Machine Vision and Applications (2011) 22:235–243
DOI 10.1007/s00138-009-0230-y

ORIGINAL ARTICLE

Comparison and classification of 3D objects surface point clouds on the example of feet

Rainer Grimmer · Björn Eskofier · Heiko Schlarb · Joachim Hornegger

Received: 31 August 2008 / Revised: 3 July 2009 / Accepted: 12 October 2009 / Published online: 26 November 2009

Abstract One of the main tasks of shoe manufacturing is the production of well fitting shoes for different specialized markets. The key to conduct this properly is the analysis of the factors that influence the variations of the foot shape. In this paper methods and results of clustering and analysis of 3D foot surfaces are presented. The data were collected from a study with more than 12,000 feet that have been laser-scanned. The database contains point clouds acquired from persons coming from different regions of the world. Furthermore, additional personal data were collected. Two different methods for quantifying the similarity of 3D surface point clouds are therefore developed. The first method generally works on nearly arbitrary 3D surface point clouds, while the second one is specialized on foot data sets. These similarity measures were used on the data sets of the foot-shape study, together with clustering and feature quality evaluation methods. The purpose was to obtain information about the impact of, and the relationship among, the different factors influencing the shape of a foot. Through the observations of the experiments presented here it was possible to build up a hierarchy of different levels of feature-groups determined by their impact on the foot shape. Furthermore, an investigation of the quality and amount of impact of the features, according to their ability to separate specific subgroups of persons, is shown. Based on these results it was possible to select those features, which result in the largest effect when designing shoes for e.g. the Asian versus European markets.

R. Grimmer (✉) · B. Eskofier · J. Hornegger
University of Erlangen-Nürnberg,
Martensstr. 3, 91052 Erlangen, Germany
e-mail: Rainer.Grimmer@imp.uni-erlangen.de

H. Schlarb
adidas AG,
Adi Dassler Strasse 24-26, 91443 Scheinfeld, Germany

Keywords 3D foot measurement · Classification · Laser scanner · 3D point cloud · Unsupervised learning

1 Introduction

As shoe manufacturing is becoming more sophisticated, the challenge of producing well fitting shoes is also becoming more pressing. For the correct fit of a shoe more factors are important than just the foot length and width [1,2]. To obtain more information about the differences that influence the foot shape a large study is necessary. Such large studies, however, cannot be manually analyzed by shoe-design experts.

An automated investigation of the correspondences and differences between the measured feet and the known facts about the corresponding persons is necessary. Laser scanners provide the technical basis to perform such a study. These scanners measure an object and generate a 3D point cloud, which can then be analyzed. For the automated analyses of the laser-scanned data, one needs algorithms which are able to compare objects on the basis of their 3D surface point clouds.

In our approach, we chose 3D laser measurements of foot surfaces as exemplary data for investigating the possible gain of information with methods of pattern recognition. To be able to get answers to the question if and how persons can be grouped together based on their feet, we developed methods for comparing foot shapes. Furthermore, our analysis should provide essential insight in developing shoe designs, which improve the fitting accuracy for specific subgroups of shoe markets.

To achieve this, we developed two methods for comparing foot surfaces. The first one only needs the measured point clouds of a laser-scanner and could potentially also be used for comparing other objects which are represented by a 3D surface model of a point cloud. Unlike the methods of Mochimaru et al. [3] and of Leng et al. [4], our method not only uses the surface information, but also the complete volume information contained in the point cloud. The second method is explicitly optimized for the purpose of measuring feet as it is based on the detection of 19 foot features defined by experts. These features have been weighted by experts according to their ability to express the fact that different regions of the foot have a different significance in the fitting accuracy. The use of manually defined features in classification tasks is quite common, particularly in applications associated with biometrics like face classification [5,6].

Furthermore, we investigated the finding of the different factors influencing the shape of feet and developed a rating of the impact of the different influences. Finally, we examined which foot features are suitable for differentiating different groups of people to give an idea of which measurements should be taken on a customer in order to determine the appropriate foot group.

The structure of the paper is as followed: Section 2 describes the materials and methods used in our analysis, followed by the performed experiments in Sect. 3. The results of the experiments are discussed in Sect. 4 and the paper concludes with a summary and brief discussion of potential further work.

