Multidimensional consideration of anthropometric diversity

E. BERTILSSON*†‡ D. HÖGBERG† L. HANSON‡ and Y. WONDMAGEGNE§

† Virtual Systems Research Centre, University of Skövde, Skövde, Sweden

‡ Department of Product and Production Development, Chalmers University of Technology, Gothenburg, Sweden

§Department of Mathematics, University of Skövde, Skövde, Sweden

Abstract

Boundary manikins, the concept of creating statistically extreme cases to accommodate a big part of the less extreme population has been known for decades. Despite this, many ergonomics simulations are done with few human models. This fact can be explained by the time consuming processes when working with many manikins in current digital human modelling (DHM) tools, but may also be a result of difficulty to understand how these manikins are calculated and defined. This paper focuses on the method of defining boundary manikins and how that functionality can be integrated into a DHM tool. Examples of boundary case methods in the literature often use Principal Component Analysis (PCA) which makes it possible to reduce the dimensions of the problem without much loss of the variance of the analysed data. Using PCA often demands some extent of manual analysis at the critical stage of reducing dimensions. This paper will explain a similar methodology for creating boundary manikins from any number of variables, i.e. anthropometric variables chosen as key measurements. This method of creating a group of manikins is intended to be used in an automatic simulation feature in the IMMA software being developed in the associated research project. By using the method, a confidence region in the standardized space is created from eigenvectors and scaled eigenvalues of a correlation matrix. Boundary manikins are chosen at the ends of the axes of the enclosing confidence region, and one manikin of mean values is also added to the group of manikins. In the method presented here, the number of manikins created depends directly on the number of variables, which lead to the fact that the decision making of which key measurements to consider has to be done carefully to not create an overwhelming number of manikins. In comparison with one method using PCA, the method presented in this paper creates more manikins with a bigger difference in the max and min values of the chosen key measurements. If a limited number of cases are of crucial interest, then using PCA to reduce the dimensions of the problem is a good method to use. But if it is possible to create automated simulations the limitation of the number of manikins might not be so important. This will, though, depend heavily on the speed of the automated simulations.

Keywords: Anthropometry, Diversity, Confidence region, Boundary manikin.

1. Introduction

Anthropometry, the study of human measurements, is central in Digital Human Modelling (DHM) due to the necessity of ensuring intended accommodation levels in ergonomics simulations and analyses. When studying larger populations, most body measurements can be considered normally distributed. Still the proportions of the human body vary from person to person, e.g. people of average height do not necessarily have an average value for all body measurements (Pheasant and Haslegrave 2006).

Different approaches exist on how to consider anthropometric diversity in a design context. Two methods frequently used and referenced to are the boundary case method and the distributed case method (Meindl et al. 1993; Speyer 1996; Bittner 2000; HFES 300 Committee 2004). The methodology for these two methods is in many ways similar and they can be used simultaneously. The concept of the methods is that a confidence interval is defined, where boundary cases are points located towards the edges of the interval, and distributed cases are spread throughout the interval randomly or by some systematic approach. Methods concerning multidimensional confidence intervals in literature often use Principal Component Analysis (PCA) which makes it possible to reduce the dimensions of the problem

without much loss of the variance of the analysed data. However, the mathematical procedures are seldom thoroughly described in a simple and complete manner in ergonomics related literature.

In DHM tools, human models can typically be created by quickly defining just stature and weight of a certain gender, age group and nationality or by defining a more complete compilation of a manikin's measurement. The latter is often used when creating a group of boundary manikins, a so called manikin family. Methods of how to create manikin families are described in literature but there are reports that indicate that these methods are seldom being used by industry users today (Bertilsson et al. 2010). Instead, ergonomics evaluations and analyses are often done with few digital human models because of the time consuming process of creating and doing analyses of a complete manikin family. Because of the fact that humans vary a lot in sizes and shapes, there is considerable uncertainty whether the expected proportion of the target population is covered by the analyses being performed by this basic approach. This problem can be reduced by facilitating an automatic generation of manikin families and by using automated batch simulation processes.

