Virtual Fitting Room using Webcam

Chinmay Barde, Soham Nadkarni, Nikhil Joshi and Saurabh Joshi

***Abstract*— Fashion coordination is one of the self-expressions which have been always in demands. Searching for perfect attire, is a time consuming task as well as many factors need to be kept in mind. In this paper we are introducing a "Virtual Fitting Room". This is an innovative Virtual shopping infrastructure, enabling customers to visualize themselves wearing garments present in traditional stores, as well as online (in internet shops). This is done by mining of the user image, alignment of models and skin color detection of image (clicked from a fix distance).The major reimbursement of the VFR include, saving time of the customer/user by avoiding don and doffing at the time of shopping, where both the virtual and physical worlds are combined. This application will be able to fill a big gap between customer and seller by showing clothes of varying size. Finally the model is superimposed on the user in real time with some manual adjustment.**

***Index Terms*— Virtual Fitting Room (VFR), Haar Classifier, Virtual try-on, 2D-model.**

I. Introduction

There is substantial, loss of time in don and doffing of clothes in stores which is one of the most time-consuming tasks. Usually long waiting periods have to be taken into account, for example, when standing in front of full fitting rooms. Even, additional time is lost while don and doffing, and also most consumers are hesitate to purchase garments from any online site or they are unsatisfied with their online shopping experience.

Clothing descriptors of anatomical types are more varied and less scientific, e.g. "outsize", "flat-chested" or, "pear-shaped". Information to date on body shapes is largely anecdotal and most clothing is made to fit a small number of stands, which are hoped to represent "average" sizes. The justification is historic custom and practice, with little consistency in the market place and continuing customer concerns about fit. Shape analysis allows the correct averaging of body shapes which fall into a particular size category, enabling improved mannequins (real and virtual) to be made.

The techniques discussed in this paper can enhance the shopping experience. In this paper we will introduce a Virtual Fitting Room system, which offers a solution for the mentioned aspects. This application is based on software which helps in representing output from the skeleton, extracted from image (taken from camera). If a person is standing in front of the camera, the person will be able to select desired clothes. Also in future we can extend our system to recommend some clothes which will suit on that particular person depending on his skin color. The selected garment is then virtually superimposed with the image recorded by the camera. A system employs one Kinect sensor and one High-Definition (HD) Camera. System has been deployed since April 2012 in one of Singapore's largest shopping centers. In order to achieve a believable virtual try-on experience for the end user, the Kinect sensors are used. These sensors are built by Microsoft and are very expensive. HD camera to replace the role of Kinect's built-in RGB camera is included for HD recording. This necessitates a calibration process between the HD camera and the Kinect depth camera in order to map the 3D clothes seamlessly to the HD video recording of the customers.

Virtual try-on system consists of a vertical TV screen, a Microsoft Kinect sensor, an HD camera, and a desktop computer. Fig. 1 shows the front view of the Interactive Mirror together with the Kinect and HD camera. The Kinect sensor is an input device marketed by Microsoft, and intended as a gaming interface for Xbox 360 consoles and PCs. It consists of a depth camera, an RGB camera, and microphone arrays. Both the depth and the RGB camera have a horizontal viewing range of 57.5 degrees, and a vertical viewing range of 43.5 degrees.

Kinect can also tilt up and down within -27 to +27 degrees. The range of the depth camera is [0.8_4]m in the normal mode and [0.4_3]min the near mode. The HD camera supports a full resolution of 2080 _ 1552, from which Virtual Try-on using Kinect and HD camera.

Fig. 1: The front view of the Interactive Mirror with Kinect and HD camera placed on top.

Manuscript received March 09, 2015.

Chinmay Barde, Soham Nadkarni, , Nikhil Joshi, and Saurabh Joshi, Department of Computer Engineering, G.H.Raisoni Institute of Engineering and Technology, Wagholi, Pune.

