INTERNATIONAL DESIGN CONFERENCE - DESIGN 2004
Dubrovnik, May 18 - 21, 2004.

AUTOMATIC LANDMARKS PREDICTION USING THE ARTIFICIAL NEURAL-NETWORK-BASED TECHNIQUE ON 3D ANTHROPOMETRIC DATA

B. Zhang, J.F.M. Molenbroek, I. Horváth and C. Snijders

Keywords: artificial neural network, landmark, configuration space, 3D anthropometric data

1. Introduction

There are a lot of 3D anthropometric techniques, such as contact measurement (template method; multiple probe method; casting method; anthropostereometry) and non-contact measurement (stereophotogrammetry; Moire contourgraphy; monophotometry / contour projector; computer vision techniques, triangulation-based spot sensing; light strip sensing; time-of-light-based ultrasonic range finder; laser range finder. Since the introduction of these 3D anthropometric techniques, engineers and ergonomists have sought to exploit the potential of this exciting technology.
The added components of shape provided by 3D measurements offer a more detailed description of human variation compared with traditional manual 1D or 2D data. However, there are few body shape description methods based on key anthropometric dimensions (circumference or height) (Mollard, 2002). Additionally, semantic descriptors such as esomorphic/dedomorphic are fuzzy and mathematically ill-defined. Incomplete human body description methods result in limitations and difficulties for the application of the 3D scanning surface. On the other hand, in the area of anthropometric modeling, the concrete problem is how to relate large amounts of 3D coordinates with the body morphology. Although anthropometric shape analysis strategy was described to analyze the full range of body sizes and shapes in terms of curvatures, it is still a method based on 2 dimensions. In other words, it is still not a proper method to apprehend the anatomical shapes and their variation relative to three-dimensional space.
With the development of computer technique, surface construction became increasingly simple and accurate . Therefore, 3D human surface anthropometric data presses for a 3D solution to provide more accurate and complete information for ergonomics and human-centered industrial design.In the field of biology, landmarks extracted from 3D scanning data can be considered as a reduced 3D configuration of the human body (Bookstein, 1991). In other words, landmarks are a reduced descriptor of 3D data. In this sense, the landmark-based shape analysis methods will simplify the modeling procedure, because it deals with the landmarks instead of the large amount of 3D coordinates. This paper is focused on the landmark-based 3D anthropometry research.

1.1 Previous methods of analysis of 1D anthropometric data and geometric methods in 3D anthropometric data

To perform a shape analysis, a biologist traditionally selects ratios of distances between landmarks or angles, and then submits these to a multivariate analysis. This approach is called 'multi-variate morphometrics' in biology. Similarly, traditional anthropometry extracts 1D or 2D measurements from samples and sends these to statistical analysis, further providing this statistical data for product design or workspace design in percentile or in multi-variable results (figure 1). In the studies of multivariate morphometrics one deals exclusively with positive variables (length, angles and ratios of lengths) (Robinette et al., 1997). However, to consider only distances and angles can be inferior to using the actual coordinates of the landmarks, because the geometry is often thrown away when using the former. Ratios of distances can easily be calculated from coordinates, whereas the converse is not generally true. A considerable amount of work was carried out in multivariate morphometrics using distances, ratios, angles, etc. and it is still very commonly used in both biology and anthropometry (Roebuck, 1995).

There are many situations, such as classification problems, where the techniques can be very powerful. However, sometimes the interpretation of the important linear combinations of ratios of lengths and angles can be difficult.

With the development of 3D anthropometric techniques, it is possible to work on the landmark coordinates directly (see figure 2 and 3). The idea is that, rather than working with quantities derived from organisms, one works with the complete geometrical object itself. In particular, we shall consider a shape space obtained directly from the landmark coordinates, which retains the geometry of a point configuration. This approach to shape analysis is called ‘geometric shape analysis’ by various authors and the subject progressed rapidly around the late 1970s/early 1980s. Bookstein (Bookstein, 1991) summarized his view of the history of geometrical shape analysis, mainly through applications in biology. The research methodology of our research regarding 3D anthropometric data is based on the theory and definitions from corresponding knowledge in biology.