This paper describes the definition, creation and implementation of a boundary manikin method. The purpose of the method is to create a confidence region of the selected variables and to define a number of manikins supposed to cover the intended accommodation level. This based on the assumption that using these boundary manikins in subsequent simulations will provide evaluation results that represents the anthropometric diversity of the targeted user group. Boundary cases created by using this method are compared with cases defined by using PCA. In addition, the described method is implemented and intended to be used in automated batch simulations in the IMMA software.

2. Materials and Methods

2.1. Resources and purpose of the method

The analysis was done using both ANSUR anthropometric data (Gordon et al. 1989) and more recent data from a Swedish anthropometric study (Hanson et al. 2009). From the anthropometric data, assumed to be normally distributed, values of interest were extracted, consisting of mean and standard deviation values and a correlation matrix. The eigenvalues and eigenvectors of the correlation matrix and the chi-square distribution were calculated using a C++ numerical analysis library called ALGLIB (Bochkanov 2011). The method for calculating the confidence regions was adopted from literature regarding multivariate statistical analysis (Johnson and Wichern 1992).

2.2. Mathematical procedure

To find the sought boundary cases the confidence region must first be defined. With p number of variables of interest (i.e. number of key measurements), the mean vector $\mathbf{\mu}$ can be denoted as

$$\mathbf{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}. \qquad (1)$$

By statistical analysis of the anthropometric data of the key measurements a correlation matrix $\mathbf{\rho}$ can be defined as

$$\mathbf{\rho} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}. \qquad (2)$$

In the correlation matrix the upper and lower triangular part can be mirrored, meaning $\rho_{ij} = \rho_{ji}$ which gives a ($p \times p$) symmetric matrix. Statistical analysis also gives the ($p \times p$) standard deviation matrix, $\mathbf{V^{1/2}}$ as

$$\mathbf{V^{1/2}} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}. \qquad (3)$$

By using the symmetric correlation matrix and the standard deviation matrix the ($p \times p$) symmetric covariance matrix, $\sum$, can be calculated with

$$\sum = \mathbf{V^{1/2}} \mathbf{\rho} \mathbf{V^{1/2}}. \qquad (4)$$

This gives the covariance matrix as

$$\sum = \begin{bmatrix} 1\sigma_1\sigma_1 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\ \rho_{12}\sigma_2\sigma_1 & 1\sigma_2\sigma_2 & \cdots & \rho_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p}\sigma_p\sigma_1 & \rho_{2p}\sigma_p\sigma_2 & \cdots & 1\sigma_p\sigma_p \end{bmatrix}. \qquad (5)$$

When comparing different normal distributions it is recommended to transform the distributions into standard normal distributions in which the mean values are 0 and standard deviation are 1 (Glenberg and Andrzejewski 2007). This gives each distribution the same significance in the calculations and also gives the covariance matrix directly from the correlation matrix as

$$\sum = \mathbf{\rho} \qquad (6)$$

Eigenvalues and eigenvector of the covariance matrix are sought to describe the confidence region. We consider the ($p \times p$) identity matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}. \tag{7}$$

The scalars $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of the covariance matrix $\sum$ and are given by studying the determinant and solving the following equation

$$|\sum - \lambda \mathbf{I}| = 0. \tag{8}$$

Each of the eigenvalues correspond to a $(p{\times}1)$ vector, the eigenvector $\mathbf{x}$, that is a nonzero vector ($\mathbf{x} \neq \mathbf{0}$) such that

$$\sum \mathbf{x} = \lambda \mathbf{x}. \tag{9}$$

The searched confidence region of a p-dimensional normal distribution is defined by the axes of the multidimensional hyper-ellipsoid. The directions of the axes are given by the normalised eigenvectors which have a length of 1 and are orthogonal to each other. The length of the axes are defined by

$$L_i = k\sqrt{\lambda_i} , i = 1, 2, \ldots, p \tag{10}$$

where the scale factor k is calculated from the chi-squared distribution by

$$k = \sqrt{\chi_p^2(1-P)} \tag{11}$$

where P is the sought accommodation level, e.g. P = 0.9 for 90 % accommodation level.