2 Materials and methods

In this section, we first show how the acquisition of the 3D point clouds of the feet was performed. In our research we first addressed the different methods for comparing 3D surface point cloud data sets in general, and then in the specific case of feet. This gave us two suitable similarity measures. Our clustering methods are then presented in the framework of these two similarity measures.

2.1 Materials

For the acquisition of the foot surface 3D point data we used a YETI foot scanner of the Vorum Research Cooperation. With this scanner it is possible to acquire about 50,000 points of the surface of the scanned foot. The used scanner has an nominal accuracy of ± 0.1 mm. Yet in the field the accuracy is reduced due to disturbing light entering the scanner. Therefore we observed as worst case accuracy ± 1 mm. Yet by knowing this light problem for the in the field measurements as much light was excluded as possible. However due to the fact that in this investigation we are interested in the general differences of the foot shape, even the worst case accuracy is sufficient. Further on even the manufacturing accuracy of shoes are in the range of ± 1 mm.

Table 1 Additionally available information for each foot shape dataset

a.	Gender of the person
b.	Person affiliation to the county
c.	Age of the person
d.	Weight of the person
e.	Height of the person
f.	Type of job (standing or sitting)
g.	Type of sport (football, tennis, running, etc.)
h.	Frequency of sport activities
i.	Time of the acquisition

An extensive data collection of feet, left and right foot per individual, has been performed with this scanner. More than 12,000 data sets have been acquired during different data collection sessions in Europe, Asia, North America and Australia. In addition to the point cloud, the results of a corresponding questionnaire are also available. The information provided by the questionnaire can be seen in Table 1.

Besides this large study, we performed a control group study to evaluate the quality of the algorithms. In this smaller study we measured 20 persons, each one 5 times, with a small variation of the acquisition conditions, like, for example, the positioning of the foot in the scanner. The goal of this study was to evaluate how accurately the proposed methods can detect similarities. It is expected that different scans of the same person should have a high similarity. The difference of more scans of the same person should be independent from the measurement conditions for a good similarity measure and, therefore, only depend on the measurement accuracy of the scanner. Investigations of the measured data sets of these 20 persons showed that 2 persons heavily changed the distribution of their weight with respect to the foot between the scans. This resulted in change in the shape of the foot to a certain extent between the scans. Especially, the parts with soft tissue are affected by weight changes. From this we concluded, that for the large study, the data collector should oversee that the examined person distributes the weight equally between the feet during the scan.

2.2 Methods for comparing foot point clouds

Both methods for comparing foot point clouds that are presented here use the same normalization steps. These are explained first in this section. Following this, we present the first method for getting a similarity measure for feet. This is a method that works solely on the measured point clouds and is therefore not limited by the usage on foot measurements. It can be used on all measured 3D surface objects, which are convex in one coordinate direction. The second method is based on adidas studies of the different interesting regions of a foot and is weighting the computed features according to their importance to shoe fitting.

2.2.1 Normalization steps

To compare different foot shape data sets, it is necessary to normalize the position and the orientation of the foot. These normalizations have to be performed because the scanned feet can have, up to a certain amount, a different location and orientation in the scanner and we want our analysis to be unaffected by these acquisition-dependent factors. For the first method, we also perform a normalization to a normalized grid.

First of all the position of the object should be independent from the scanner geometry. Therefore, the origin of the coordinate system is set to the center of gravity of the point cloud. Equation (1) shows the calculation of the center of gravity and Eq. (2) shows how the origin of the coordinate system can be translated from the scanner coordinate system to the new system, where n is the number of the measured points $\mathbf{p}$ in the data sets and $\mathbf{p}_i = \{x_i, y_i, z_i\}$ are the coordinates of a specific point $\mathbf{p}_i$ with $i = 1, \ldots, n$.

$$\mathbf{m} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{p}_i \tag{1}$$

$$\hat{\mathbf{p}}_i := \mathbf{p}_i - \mathbf{m} \quad \text{for } i = 1, 2, \ldots, n \tag{2}$$

The second normalization step is the orientation of the objects. In our case the feet are standing on a bottom plate. Due to this only the rotation on the plane of this plate has to be investigated. We defined the following coordinate system: The z-axis is along the maximal extend of the foot. The x-axis is perpendicular to the z-axis and together they define the bottom plate. The y-axis the perpendicular to the bottom plate and therefore defines the up-vector (see Fig. 1).