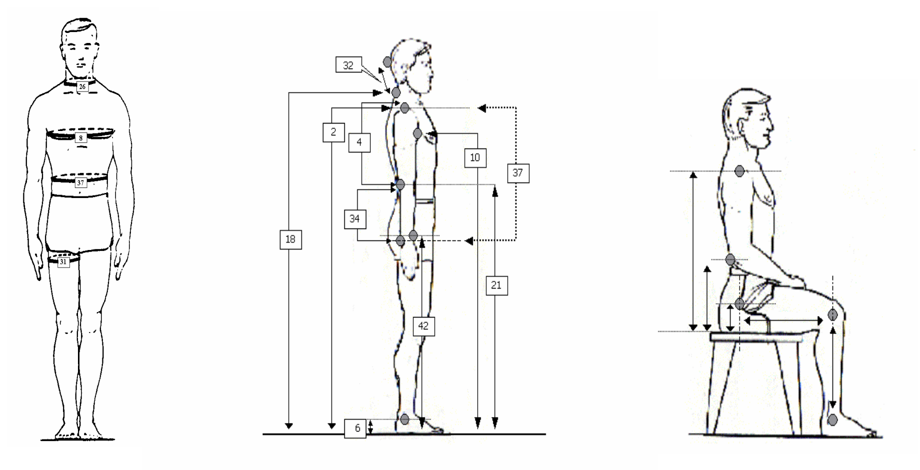

Figure 1. Traditional anthropometric measurements in one dimension

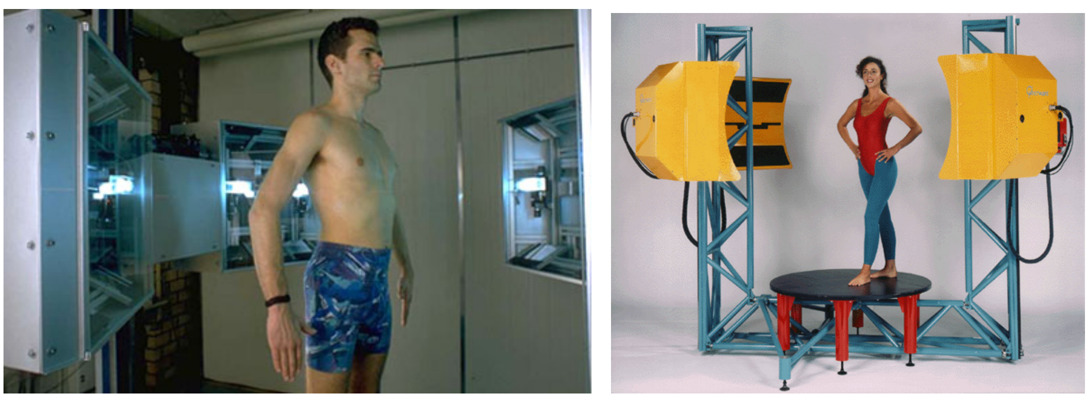

Figure 2. (top) Tecmath/Vitronic/Vitus Pro (The Netherlands). The resolution of the system was 2mm vertically and 1mm horizontally); (bottom) Cyberware WB4 (North America/Italy)

Figure 3. Three postures in scanning (CAESAR project)

The raw 3D surface of human body anthropometric data is available in many common known formats which are easily transferable and can be used in computer image processing (Robinette et al., 1997). The landmarks can be recognized directly by the computer program. However, current research on the application of 3D human body data is limited in acquirement of 1D or 2D measurements as the one of traditional anthropometry.

1.2 Introduction to Artificial Neural Networks and their application in ergonomics

Artifical Neural Network has gained much attention in the last decade as an effective artificial intelligent technique. ANN is used in many kinds of applications. For example, Marcle et al. (1999) employed a constrained generative ANN model to hand posture recognition for real-time computer visualization purpose. The goal of the constrained generative learning is to closely fit the probability distribution of the set of hands using a non-linear compression neural network. A small set of hand postures was selected (5 kinds of postures). And a database of thousand different hand posture images with both uniform and complex backgrounds was built. The research result shows ANN can effectively recognize these hand postures. The mean detection rate is 93.4%.