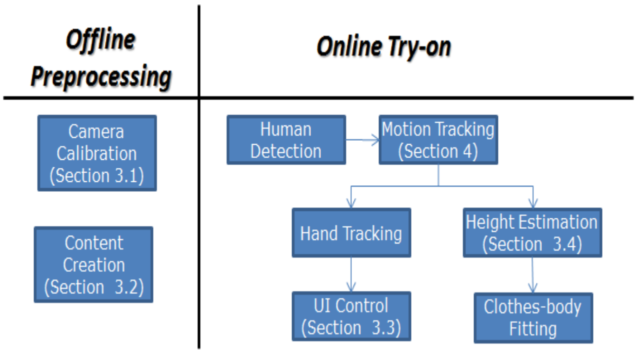

Fig. 2: Major software components of the virtual try-on system.

Fig. 2 illustrates the major software components of the virtual try-on system. During the offline preprocessing stage, we need to calibrate the Kinect and HD cameras, and create 3D clothes and accessories. These two components will be discussed in more details in Sections 3.1 and 3.2 respectively. During the online virtual try-on, we first detect the nearest person among the people in the area of interest. This person will then become the subject of interest to be tracked by the motion tracking component implemented on two publicly available Kinect SDKs, as will be discussed in Section 4. The user interacts with the Interactive Mirror with her right hand to control the User Interface (UI) and select clothing items.

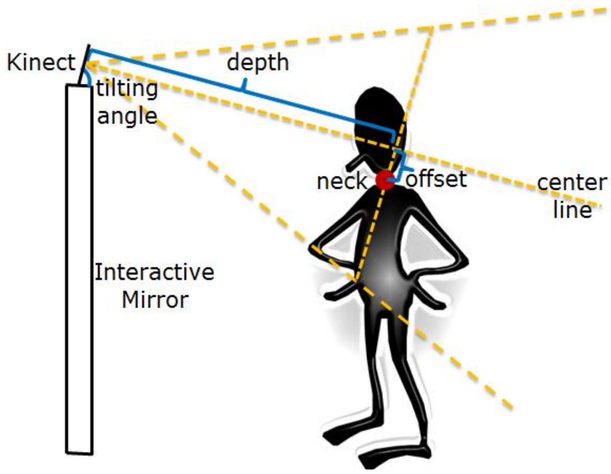

Fig.3 Shoulder height estimation when the user's feet are not in the field of view of Kinect. The tilting angle of the Kinect sensor, the depth of the neck joint, and the offset of the neck joint with respect to the center point of the depth image can jointly determine the physical height of the neck joint in the world space.

II. Proposed System

Proposed VFR is software based and designed to be universally compatible as long as the device has a camera. The use of web camera is a cheaper alternative to Kinect sensors. It does not require extra hardware support. The users can use the proposed system from their home itself. It provides real time access. Compared to other existing VFR systems, key difference is the proprietary hardware components or peripherals. The system makes the use of web cam to detect the human body. The body is then divided into upper body and lower body. Resizing of the images is done to superimpose the cloth image on the human body. This is cheaper version of the existing system which uses lot of hardware and cannot be used at home.

Fig.4 proposed view of the system

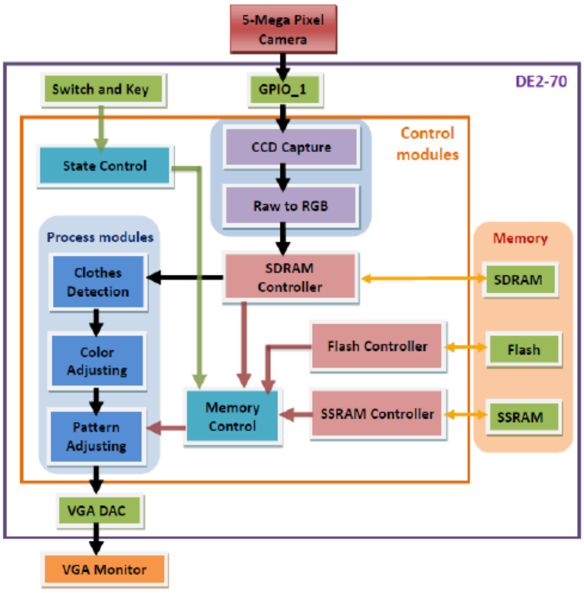

Fig.5 System architecture

III. Methodology

A. User Extraction

Extraction of user allows us to create an augmented reality environment by isolating the user area from the image and superimposing it onto a virtual environment in the user interface. Furthermore, it is here a useful way to determine the region of interest that is also used for skin detection which is explained in further section The camera provides the image. When the device is working, image is segmented in order to separate background from the user [7]. The background is removed by blending the RGBA image with the segmented image for each pixel by setting the alpha channel to zero if the pixel does not lie on the user.