For the rotation, a robust measure for a rotation axis depending on the point cloud is needed. We applied principal component analysis (PCA) in order to calculate the axis of elongation of the point cloud. In PCA, the axis of elongation is obtained by the eigenvector of the covariance matrix which corresponds to the largest eigenvalue. The other eigenvector provides the minor axis of the ellipsoid. The usage of the PCA is a very common and stable method for the rotation normalization [7] and works very well also in our case.

The third normalization step is to map the measured point clouds to a normalized grid. As mentioned before, this normalization is applied only on our technique for general point cloud comparisons. This normalized grid means a grid with defined distances, in our case 2 mm, which is overlaid in the captured volume and the measured points are mapped to the nearest grid-point.

2.2.2 Method for comparing general point clouds

The intention of the first method was to derive an object-independent technique for quantifying the differences between two objects available as point clouds as measured by a laser scanner or another device which gives 3D surface point information. Therefore, the method should only use the information of the 3D points and not require further human interaction. The similarity measure we chose is an approximation of the volume difference between the volumes of two aligned point clouds. To measure this, we first normalize the objects and then we search for the parts of the two objects with a difference in the overlap. The volumes of the nonoverlapping parts are then summed up.

After the normalization (see Sect. 2.2.1) the volume discrepancies can be calculated by performing the following steps. The 3D object is first expressed as a collection of cut-views perpendicular to the axis of elongation to reduce the problem to the 2D case. This can be done since with the normalized grid the cut-views of both objects are exactly at the same position. So the single cut-views (see Fig. 2

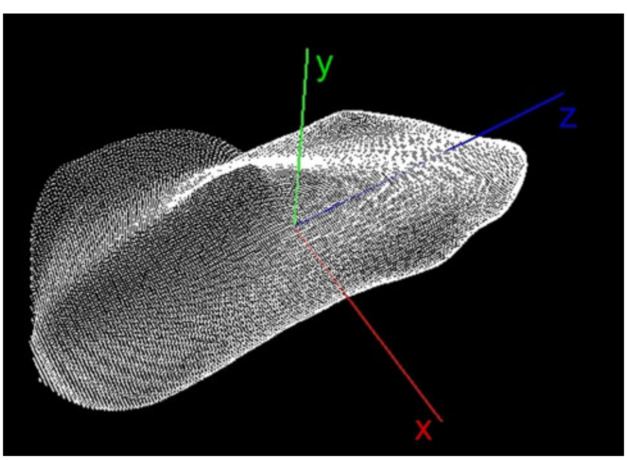

Fig. 1 Point cloud of measured foot with used coordinate system

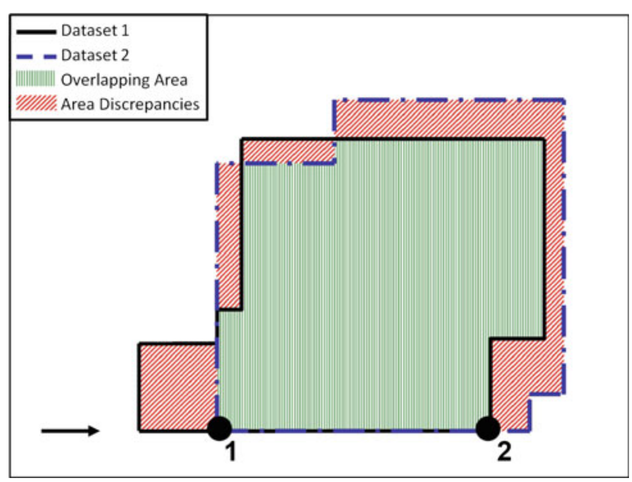

Fig. 2 Calculation of the area discrepancies on one cut view

for a stylized cut-view) of the same position are compared with each other. This has the advantage that the 3D problem can be reduced to a sum of 2D problems, which reduces the complexity and therefore the calculation time. On these cut-views, the area of each object is first calculated separately and denoted by A_{F1} and A_{F2}. This can be done by the Gaussian trapezoid formula

$$2A_{F1} = \sum_{i=1}^{n}(x_i + x_{i+1})(y_{i+1} - y_i), \tag{3}$$

where again n is the number of points in the cut-view, x denotes the x-component of a point $\mathbf{p}$ and y is its y-component. A_{F2} is computed accordingly.