Lim (1996) examined the potential of neural network analysis to predict the range of anatomical joint motion for the design/layout of workstation and tasks (Lim, 1996). The posture and motion data were recorded with a flexible electrogoniometric system. A feed-forward back-propagation neural network was employed to predict the joint motions. In fact, various methods based on artificial intelligence techniques, claiming to be a universal approximator, are proposed as alternatives to statistical methods, especially to model highly non-linear functional relationships. An adaptive neuro-fuzzy inference system (ANFIS) was employed to estimate anthropometric measurements (Kaya, 2003). It was found that ANFIS performs better than the stepwise regression method (traditional statistical method).

A pilot study of the human head form with Radial Basis Neural Network (RBNN) was employed to learn 3D human head surface landmarks (Zhang et al., 2002). The original scanned coordinates are shown almost overlapping new coordinates predicted by RBNN. This study proved that the neural network could be used to approximate and predict human head surface landmarks precisely with limited performance.

As a result, neural networks could be a possible candidate approach to study 3D landmarks from 3D scanned anthropometric data due to the capabilities of the learning mechanism. However, different types of neural networks have different characteristics and therefore different fields of application.

Basically, RBF neural networks and BP neural networks are all useful to approximate any non-linear functions. The main difference between them is that due to the different ‘working functions’, the neurons of the hidden layers of BP neural network employ sigmoid functions, where the value of the functions are non-zero values in the infinite area of the input space, while the ‘working functions’ of RBF neural network are local. From the above analysis, we chose BP neural networks in our study of Automatic Landmarks Prediction based on 3D anthropometric data.

1.3 Aim of this research and structure of the paper

This research is one part of a research project which is aimed at building a module of a computer-aided ergonomic tool based on 3D anthropometric data (figure 4).