B. Skin Segmentation

Since the model is superimposed on the top layer, the user always stays behind the model which restricts some possible actions of the user such as folding arms or holding hands in front of the t-shirt. In order to solve that issue skin colored areas are detected and brought to the front layer [12]. HSV and YCbCr color spaces are commonly used for skin color segmentation. In this work we preferred YCbCr color space and the RGB images are converted into YCbCr color space by using following equations:

C_b =128-0.169R-0.33G+0.5B
Y = 0.229R+0.587G+0.114B
C_r = 128+0.5R-5.419G

Fig. 7. Background removal. Depth image (left), color image (middle), extracted user image (right)

Chai and Ngan reports that the most representative color ranges of human skin on YCbCr color space [5]. A threshold is applied to the color components of the image within the following ranges:

$77 < C_b < 127$
$177 < C_r < 173$
$Y < 70$

Since we have the extracted user image as a region of interest the threshold is applied only on the pixels that lie on the user. Thus, the areas on the background which may resemble with the skin color are not processed. The skin color segmentation is illustrated in Figure 8.

Fig. 8. Skin colour segmentation. User image (left),segmented image(right)

TABLE 1:

STANDARD BODY PARAMETERS FOR FEMALE

Size	Height (In/Cm)	Weight (Lbs/Kgs)	Chest (In/Cm)	Waist (In/Cm)	Hips (In/Cm)
4	5.4 - 5.6	95 - 105	29 - 31	24 - 26	32 - 34
	162 - 167	42 - 47	73 - 78	60 - 66	81 - 86
6	5.5 - 5.7	105 - 115	31 - 33	26 - 28	34 - 36
	165 - 170	47 - 51	78 - 83	66 - 71	86 - 91
8(1)	5.6 - 5.8	115 - 130	33 - 35	27 - 29	36 - 38
	167 - 172	51 - 58	83 - 88	68 - 73	91 - 96
8(2)	5.8 - 5.1	120 - 135	33 - 35	27 - 29	36 - 38
	172 - 177	54 - 60	83 - 88	68 - 73	91 - 96
10(1)	5.7 - 5.9	125 - 145	35 - 37	30 - 32	38 - 40
	170 - 175	56 - 65	88 - 93	76 - 81	96 - 101
10(2)	5.9 - 5.11	130 - 140	35 - 37	30 - 32	38 - 40
	175 - 180	58 - 63	88 - 93	76 - 81	96 - 101
12(1)	5.8 - 5.1	135 - 150	37 - 39	32 - 34	40 - 42
	172 - 177	60 - 67	93 - 99	81 - 86	101 - 106
12(2)	5.6 - 5.8	130 - 140	37 - 39	32 - 34	40 - 42
	167 - 172	58 - 63	93 - 99	81 - 86	101 - 106
14(1)	5.9 - 5.11	145 - 155	39 - 41	34 - 36	42 - 44
	175 - 180	65 - 69	99 - 104	86 - 91	106 - 111
14(2)	5.7 - 5.9	140 - 150	39 - 41	34 - 36	42 - 44
	170 - 175	63 - 67	99 - 104	86 - 91	106 - 111

IV. HAAR CLASSIFIER

The core basis for Haar classifier object detection is the Haar-like features. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. Haar-like features, as shown in Figure 1 are used to detect an image [5]. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined. This allows features to be used to detect objects of various sizes.

A. Integral Image

The simple rectangular features of an image are calculated using an intermediate representation of an image, called the integral image [9]. The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (x, y) inclusive [11][12]. So if A[x,y] is the original image and AI[x,y] is the integral image then the integral image is computed as shown in equation 1 and illustrated in Figure 9.