The overlapping area of both slices is calculated line by line as followed: for each grid line that occurs in both cuts, we identify the set of the overlapping points and take the first and the last of these on this line. For example, in the schematic in Fig. 2, in the bottom line pointed by the arrow, the set contains the points (1) and (2). This is repeated for each overlapping line. In case of no overlapping line no point is set. Once all border points of overlapping area are identified, the overlapping area itself is calculated by the Gaussian trapezoid formula. This overlap area is denoted by A_O.

The next step is to calculate the discrepancy area V_{Disp}. This is done by subtracting the overlapping area twice from the sum of A_{F1} and A_{F2}. To get the desired volume discrepancies we then multiply the calculated area discrepancies by the z-axis increment Δz_i between two adjacent cut views. This give us the final formula

$$V_{\text{Disp}} = \sum_{i=1}^{N}(A_{F1_i} + A_{F2_i} - 2A_{O_i})\Delta z_i \tag{4}$$

where N denotes the number of cut-views.

This method has the advantage that it gives a relatively intuitive and meaningful measure of the similarity of two 3D objects. However, it has the drawback that all differences are treated equally and weighting of different regions cannot be easily incorporated.

2.2.3 Method for comparing feet on basis of foot features

The second method was designed in the way that it allows the inclusion of information about the unequal importance of different parts of a foot by weighting them. Moreover, it should be possible to calculate the measure on either the acquired point cloud of a laser-scanner or on measurements taken manually. This provides the possibility to manually match measured feet with point clouds in the database.

To accomplish this task we defined 19 foot features. These features were build up by an expert board of experienced last makers and sport scientists in the field of foot ergonomics. The features displayed in Table 2 were selected on basis of these experts experience on last making and foot physiology. These features have further been rated by these experts based on their importance with respect to the sensitivity of the fitting accuracy. The weighting of the features has been performed by distributing the features to four different weight groups. Group A is the one with the highest weight and D the one with the lowest weight. The distribution of the features to the different weight groups can also be found in Table 2.

Table 2 The features that were manually selected by experts for the feat classification. A–D represents the importance of the individual features with respect to feet shapes with A being the highest weight

Group	Number	Feature description
A	1st	Length of the foot
A	2nd	Ball width
A	3rd	Circumference of the ball
B	4th	Length to the medial ball point
B	5th	Length to the lateral ball point
B	6th	Height to medial ankle
B	7th	Height to lateral ankle
B	8th	Heel width
B	9th	Circumference of the wrist
B	10th	Circumference of the ankle
B	11th	Circumference from heel to wrist
B	12th	Width between the ankles
C	13th	Maximal toe height
C	14th	Ball height
C	15th	Wrist height
C	16th	Height of the arch
C	17th	Circumference of the inboard foot
D	18th	Angle of the heel
D	19th	Width changes of the toe

Due to the fact that the features have different value ranges there is a need for another kind of normalization besides the two first normalization steps described in Sect. 2.2.1. All features are normalized to a standard deviation of 1. For the experiments the weighting of the features is introduced in a way that the single feature values are multiplied by a weight group dependent factor. For our experiments we chose the following factors:

– A features are multiplied by 4
– B features are multiplied by 3
– C features are multiplied by 2
– D features are multiplied by 1.

The similarity measure itself interprets the features as a 19 dimensional vector $\mathbf{c}$ representing a point in a 19 dimensional space. The elements of this vector are the individual weighted feature values. The Euclidean distance between two of these vectors is used as measure of the similarity of two data sets.

2.3 Clustering and quality investigation methods

We have implemented two different types of methods that each use the proposed similarity measures to get valuable information out of the data sets. The first methods are based on unsupervised learning. There we chose two techniques: a hierarchical and an iterative clustering. The second kind of method acts as a quality criterion and calculates the ability of our similarity measures to group the data into defined subgroups. This means person-specific data are given and one is interested in how suitable different features and measures are in assigning the individual data to the appropriate group.