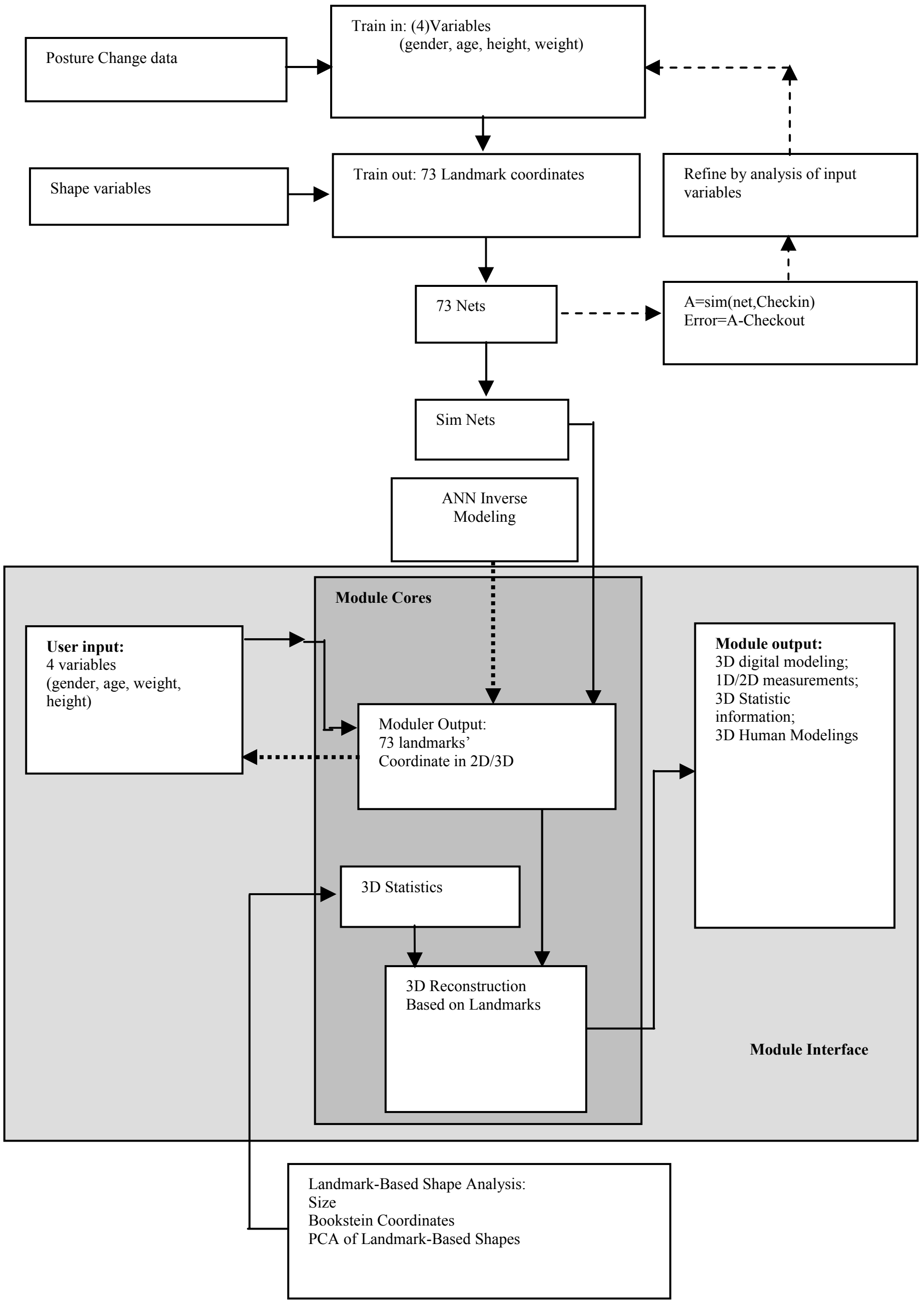

Figure 4. Research flow chart based on 3D anthropometric data towards a module of a computer-aided ergonomic tool

The goal of this first research is to get one of the 73 landmarks sim nets. With these sim nets, the user is able to obtain the anthropometric data in 3D for their product design. In this research, due to the number of input samples, the input variables are gender, age, weight and height. The output at this moment consists of the expected polar coordinates of the corresponding landmarks. These polar coordinates of landmarks can be used to reconstruct a 3D digital human modeling system for design purposes.

Figure 5 illustrates the ANN roles in automatically predicting the 3D landmark coordinates of human body space and when posture variables are used as input to ANN, the landmarks coordinates can be predicted automatically in different postures. The structure of this paper is as follows. Section 1 introduces the basic theory and definitions. Section 2 describes the research method. The results and fitting errors are analyzed in section 3. Finally, section 4 offers some suggestions to increase the accuracy and stability of ANN. Additionally, future research contents are described in section 5.

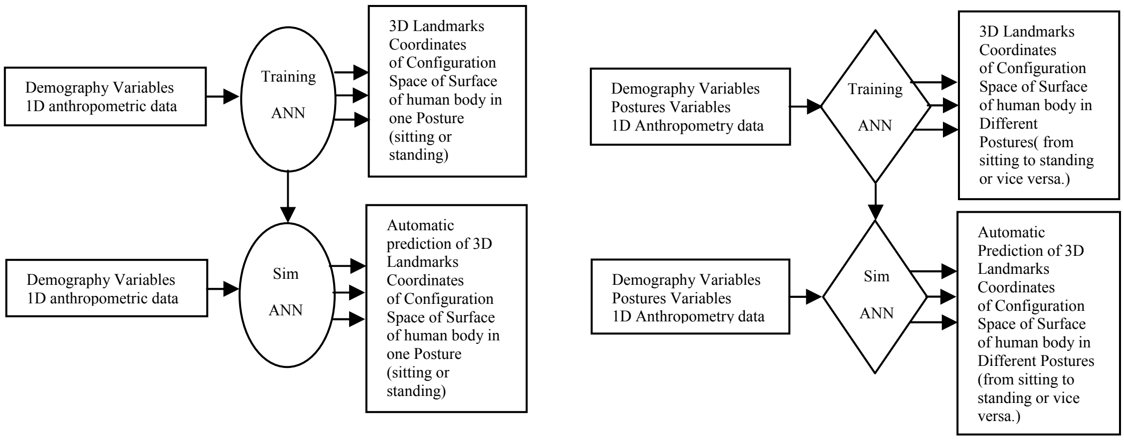

Figure 5. (left) ANN application for Automatic Landmarks Prediction coordinates of the human body based on 3D scanning techniques;(right) ANN application for Automatic Landmarks Prediction with coordinates of the human body from sitting posture to standing posture (and vice versa), based on 3D scanning techniques

