, Tech. rep., Technical Report, University of Bonn (2015).
URL https://www.researchgate.net/profile/Christian_Bauckhage/publication/272351873_NumPy_SciPy_Recipes_for_Data_Science_k-Medoids_Clustering/links/54e244d10cf29666379649bc.pdf
- [26] S. Ray, R. H. Turi, Determination of number of clusters in k-means clustering and application in colour image segmentation, in: *Proceedings of the 4th international conference on advances in pattern recognition and digital techniques*, India, 1999, pp. 137–143.

- [27] B. Desgraupes, Clustering Indices, University of Paris Ouest-Lab ModalX (2013) 1–34.
URL <ftp://apache.cs.uu.nl/mirror/CRAN/web/packages/clusterCrit/vignettes/clusterCrit.pdf>
- [28] EMOTIV Epoc - 14 Channel Wireless EEG Headset.
URL <http://emotiv.com/epoc/>
- [29] Continuous saturation of curved surfaces, IEEE Transactions on Computers 20 (6).