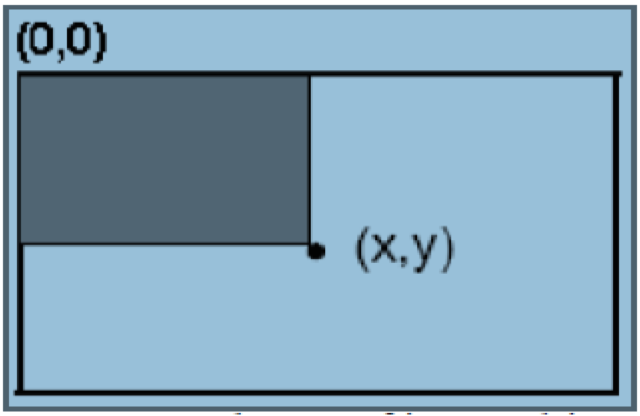

Fig.9. summoned are of integral image

B. Classifiers Cascaded

Although calculating a feature is extremely efficient and fast, calculating all 180,000 features contained within a 24 × 24 sub-image is impractical [Viola 2001, Wilson 2005]. Fortunately, only a tiny fraction of those features are needed to determine if a sub-image potentially contains the desired object [6]. In order to eliminate as many sub-images as possible, only a few of the features that define an object are used when analyzing sub-images. The goal is to eliminate a substantial amount, around 50%, of the sub-images that do not contain the object.

Haar classifiers continue, increasing the number of features used to analyze the sub-image at each stage. The cascading of the classifiers allows only the sub-images with the highest probability to be analyzed for all Haar-features that distinguish an object. It also allows one to vary the accuracy of a classifier. One can increase both the false alarm rate and positive hit rate by decreasing the number of stages.

The inverse of this is also true. Viola and Jones were able to achieve a 95% accuracy rate for the detection of a human face using only 200 simple features [9]. Using a 2 GHz computer, a Haar classifier cascade could detect human faces at a rate of at least five frames per second [5].

V. EXPERIMENTS

The result of the experiment is evaluated by the average error rate of 10-fold cross validation of each data set for 10 runs. 10-fold cross validation is a process which divides the data set into 10 blocks.9 blocks are merged for training data and the rest for the testing data. We test every data set for ten times and then calculate the average accuracy. Every time we randomly choose the training data and testing data.

TABLE 2:

STANDARD BODY PARAMETERS FOR MALE

Size	Height (In/Cm)	Weight (Lbs/Kgs)	Chest (In/Cm)	Waist (In/Cm)	Hips (In/Cm)
XS	5.4 - 5.7 16[illegible] - 169	120 - 145 54 - 66	33 - 36 84 - 91	27 - 29 68 - 73	32 - 35 81 - 88
S	5.6 - 5.9 167 - 175	135 - 155 61 - 70	34 - 37 86 - 94	28 - 31 70 - 78	34 - 37 86 - 94
M	5.7 - 5.11 170 - 180	150 - 175 68 - 80	37 - 39 94 - 99	29 - 33 73 - 83	35 - 39 89 - 99
M-L(1)	5.1 - 6 177 - 183	170 - 195 77 - 89	38 - 42 196 - 106	32 - 36 81 - 91	38 - 42 96 - 106
M-L(2)	6 - 6.2 183 - 188	175 - 200 80 - 91	38 - 42 96 - 106	32 - 36 81 - 91	38 - 42 96 - 106
M-L(3)	5.1 - 6 177 - 183	180 - 205 82 - 93	40 - 44 101 - 111	32 - 36 81 - 91	40 - 44 101 - 111
L	6 - 6.2 183 - 188	185 - 210 84 - 96	40 - 44 101 - 111	34 - 37 86 - 94	41 - 44 101 - 111
L-XL(1)	6 - 6.4 187 - 193	195 - 220 89 - 100	41 - 44 101 - 111	34 - 37 86 - 94	42 - 44 101 - 111
L-XL(2)	6 - 6.2 183 - 188	200 - 230 91 - 105	42 - 46 106 - 116	34 - 37 86 - 94	43 - 47 109 - 119
XL	6.2 - 6.4 187 - 193	210 - 240 96 - 109	42 - 46 106 - 116	37 - 40 94 - 101	43 - 47 109 - 119
XL-XXL	6.2 - 6.4 187 - 193	200 - 250 100 - 114	44 - 48 111 - 121	37 - 40 94 - 101	45 - 49 114 - 124
XXL	6.3 - 6.5 190 - 196	235 - 265 107 - 121	44 - 48 111 - 121	39 - 43 98 - 109	45 - 49 114 - 124

VI. CONCLUSION

After an introduction, the related work was presented; starting with cloth selection and virtual try-on, cloth recommendation system is also available. Subsequently a closer look on the technologies and frameworks that were used for the implementation, like Haar classifier algorithm, of the Tailoring Measurement and Virtual Try-on was taken. After this the different aspects of the design process up to the construction of the garment models were highlighted. This is followed by the implementation, for instance.