Values for every variable for each boundary manikin, M_i, can now be defined as the point at the ends of each axis, in both positive and negative direction. When a standard normal distribution is used the axes points defines the value for the manikins in standardized space, and needs to be multiplied by the standard deviation and added to the mean value to get the real anthropometric measurement. These calculations can be combined for one of the directions by

$$\mathbf{M_{1\text{-}i}} = \begin{bmatrix} \mu_1 + x_{1i} \cdot L_i \cdot \sigma_1 \\ \mu_2 + x_{2i} \cdot L_i \cdot \sigma_2 \\ \vdots \\ \mu_p + x_{pi} \cdot L_i \cdot \sigma_p \end{bmatrix}, i = 1, 2, \ldots, p \tag{12}$$

and for the other direction by

$$\mathbf{M_{2\text{-}i}} = \begin{bmatrix} \mu_1 - x_{1i} \cdot L_i \cdot \sigma_1 \\ \mu_2 - x_{2i} \cdot L_i \cdot \sigma_2 \\ \vdots \\ \mu_p - x_{pi} \cdot L_i \cdot \sigma_p \end{bmatrix}, i = 1, 2, \ldots, p \tag{13}$$

where x_{pi} is the p:th value in the eigenvector of the eigenpair with eigenvalue λ_i. It is also suggested that a mean case is created, defined by the mean values for each key measurement as

$$\mathbf{M_3} = \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}. \tag{14}$$

By creating a group of manikins by this method the number of manikins n are directly connected to the number of key measurements p chosen as

$$n = 2 \cdot p + 1. \tag{15}$$

If gender is treated separately, as it is done in many analyses, the number of manikins will be doubled.

3. Results

The results created by the described method are divided into three parts. First part shows an example where the method has been used to calculate and plot an ellipsoid in three dimensions with defined boundary cases at the ends of each axis. The second part covers a comparison with the described method and a similar method which uses PCA to reduce the dimensionality of the chosen key variables. The last part concerns the implementation of the method into the IMMA software.

3.1. Creation of boundary ellipsoid

The described method is here used to create a group of male manikins based on a three dimensional ellipsoid. Key variables are chosen as body weight, stature and sitting height. Anthropometric data is taken from the ANSUR data set. Accommodation level is set to 90 %. Values for the mean and standard deviation for each variable are given in Table 1. Table 2 shows the correlation matrix, from which eigenvalues and eigenvectors are calculated. Values for each manikin are calculated using equation (12), (13) and (14) and can be seen in Table 3. The resulting confidence ellipsoid is shown in Figure 1 in which the ellipsoid's axes in real space can be seen as dashed lines and differs from the boundary cases defined in standardised space.

Table 1: Means and standard deviations of chosen anthropometric variables.

	Mean	Standard deviation
Body weight (kg)	78.49	11.11
Stature (mm)	1755.81	66.81
Sitting height (mm)	913.93	35.58

Table 2: Correlation matrix for the chosen variables.

	Body weight	Stature	Sitting height
Body weight	1	0.546	0.447
Stature	0.546	1	0.741
Sitting height	0.447	0.741	1

Table 3: List of created boundary manikins and values for each key variable.

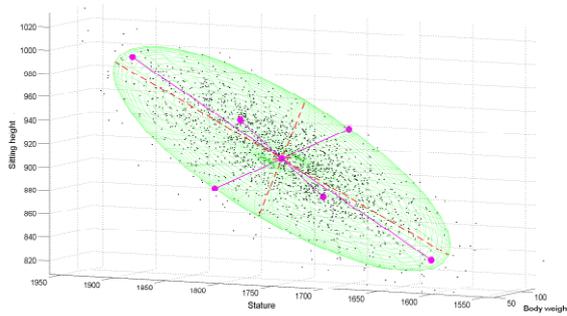|	Body weight (kg)	Stature (mm)	Sitting height (mm)
1	76.24	1818.09	885.54
2	96.37	1725.68	880.64
3	99.61	1907.84	991.39
4	80.73	1693.52	942.31
5	60.60	1785.93	947.21
6	57.36	1603.77	836.46
7	78.49	1755.81	913.93

Figure 1: Depiction of boundary cases and the confidence ellipsoid enclosing 90 % of the multivariate normal distribution for the three key variables: Body weight, Stature and Sitting height.

3.2. Comparison with a method using PCA

Table 4 shows boundary cases created by the described method, here based on the selection of key measurements and data given in Meindl et al. (1993). These cases can be compared with the variable values for the "2-D Man" defined by Meindl et al., seen in Table 5. The method described in this paper creates thirteen cases while the PCA based method defines values for eight cases. There is also a difference in max- and min-values for each variable, where the 13 boundary cases show a bigger span compared to the cases defined by Meindl et al.