2. Methodology

2.1 Approaches

Because of the ongoing research, this paper only focuses on the first step of ANN application on 3D anthropometric data. In other words, this paper is concerned with the prediction of landmarks coordinates on the basis of 4 input variables, which are gender, age, height and weight. The following are approaches which were employed in this research: 1) design of the ANN construction; 2) analysis inputs; 3) changing of coordinates; 4) normalization; 5) sample extension; 6) training; 7) sim net; 8) error analysis; 9) net performance test.

2.2 Definition of shape, size-and-shape, landmarks and configuration space

Advances in technology have led to the routine collection of geometric information and the study of the shape of objects has become increasingly important. Shape analysis is of great interest in a wide variety of disciplines. Some specific applications follow from biology, medicine, image analysis, archaeology, geography, geology, agriculture, and genetics. The definition of shape is considered in our research.

Definition 1. Shape is all the geometric information that remains when location, scale and rotational effects are filtered out from an object.Definition 2. Size-and-shape is all the geometric information that remains when location and rotational effects are filtered out from an object.We will describe shape by locating a finite number of points on each human body which are called landmarks.Definition 3. A landmark is a point of correspondence on each object that matches between and within populations. Definition 4. *Type A landmarks* occur at the joints of tissues/bones; *type B landmarks* are defined by local properties such as maximal curvatures and *type C landmarks* occur at extremal points or constructed landmarks, such as maximal diameters and centroids.

Normally, anatomical landmarks are usually of type A or B and mathematical landmarks are usually of type B or C. Pseudo-landmarks are commonly taken as equi-spaced along outlines between pairs of landmarks of type A and B, and in this case the pseudo-landmarks are type C landmarks. Type A landmarks are usually the easiest and most reliable to locate and type C are the most difficult and least reliable to locate. The research of this paper is based on landmarks following the above definitions.Definition 5. The configuration is the set of landmarks on a particular object. The configuration matrix X is the $k \times m$ matrix of Cartesian coordinates of the k landmarks in m dimensions. The configuration space is the space of all possible landmark coordinates(Dryden and Mardia,1998). In our research we have k=73 landmarks in m=2 or m=3 dimensions and the configuration space is typically R^{km}.