2.3.1 Hierarchical clustering

The idea of hierarchical clustering is that each foot-data starts in its own class. Then in every iteration the two classes are merged that fulfill an optimization criterion best. The algorithm stops when a preset amount of classes is reached.

As optimization criterion we chose the mean intraclass distance of all classes after merging two classes. In every step, the two classes that fulfill this optimization criterion best are chosen and then merged. This is a greedy algorithm because once made decisions are never reconsidered. It works well for not clearly separated classes. Furthermore it is possible to identify measurements, which are probably outliers, since they stay very long in a separate class of one element. These measurements can then be excluded from the study to get rid of acquisition errors, which would unnecessarily worsen the results.

2.3.2 Iterative clustering

The second clustering method is an iterative clustering, where the amount of classes is chosen first. Then the samples are randomly distributed over these classes. After this in every iteration step it is examined for every data set if the assignment of this data set to an other class would fulfill the optimization criterion better than the current assignment. If there is a better class the data set will be assigned to this class in the next iteration step.

For this clustering method, we choose different optimization criteria for the two similarity measures. For the features the criterion is the minimization of the mean error. Therefore, for every class k the mean feature vector

$$\boldsymbol{\mu}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbf{c}_i \tag{5}$$

is calculated, where n_k denotes the number of elements in class k.

For the volume discrepancies the calculation of the mean point cloud is very complex and not feasible for clustering in a reasonable time. In consequence we calculate the mean distance to the elements of the different classes for the inspected element. These mean intraclass distances to the inspected element are compared. The class is chosen with the resulting minimal mean intraclass distance change, if the inspected element would be introduced to this class.

2.3.3 Quality test by the nearest–neighbor classifier

The purpose of this method is to quantify the ability of a similarity measure to separate defined groups of data sets. Therefore, a nearest-neighbor (NN) classifier is used in the following way: At first, one data set is removed from the investigated sample. Second, the NN classifier is used to find the best corresponding data set in the remaining sample. If the result of this classification is from the same group of data sets as the excluded element, then the assignment is regarded as correct, if not as wrong.

This is done for every data set of the investigated sample and the ratio of correct assignments to all assignments is taken as quality criterion.

3 Experimental results

The experiments we performed can be divided into three groups. At first we made some tests to verify the methods we developed. Second, we conducted a large number of clustering experiments from which we here present the general results. Finally we present two of the experiments where we test how good predefined groups can be discriminated by different features.

3.1 Method evaluation experiments

First, we investigated if it is possible to calculate reasonable values using the volume discrepancy similarity measure. To do this, we randomly chose a sample of 1,200 elements consisting of left and right feet in about the same quantity. If the measurement is correct one would expect large differences between left and right feet in the results. To visualize the calculated similarity we calculated a 2D Sammon map [8]. The Sammon transformation maps the high dimensional space of the similarity measure to the 2D space by trying to keep the dimension reduction error as small as possible. The result is shown in Fig. 3, where the separation of left and right feet can be clearly seen. This can be regarded as good evidence for the potential of the similarity measure.

In the second step, we wanted to test how good the similarity measures work. For this experiment we took the test study of the 20 persons with five measurements each. We used the quality test described in Sect. 2.3.3 where we regarded the assignment of a data set to an other measurement of the

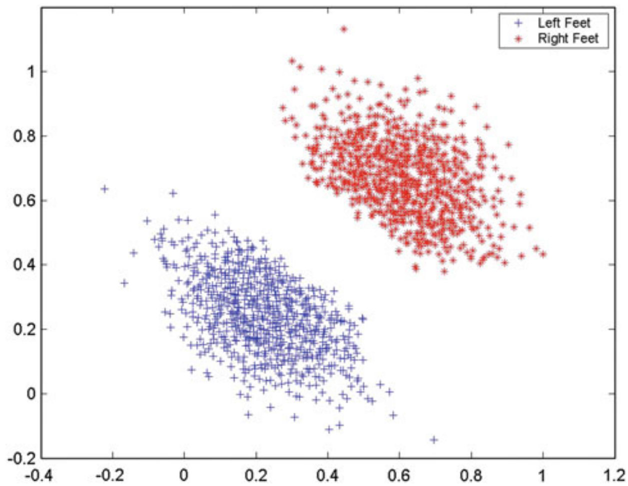

Fig. 3 Sammon map of right and left feet

same person as a correct assignment. Using the volume discrepancies we reached a success rate of 100% and for the weighted feature combination a success rate of 93.7%. This proved that both similarity measures can identify corresponding feet, where the volume discrepancies were more suitable for exact discrimination of persons.