Table 4: 13 boundary cases created based on values given in Meindl et al. (1993) with 95 % accommodation level. Values in mm, M1: Thumb tip reach, M2: Buttock knee length, M3: Knee height/Sitting, M4: Sitting height, M5: Acromion height/Sitting, M6: Eye height/Sitting.

	M1	M2	M3	M4	M5	M6
1	803.0	605.5	555.7	953.4	608.3	791.3
2	817.0	627.4	526.3	929.6	613.4	814.0
3	802.4	604.0	556.0	950.3	569.0	834.7
4	874.4	569.1	547.6	932.9	610.0	808.4
5	734.5	552.3	518.3	979.7	655.3	855.9
6	902.9	672.8	630.0	1029.9	690.8	900.9
7	803.2	602.5	559.5	910.2	612.3	827.7
8	789.2	580.6	588.9	934.0	607.2	805.0
9	803.8	604.0	559.2	913.3	651.6	784.3
10	731.8	638.9	567.6	930.7	610.6	810.6
11	871.7	655.7	596.9	883.9	565.3	763.1
12	703.3	535.2	485.2	833.7	529.8	718.1
13	803.1	604.0	557.6	931.8	610.3	809.5
Max	902.9	672.8	630.0	1029.9	690.8	900.9
Min	703.3	535.2	485.2	833.7	529.8	718.1

Table 5: Eight boundary cases defined by Meindl et al. (1993) with 95 % accommodation level. Values in mm, M1: Thumb tip reach, M2: Buttock knee length, M3: Knee height/Sitting, M4: Sitting height, M5: Acromion height/Sitting, M6: Eye height/Sitting.

	M1	M2	M3	M4	M5	M6
A	885,4	663,2	612,5	956,5	627,8	831,6
B	787,7	595,6	541,6	860,3	548,8	742,1
C	720,8	544,9	502,8	907,2	593,3	787,4
D	818,5	612,5	573,7	1003,4	672,3	876,9
W	872,2	651,8	607,7	999,9	666,4	872,8
X	850,4	639,9	585,1	898,7	579,1	777,5
Y	734,0	556,2	507,6	863,8	554,7	746,2
Z	755,8	568,2	530,2	965,0	642,0	841,5
Max	885,4	663,2	612,5	1003,4	672,3	876,9
Min	720,8	544,9	502,8	860,3	548,8	742,1

3.3. Implementation into IMMA software

The method described in this paper is implemented into the IMMA software through a graphical user interface shown in Figure 2. The manikin family functionality is combined with regression equations

that calculates the omitted anthropometric variables needed to define a complete manikin. These equations are based on a hierarchical estimation method for anthropometric variables described by You and Ryu (2005). Weight and stature are on the top of the hierarchy and therefore necessary as key variables due to necessity to know these two variables to be able to compute all regression equations. In addition to weight and stature more key measurements can be added to also be taken into consideration. Added measurements will together with weight and stature define the number of dimensions and therefore also the number of defined boundary cases. It will also be possible to define if manikins of both or only one gender shall be created. Because of the IMMA project's strong focus on Swedish production currently only Swedish anthropometrics is implemented in the boundary case creation functionality.
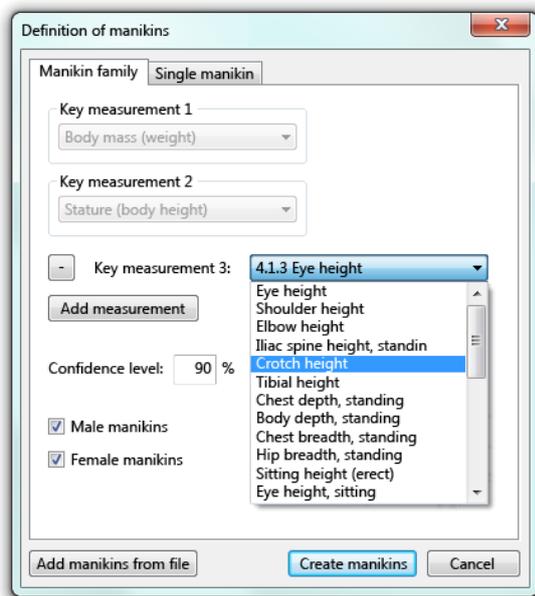

Figure 2: Graphical user interface of the boundary manikin creation functionality in the IMMA software.