2.3 Samples

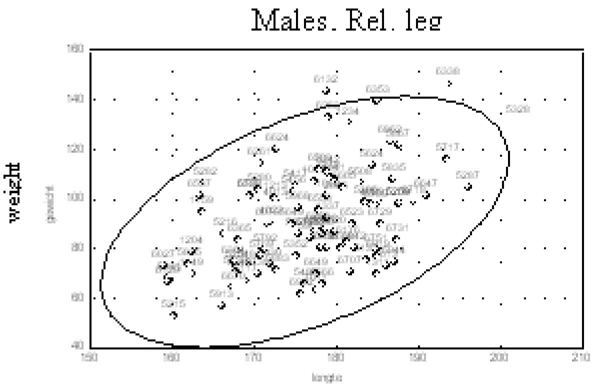

Figure 6. Samples with variables of leg length and weight

Table 1. Samples with 4 input variables and 3 output variables

A	B	C	D	E	F	G	H
Sample number	x1Gender	x2 Age	x3 Height	x4 Weight	x	y	z
1248	1	42	190	65	-39,25	8,57	777
1251	1	47	194	85	-39,8	-9,39	816,82
1449	2	26	143	52	60,69	27,03	353,04
4042	1	53	178,5	87	-21,15	43,72	658,82
5208	2	36	180	88	-2,19	19,9	673,24
5282	1	59	167	105	44,81	48,2	543,19
5287	1	47	196	100	-11,94	9,44	866,37
5317	1	46	185	91	-20,26	17,6	757,49
5440	2	59	166	122	41,63	18,85	537,61
5514	2	39	186	70	-4,33	12,82	758,62
5525	2	33	176	105	1,31	31,35	666,44
5590	2	51	168	92	-2,06	3,9	558,02
5649	1	45	178	87	25,6	57,86	673
5903	2	40	195	81	11,47	8,6	837,08
5913	1	68	170	55	-60,32	-12,88	564,76
5953	2	60	168	78	33,2	22,33	528,18
6023	2	27	174	60	-6,86	-12,37	598,84
6027	1	50	160	74	50,79	23,15	502,4
6114	1	38	205	100	-12,86	-2,2	916,57
6268	2	62	155	67	1	43,07	461,22
6353	1	39	187	135	-17,02	18,67	733,44
6486	2	22	169	50	1,46	8	578,43
6550	2	47	175	126	-1,88	39,09	629,17
6551	1	43	203	80	-23,14	12,23	911,71
6562	2	55	150	92	47,6	47,99	411,66
6624	1	57	172	110	-13,15	-28,52	609,48
6701	2	22	182	57	-20,46	-28,83	682,26
6738	1	23	174	65	-11,65	4,42	664,92
6754	2	23	159	49	32,83	31,22	526,85
6803	1	22	192	80	-14,55	6,44	774,01
6950	2	32	176	63	4,73	7,93	618,15
6992	1	21	180	65	17,81	46,45	655,68

The raw data are from the CAESAR project (The Civilian American and European Surface Anthropometry Resource), a survey of body measurements for people aged 18-65 in three countries, the United States, the Netherlands, and Italy (Figure 6). The raw data of the scanned 3D surface of a human body is in STL format which contains polygon mesh objects. The research variables in this paper include demography information, 1D anthropometric data, and coordinates of one landmark of 73 whole body landmarks scanned by laser technique. The total number of the scans in our research is 32, among those, 28 scans were used to train the neural network, 4 scans were used to test the performance of the neural network (table 1).

2.4 Neuron model and network architecture

In this research, with the most commonly used back-propagation algorithm, the multiple layer feed-forward network was employed. An elementary neuron with R inputs is shown in figure. Each input is weighted with an appropriate ω. The sum of weighted inputs and the bias forms the input to the transfer function f. In this research, the neurons used two differentiable transfer functions f to generate their output. One is tansig (tan-sigmoid transfer function), the other is purlin (linear transfer function).

$$a = f(W_p + b) \tag{1}$$

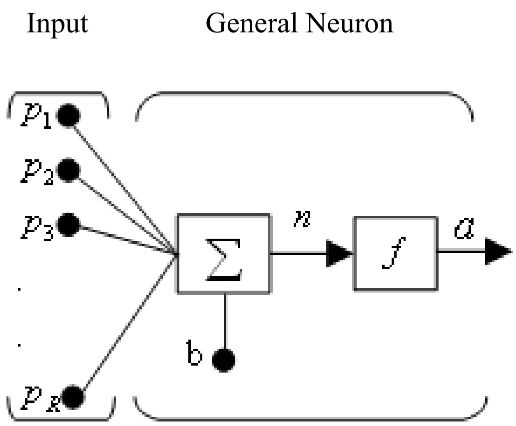

Figure 7. Neuron model

Where R is the number of elements in the input vector (figure 7).

Feed-forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. We chose a feed-forward network with one input layer, one hidden layer, and one output layer. Both input and hidden layer have sigmoid neurons with tansig transfer functions; the output layer has linear neurons with purlin transfer functions. As the algorithm of TrainLM appears to be the fastest method of training moderate-sized feed-forward neural networks (up to several hundred weights), we employed TrainLM in our research. Three-layer feed-forward networks were used in our research. The architecture is shown in the following figure (figure 8). There are two hidden layers and one output layer, which have 4, 8 and 3 neurons respectively. In other words, S1=4, S2=8 and S3=3 in figure 8.