Finally, we tested the clustering systems with the goal that the different measurements of the same persons should cluster together. Using the hierarchical clustering we found that for 18 persons clusters were created where all data sets were put together in one cluster of their own. The same was valid for the iterative clustering. The reason why two persons were not completely clustered correctly, results from the acquisition. These two persons are these two, who significantly changed their weight on the measured foot between the different measurements, as already mentioned in Sect. 2.1. This weight change caused for example a standard deviation of the ball width of the two persons, which is more than five times higher than the mean standard deviation of the other persons. This fact shows the need for strict controlling of the equivalent weight distribution between both feet during other measurements.

3.2 Clustering experiments

After evaluating the usability of our methods we started the investigation on the complete study with the unsupervised learning methods. We used the volume discrepancies as well as the weighed combination of the measured features as similarity measure. Both proposed clustering methods have been used. The hierarchical clustering was used to give the overview of the current interconnections of the data. The iterative clustering was used afterwards with the information gained by the hierarchical clustering to give a more precise clustering result.

The first result of these evaluations was that the factors which are influencing the shape of the foot are not well separated. This means that normally a whole bundle of person specific characteristics influence each other and form together a specific foot shape feature. So our next task was to separate the strong interaction between the observed factors. To be able to do so we created 10 different subsets of the complete study. The data sets of the different subsets are chosen to fulfill various constraints to separate the different effects.

We investigated the clusters from the subsets according to the following factors:

– Gender
– Age
– Height
– Weight
– Overweight person, which means person over 100 kg
– Continent affiliation
– Country affiliation
– Kind and frequency of sport
– Acquisition time of the day
– Job (sitting or standing)

By investigating the subsets from the large study we found that the effects from the different factors strongly interact with each other. Because of this it is not possible to find a large cluster, which is only caused by only one factor. However we build up a hierarchy of different strength groups of the factors influencing the foot shape. To be able to do this we first build subsets where the stronger effects showed their influence. After this step by step the variation in stronger effects were excluded from the subsets by taking only data sets, which have the same parameter value for the strong effect. Due to this it became possible also to observe the weaker effects.

Unfortunately, it was impossible to build up a complete hierarchy of the single effects since in several groups the effects affect each other and some have nearly the same strength. Nevertheless, since some parameters have strong effect on the food shape while others have smaller effects, we were able to well separate between these groups of effects.

3.3 Discrimination on the basis of features

Our final group of investigations faces the task of discriminating between defined groups of data sets. In this experiment, we wanted to know how good our proposed similarity measures are able to perform this discrimination. Furthermore, we tested the single features and searched with the help of a beam search [9] for the best feature combination. For the beam search itself we took the best 100 feature combinations of the antecedent step for every new step. To gain the quality rates we used again the methods introduced in Sect. 2.3.3.

We chose two defined groups as example. The first group is the discrimination between Asia and Europe, a group where strong effects have been observed. The second group is the discrimination between football players and nonfootball players. This is a group with a lot smaller effects. To be independent from the continent effects tested in the first example, we limited this football subset to data sets from Germany.

3.3.1 Discriminating between Asia and Europe

The first example addresses the discrimination between Asian and European feet. The group for this test consists of 2,153 data sets with 52% data sets from Asia and the rest from Europe. Due to the fact that the amount of data sets is not completely equally distributed between Asia and Europe one can only regard values above the smart guess as a benefit of the system. On this sample a smart guess would give a success rate of 52% since this is the amount of Asian data sets.

First of all we investigated the quality of the single features. Table 3 shows the best six results if only one feature is used for the discrimination between Asian and European foot data sets. This table shows that using only one feature gives an indication if the person is from Asia or Europe, but is not enough to provide a clear separation between both groups. These success rates and all other in this paper are classwise averaged rates.