4. Discussion

If it is crucial to keep the boundary cases to a minimum, then using PCA to decrease the number of dimensions is adequate to use. Though, when using PCA some extent of manual analysis often has to be done when reducing the number of dimensions. It can also be somewhat difficult to grasp how the boundary manikins are defined in the dimensionality of the principal components.

However, in the IMMA software it will be possible to, by defining some constraints, make automated batch simulations with a group of manikins. This functionality diminishes the need for decreasing the number of manikins, if the simulation time can be kept at a reasonable level. Still it will be important to choose key variables with care. Both because

choosing too many measurements will lead to an overwhelming number of manikins and in some simulations variance of some measurements will not be of interest. Because the manikin family functionality is combined with a hierarchical estimation method it is rational if key variables are chosen in respect to these regression equations. This meaning that measurements higher up in the hierarchy are more suited as key variables.

It can also be argued that choosing just one case at the ends of each axis is not enough to fully consider the anthropometric diversity. Reasons why this approach was chosen is both that these cases easily can be identified and also that the number of cases will not increase exponentially when increasing the number of key measurements. If the approach of just defining axis points without adding intermediate octant points (for example) is adequate can be debated and needs to be further investigated. Therefore, future research will investigate how boundary and distributed cases are to be defined within multidimensional normal distributions.

5. Conclusion

This paper describes the definition, creation and implementation of a boundary manikin method which creates a confidence region of selected variables and defines a number of manikins supposed to cover the intended accommodation level. The method is intended to be used in automated batch simulations, which will decrease the simulation time when using DHM tools for the design of products and workplaces.

Acknowledgement

References

Bertilsson E, Svensson E, Högberg D, Hanson L, 2010. Use of digital human modelling and consideration of anthropometric diversity in Swedish industry. Proceedings of the 42nd annual Nordic Ergonomic Society Conference, Stavanger, Norway, September 6-8, ISBN 9788299574723.

Bittner AC, 2000. A-CADRE: Advanced family of manikins for workstation design. In: Proceedings of the IEA/HFES 2000 Congress. Human Factors and Ergonomics Society, Santa Monica, CA.

Bochkanov SA, 2011. ALGLIB® - numerical analysis library. Retrieved February 21, 2011, from http://www.alglib.net/

Glenberg A, Andrzejewski M, 2007. Learning from Data: An Introduction to Statistical Reasoning, 3rd edition. Routledge Academic, New York, ISBN 9780805849219.

Gordon CC, Churchill T, Clauser CE, Bradtmiller B, McConville JT, Tebbetts I, Walker RA, 1989. 1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics. Technical Report Natick/TR-89-044. U.S. Army Natick Research, Development and Engineering Center, Natick, MA.

Hanson L, Sperling L, Gard G, Ipsen S, Vergara CO, 2009. Swedish anthropometrics for product and workplace design. Applied Ergonomics, Volume 40 (4), 797-806.

HFES 300 Committee, 2004. Guidelines for Using Anthropometric Data in Product Design. Human Factors and Ergonomics Society, Santa Monica, CA.

Johnson RA, Wichern DW, 1992. Applied multivariate statistical analysis (3rd ed.). New Jersey: Prentice-Hall, Inc.

Meindl RS, Hudson JA, Zehner GF, 1993. A Multivariate Anthropometric Method for Crew Station Design. Crew Systems Directorate, Human Engineering Division, Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio. Technical Report AL-TR-93-0054.

Pheasant S, Haslegrave CM, 2006. Bodyspace: Anthropometry, Ergonomics and the Design of Work, 3rd ed., Taylor & Francis, Boca Raton. ISBN 0415285208.

Speyer H, 1996. On the definition and generation of optimal test samples for design problems. Kaiserslautern, Human Solutions GmbH.

You H, Ryu T, 2005. Development of a hierarchical estimation method for anthropometric variables. International Journal of Industrial Ergonomics 35 (4), 331-343.