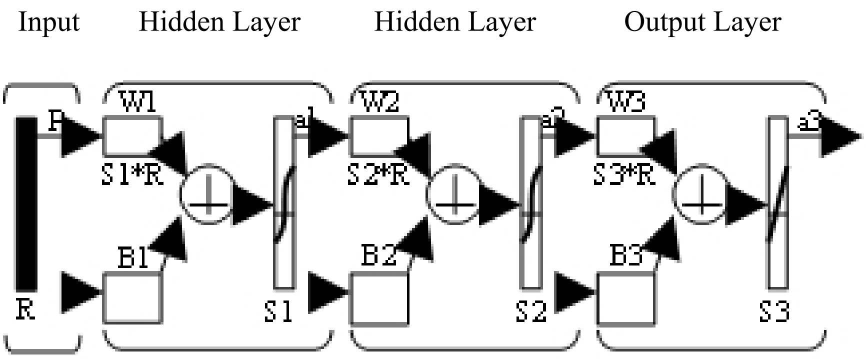

Figure 8. Network architecture design

Where P is input, R is the number of input vectors, S is the number of neurons in layers, W is the input weight, B is the input bias.

3. Results and Discussion

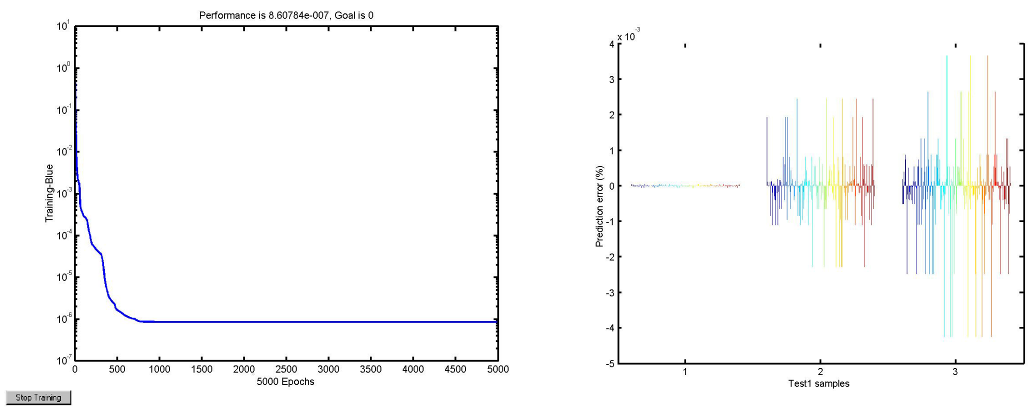

Figure 9. (left) Development of the error while training; (right) The prediction error (x 10^{-3})

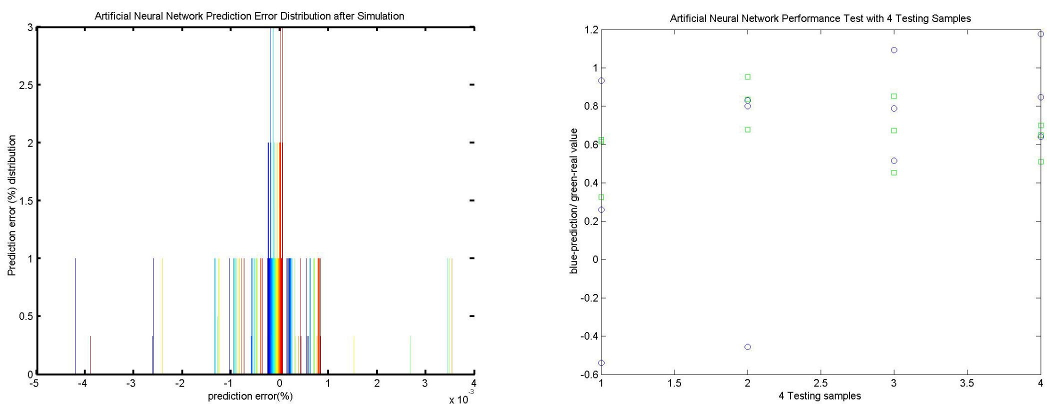

Figure 10. (left) The distribution of prediction errors; (right) Feed-forward back-propagation artificial neural network performance checked with 4 testing samples (squares are prediction value, circles are real value)

In total, three different types of BP neural networks with different numbers of neurons in the hidden layer were trained with the same inputs. They are 4*6*3 neurons BP NN, 4*8*3 neurons BP NN, and 4*10*3 neurons BP NN. Training results show that the 4*6*3 neurons BP NN is more stable than the other two kinds of BP NN in this research.