Size	Height (In/Cm)	Weight (Lbs/Kgs)	Chest (In/Cm)	Waist (In/Cm)	Hips (In/Cm)
XS	5.4 - 5.7 16[illegible] - 169	120 - 145 54 - 66	33 - 36 84 - 91	27 - 29 68 - 73	32 - 35 81 - 88
S	5.6 - 5.9 167 - 175	135 - 155 61 - 70	34 - 37 86 - 94	28 - 31 70 - 78	34 - 37 86 - 94
M	5.7 - 5.11 170 - 180	150 - 175 68 - 80	37 - 39 94 - 99	29 - 33 73 - 83	35 - 39 89 - 99
M-L(1)	5.1 - 6 177 - 183	170 - 195 77 - 89	38 - 42 196 - 106	32 - 36 81 - 91	38 - 42 96 - 106
M-L(2)	6 - 6.2 183 - 188	175 - 200 80 - 91	38 - 42 96 - 106	32 - 36 81 - 91	38 - 42 96 - 106
M-L(3)	5.1 - 6 177 - 183	180 - 205 82 - 93	40 - 44 101 - 111	32 - 36 81 - 91	40 - 44 101 - 111
L	6 - 6.2 183 - 188	185 - 210 84 - 96	40 - 44 101 - 111	34 - 37 86 - 94	41 - 44 101 - 111
L-XL(1)	6 - 6.4 187 - 193	195 - 220 89 - 100	41 - 44 101 - 111	34 - 37 86 - 94	42 - 44 101 - 111
L-XL(2)	6 - 6.2 183 - 188	200 - 230 91 - 105	42 - 46 106 - 116	34 - 37 86 - 94	43 - 47 109 - 119
XL	6.2 - 6.4 187 - 193	210 - 240 96 - 109	42 - 46 106 - 116	37 - 40 94 - 101	43 - 47 109 - 119
XL-XXL	6.2 - 6.4 187 - 193	200 - 250 100 - 114	44 - 48 111 - 121	37 - 40 94 - 101	45 - 49 114 - 124
XXL	6.3 - 6.5 190 - 196	235 - 265 107 - 121	44 - 48 111 - 121	39 - 43 98 - 109	45 - 49 114 - 124

e g d a n y o e a

	MALE	FEMALE
Test 1 accuracy	80.13%	86.13%
Test 2 accuracy	79.27%	70.22%
Test 3 accuracy	74.12%	76.23%
Test 4 accuracy	72.22%	63.87%
Test 5 accuracy	77.93%	73.46%
Test 6 accuracy	66.71%	72.33%
Test 7 accuracy	71.53%	71.53%
Test 8 accuracy	88.30%	78.61%
Test 9 accuracy	77.43%	77.12%
Test 10 accuracy	70.02%	80.59%
TOTAL	74.56%	75.08%

SHOWS THE ACCURACY of CLASSIFYING MALE DATA SET AND FEMALE DATA SET.

Beyond that a simple setup of the system can also be assembled at home since the minimum requirements are a computer with a screen and a Camera. This can also result in an additional feature for a web shop, for instance. It would allow a virtual try-on of clothes before people are buying it online, taking a closer look at the garment and even conveying the actual behavior of the real cloth. This demonstrates a huge advantage over the common web shopping experience.

VII. FUTURE WORK

In this paper, we have developed a methodology, to put on some clothes on the image from our database in 2D module. This is just a small step toward the Virtual Fitting application. Here we have, front image for each dress which is superimposed on the user and the 2D graphics of the product seem to be relatively satisfactory and practical for many uses such as jewelry, glasses, hair style, fitness, and gaming.

There are many possible implementations regarding the model used for fitting. It is possible to apply a homographic transformation to the images rather than the simple scale-rotate technique in order to match multiple joints altogether although it would require more computation.