Second, we tried to find the best feature combination by applying a beam search, promoting the best 100 combinations of the antecedent step. Not using all combinations reduces the computation load to a reasonable amount but implies that the result may only be close to the global optimum. The best result found is reached by the combination of five features. This has a success rate of 78.5%, which is far better than the single features. This result is achieved by the combination of the features

- the ball width (2nd feature),
- length to the medial ball point (4th feature),
- the height to medial ankle (6th feature),
- the height to lateral ankle and (7th feature) and
- the angle of the heel (18th feature).

Table 3 Separation of European from Asian feet using only one feature

Height to lateral ankle (7th feature)	63.4%
Ball width (2nd feature)	61.7%
Height of the arch (16th feature)	58.8%
Circumference from heel to wrist (11th feature)	58.8%
Height to medial ankle (6th feature)	58.6%
Circumference of the wrist (9th feature)	55.7%

These results imply that for the differences in shoes for the Asian and European market primarily those features should be taken into consideration.

Finally, we analyzed the weighted combination of the features and the similarity measure of the volume discrepancies with respect to their ability to separate Asia from Europe. Using all features led to a success rate of 74.6%, which is below the best feature combination rate found. The reason for this is that here also features are used which are not appropriate for the separation between Asia and Europe. An example for such a feature is the heel width, which has a success rate on its own of 47.7% which is below the value of the smart guess. The volume discrepancies reach a success rate of 80.6%, which shows that there are significant shape difference in the volume of the foot between the European and Asian foot.

3.3.2 Discriminating between football players and nonfootball players

The second example addresses the separation between football players and nonfootball players. The sample for this test was limited to data sets from Germany to be independent of continent and country effects. The sample consists of the total amount of 406 data sets where 57% of these were playing football. Therefore a smart guess would reach a success rate of 57% and so only rates above this value can be regarded as a real benefit.

Here again we first looked at the single features. Table 4 shows the best six features for discrimination between football players and nonfootball players. In this table, one can easily see that only the first three features would give a real benefit on their own. However, one can state that football playing has a small effect on the position of the ankle.

After this once again a beam search was used to find a good feature combination. Again, the best 100 combinations were considered for the calculation of the next step. Here the success rate improved until the combination of the following nine features were it reached a rate of 72% correct assignments with respect to football playing. The features are

- the length of the foot (1st feature),
- the ball width (2nd feature),

Table 4 Separation of football players from not football players using only one feature

Height to lateral ankle (7th feature)	64.3%
Width between the ankles (12th feature)	63.3%
Width changes of the toe (19th feature)	57.4%
Heel width (6th feature)	55.2%
Length to maximal lateral ball point (4th feature)	54.4%
Circumference of the wrist (9th feature)	53.4%

– the circumference of the ball (3rd feature),
– the length to maximal medial ball point (4th feature),
– the length to maximal lateral ball point (5th feature),
– the width between the ankles (12th feature),
– the wrist height (15th feature),
– the height of the arch (16th feature) and
– the angle of the heel (18th feature).

This shows that a large number of features depend on whether a person is playing football, but the changes are not really strong and only of low significance.

Finally, we used the weighted combination of the features and the similarity measure of the volume discrepancies for classification. The combination of all features led to a success rate of 63%. This is again lower than the rate of the combination of the best feature set found. The explanation for this is again the fact that for the good combination only the features are used, which are appropriate for the task while in the combination of all features also inappropriate ones are used. The volume discrepancies lead to a success rate of 64.5%, which shows that there are no large shape differences in the volume of feet based on whether a person is playing football.

4 Discussion

Investigating the ten build up subsets of the large study we can first of all state that our experiments showed that it is possible to arrange the investigated factors according to their influence on the foot shape. Yet the different factors are not completely well separated, so it was only possible to build up a hierarchy of four different strength groups in which we sorted the investigated features. For building up this hierarchy we used the information which effect dominated the clustering in the different subset and which effect only emerged in subsets where stronger effects were eliminated by construction of the subset. The group with the strongest effect consist of the gender, the weight of the person and if the person is from Europe or Asia. Furthermore it was found that persons with a weight over 100 kg cluster together.

The second group consists of factors with a significantly weaker effect. In this group are the age of the person, the kind of job, the kind of sport and the country that the person is from. Concerning the age it has to be admitted that the large study mainly consists of younger persons. Only less than 5% of the data sets are from persons older than 40 years. However, since already in this study with younger persons an effect can be observed leads to the assumption that with an appropriate study of older persons probably a stronger effect could be observed.