Figure 10 illustrates that the trainings have been constringency to 10^{-6} with 5000 epochs. The training results of 4*8*6 BP NN show the feed-forward BP neural networks can be employed to predict landmark coordinates using demography information and 1D anthropometric data effectively. Figure 10 (right) shows the polar coordinates of 4 testing scans. The blue squares are prediction values and the green circles are the real values of the 4 test samples. The prediction error was analyzed (figure 10 left). The prediction error is in normal distribution and its value is mostly lower than 0.001. The insufficiency of the input data is the main cause of inaccurate results. In other words, the number of scans is very important for the training accuracy. In this research, only 28 scans enlarged 50 times were done first and then randomly sent to input neurons to train them.

4. Conclusion and Future work

Results of this research indicate that BP neural networks are capable of memorizing and predicting the landmarks of the surface of the human body. The learning of ANN is not stable with different numbers of neurons in the hidden layer. In other words, once the function of Newff runs, the initialization of ANN kept changing, and this leads to inaccurate results. The BP neural network analysis is being refined to improve the prediction.

When ANN are trained and saved, by Sim (nets), once users input demographic variables, the networks will provide users with 73 corresponding 3D landmarks coordinates in 2D or 3D; furthermore, a 3D digital modeling will be constructed based on these landmarks. 73 landmarks express the new and reduced human surface characteristics compared with a 1D measurements extract in traditional anthropometry. However, ANN can further help to look for simplified characteristics (which could amount to less than 73 landmarks). With inverse modeling, the ANN can predict the individual's gender, age, weight and other characteristics of demography.

Acknowledgments

I would like to thank Dr. B.Yu who was helpful in this research regarding neural network design.

References

Bookstein, F.L., Morphometric tools for landmark data, Cambridge University Press.1991.

Dryden, I.L., Mardia, K.V., Statistical Shape Analysis, John Wiley & Sons Ltd.,UK, 1998, p. 3.

Daanen, H. A. M., Robinette, K. M., CAESAR, the Dutch dataset, Report TNO-TM-01-C026. 2001.

Kaya, M.D., Hasiloglu, A.S., Bayramoglu, M., Yesilyurt, H. and Ozok, A.F., A new approach to estimate anthropometric measurements by adaptive neuro-fuzzy inference system, International Journal of Industrial Ergonomics, 32,2003, p. 105-114.

Lim, S.Y.E., Neural Network Investigation of Posture and Motions, Contemporary Ergonomics,1996, p. 539-544.

Marcel, S. and Bernier, O.,Hand Posture Recognition in a Body-Face Centered Space, GW`99, LNAI 1739,1999, p. 97-100, Springer-Verlag, Berlin, Heidelberg.

Mollard, R., Shape modeling driven by the product design, Computer-Assisted Radiology, 16th International Congress and Exhibition, Towards a World Engineering Anthropometry Resource-WEAR, 2002.

Robinette, K.M., Vannier, M.W., Rioux, M., and. Jones, P.R.M, AGARD Advisory Report 329: 3-D Surface Anthropometry: Review of Technologies, Advisory Group for Aerospace Research & Development, Neuilly-Sur-Seine, France.1997.

Roebuck, J.A.,Devices and Procedures for Measurements, in: Anthropometric Methods: Designing to Fit the Human Body, 1995, p. 37-47.

Zhang, B., Molenbroek, J.M.F., Snijders, C.J.,Fundamental Research of Mathematical Model of the Human Body, In: Proceedings of the Digital Human Modeling Conference, Munich, 18-20 June, 2002.

Zhang B. Msc.
Delft Unniversity of Technology, Industrial Design Engineering
Landbergstraat 15, 2628CE, Delft. The Netherlands
Telephone: 0031-15-2787195, Telefax: 0031-15-2787179
E-mail: b.zhang@io.tudelft.nl