Another alternative could be using many pictures at different angles so that it would be possible to create more realistic video streams. One could achieve a similar effect using 3D models and rendering them according to the current angle and positions. Second approach would also make it possible to implement a physics engine to go along with the model.

The 2D patterns can be generated from the personally sized garments or by using the generic body measurements as shown in table 1 and table 2. These 2D measurements could be directly sent to the cloth manufactures. The speed optimization for on-line calculation comes from wide use of generic database of bodies and garments [3].

REFERENCES

[1] S. M. Metev and V. P. Veiko, "Laser Assisted Micro technology," 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.
[2] J. Breckling, Ed., "The Analysis of Directional Time Series: Applications to Wind Speed and Direction," ser. Lecture Notes in Statistics. Berlin, Germany: Springer, , vol. 6, 1989.
[3] D. Protopsaltou, C. Luible, Marlene Arevalo, Nadia Magnenat-Thalmann "A body and garment creation method for an Internet based virtual fitting room," MIRALab CUI, University of Geneva, Switzerland 2006.
[4] M. Fukuda, Y. Nakatani "Clothes Recommend Themselves: A New Approach to a Fashion Coordinate Support System," Proceedings of the World Congress on Engineering and Computer Science 2011 Vol I WCECS 2011, October 19-21, 2011, San Francisco, USA.
[5] P. Ian Wilson and Dr. J. Fernandez "Facial feature detection using Haar classifiers," JCSC 21, 4 (April 2006)
[6] F. Isikdogan and G. Kara "A Real Time Virtual Dressing Room Application using Kinect," Bogazi ci University, Istanbul, Turkey 2012.
[7] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-Time Human Pose Recognition in Parts from Single Depth Images," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,2011.
[8] F. Cordier, W. Lee, H. Seo, N. Magnenat-Thalmann "Virtual-Try-On on the Web," in proceedings of International Conference on Virtual Reality, Laval Virtual, University of Geneva 2, May 16-18, MIRALab, 2001.
[9] K. Kjærside, K.J. Kortbek, H. Hedegaard, "ARDressCode: Augmented Dressing Room with Tag-based Motion Tracking and Real-Time Clothes Simulation," in Proceedings of the Central European Multimedia and Virtual Reality Conference, 2005.
[10] Philipp Presle "A Virtual Dressing Room based on Depth Data," Vienna University of Technology, Klosterneuburg.
[11] K. Onishi, T. Takiguchi, and Y. Ariki, "3D Human Posture Estimation using the HOG Features from Monocular Image," in proceedings of 19th International Conference on Pattern Recognition, 2008.
[12] D. Chai, and K. N. Ngan, "Face Segmentation using Skin-Color Map in Videophone Applications," IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 4, June 1999.
[13] J. Young Choi, Y. Man Ro and Konstantinos N. Plataniotis," Color Local Texture Features for Color Face Recognition" in proceedings of IEEE Conference, 2011.
[14] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face recognition algorithms," IEEE Trans Pattern Anal. Mach. Intell., vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[15] Euratex (2000), Bulletin 2000/5, "The European Textile/Clothing Industry on the eve of the New Millennium," Brussels.
[16] Blanz V., Vetter T. (1999)," A morphable model for the synthesis of 3D faces, in Computer Graphics," (Proc.SIGGRAPH'99, Los Angeles California, USA), ACM Press New York, pp. 187-194.3.
[17] Volino P., Magnenat-Thalmann N. (2000) "Virtual Clothing-Theory and Practice", Springer, Berlin Heidelberg.
[18] Carter J.E.I, Heath, B.H. (1990), "Somatotyping–development and applications," Cambridge University Press, Cambridge.
[19] Liechty E.G., Liechty E.L., Pottberg D.N., Judith A. (1992)," Fitting and Pattern Alteration: A Multi-Method Approach, Fairchild Publications," Chicago.
[20] Roebuck, J.A., Jr. (1993), "Anthropometric Methods: Designing to fit the Human Body," Monographs in Human Factors and Ergonomics, Human Factors and Ergonomics Society, Santa Monica.
[21] Cootes T.F., Taylor C.J., (1998), "Active Shape Models, Work Report" Department of Medical Biophysiscs, University of Manchester, August 1998.