The third stage regards the factors with small effect, which can only be observed if all factors of stage one and many of stage two have been excluded from the subset. The only factor in this stage is the time of the day when the acquisition is taken. From this it can be stated that only for investigations of other small effects it has to be taken into consideration that the data sets were acquired at the same time of the day.

Finally, also the frequency of the sport and the height of the persons have to be mentioned. For the different frequencies of sport we were not able to find any significant effect so we conclude that it has no major general effect on the foot shape. The height of the persons is a special case since when only regarding the effect of the height in respect to its standard deviation it should have been regarded as an effect of the second stage. Yet due to the fact it has a very small absolute effects of below 1 cm it has to be regarded as a effect below the measurement accuracy of this factor.

The second part of the results is based on the experiments of the separation quality according to defined criteria. There it came up that the volume discrepancies are well suited to separate between Asia and Europe which shows that there is a larger shape difference between these person groups. In counterpart to this is the experiment of separating football players from nonfootball players where the volume differences gave only a very small benefit. By additionally taking the results of the feature comparison into account, it leads to the fact that for the continent differences major effects in the foot shape can be found while for the discrimination between football and nonfootball only on some special areas of the foot differences can be found.

Generally, it can be said that for building shoes concerning the differences between Asia and Europe larger volume differences in the foot shape have to be taken in consideration and for the special football shoes only on some small places a specialized change is necessary.

5 Conclusion and further work

Summing up this work one can say that it is possible to compare general 3D point data sets only with the information given by the data set as well as by defining specialized features according to the regarded object. Both methods proved their ability that they are in principle very suitable for measuring the similarity between 3D objects and therefore give the basics for learning algorithms.

By performing an unsupervised learning with these similarity measures, one can obtain a general overview about the facts and the different influences, which have to be considered for the investigation of the object.

However, in the example case of feet it also shows besides the success about the basic intuition that the interaction of different factors make such kind of problem much more complicated and makes it difficult to provide linear dependencies. Therefore there is the need for a large number of data sets to be able to exclude some of the factors to make independent observations. This means that first specialized studies have to be performed on defined person groups to make reliable investigations on these effects.

For a more detailed study more data sets are necessary, which then include a preselection of the measured persons adjusted to the specific problem. By this procedure, then one could observe the factors completely independent and would be able to make quantitative statements.

Acknowledgments The authors would like to thank the adidas AG for giving access to their laser scanner, the large study of measured feet and a good introduction in the field of shoe making and the problems linked to this field.

References

1. Wunderlich, R., Cavanagh, P.: Gender difference in adult foot shape: implications for shoe design. Med. Sci. Sports Exerc. **33**(4), 605–611 (2001)
2. Luximon, A., Goonetilleke, R., Tsui, K.: The customer centric enterprise. Advantages of mass customization and personalization. Footwear Fit Categorization **29**, 491–499 (2003)
3. Mochimaru, M., Kouchi, M., Dohi, M.: Analysis of 3-D human foot forms using the free form deformation method and its application in grading shoe lasts. Ergonomics **43**(9), 1301–1313 (2000)
4. Leng, J., Du, R.: A deformation method for shoe last customization. Comput. Aided Des. Appl. **2**(1–4), 11–18 (2005)
5. Chellappa, R., Wilson, C., Sirohey, S.: Human and machine recognition of faces: a survey. Pro. IEEE **83**(4), 705–740 (1995)
6. Heisele, B., Serre, T., Pontil, M., Vetter, T., Poggio, T.: Categorization by learning and combining object parts. Adv. Neural Inf. Process. Syst. **2**(14), 1239–1245 (2002)
7. Bustos, B., Keim, D., Saupe, D., Schreck, T., Vraniċ, D.: An experimental effectiveness comparison of methods for 3D similarity search. Int. J. Digit. Libr. **6**(1), 39–54 (2006)
8. Sammon, J., John, W.: A nonlinear mapping for data structure analysis. IEEE Trans. Comput. **C-18**, 401–409 (1969)
9. Tillmann, C., Ney, H.: Word reordering and a dynamic programming beam search algorithm for statistical machine translation. Comput. Linguist. **29**(1), 97–133 (2003)