Automatic Detection of 2D Human Postures Based on Single Images

Humberto Souto Júnior, Soraia Raupp Musse
Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS
Graduate Programme in Computer Science
Virtual Human Laboratory -www.inf.pucrs.br/∼vhlab
Porto Alegre, Brazil
E-mail:soraia.musse@pucrs.br

Figure 1. Detection of human postures in single images.

***Abstract*—Estimating human pose in static images is a challenging task due to the high dimensional state space, presence of image clutter and ambiguities of image observations. In this paper we propose a method to automatically detect human poses in a single image, based on a 2D model combined with anthropometric data. Furthermore, we use artificial neural networks to detect high level information about the human posture. Experimental results showed that the proposed technique performs well in non trivial images.**

***Keywords*-posture detection; single image; artificial neural network.**

I. Introduction

Estimating human pose in static images is challenging due to the influence of numerous real-world factors such as shading, image noise, occlusions, background clutter and the inherent loss of depth information when a scene is captured onto a 2D image [1]. Human pose estimation can be made more robust by integrating the detection of body components such as face and limbs, with the highly constrained structure of the articulated body [2]. As related by Agarwal and Triggs [3], there are two main schools of thought on this. Model-based approaches presuppose an explicitly known parametric body model, and estimate the pose either by directly inverting the kinematics or by numerically optimizing some form of model-image correspondence metric over the pose variables. In contrast, learning based approaches try to avoid the need for explicit initialization and accurate 3D modeling and rendering, and to capitalize on the fact that the set of typical human poses is far smaller than the set of kinematically possible ones, by estimating (learning) a model that directly recovers pose estimates from observable image quantities.

In this paper we propose a learining based method to automatically detect human postures in single images, based on Artificial Neural Networks (ANNs) and heuristics. Requirements for our model are the segmented image (that can be obtained using an interactive software (e.g. based on graph or grab cuts [4], [5]) or fully automatic approaches (e.g. [6], [7],). We propose the classification of HLPs (High Level Posturs), which aims to characterize high level information about the posture of person in the image.

The next sections present details about our method and obtained results. In Section II we describe competitive approaches intending to contextualize our work in the state-of-the-art. Section III presents details of our method, while Section IV presents results using manually and automatically segmented images as input data. Also, we discuss some limitations and contributions of our work.

II. Related Work

In [8], the goals were to detect a human figure, and localize his/her joints and limbs, along with their associated pixel masks. For that purpose, authors start with a low-level process that consists of Canny edge detection followed by a boundary finder, and also the Normalized Cuts segmentation algorithm. The segmentation is used to obtain the torso and head, while shape, shading and focus cues are used to isolate the limbs from the background. Finally, all body parts are put together based on global anthropometric and kinematic constraints.

The problem tackled in [9] was to locate the positions of the joints of a person in a still image, and to use this information to estimate the configuration and posture of the body in 3D. The basic approach was to store a number of examples of 2D images of human bodies in different configurations and camera viewpoints. In each of these stored images, the location of joints (elbows, knees, etc.) was manually marked and labeled to be used in future. The input image is then tested with each of the samples stored, and the most appropriate example is found.

The work of Lee and Cohen [10] proposes the use of an adaptive approach, where a human model is used to synthesize the image regions corresponding to human forms (given a hypothesis) and thus separating the person from the background. For hypotheses generation, the MCMC model is used.

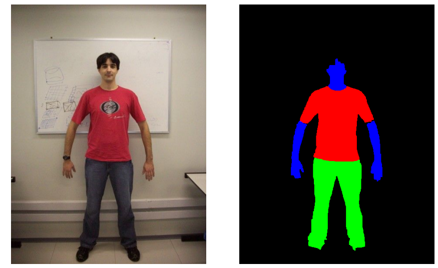
McIntosh et al. [11] propose a technique for self-initializing a kinematic tracker that automatically discovers its part appearance models from a video sequence. Through its unique combination of an existing global joint estimation technique and a robust physical deformation based local search method, the tracker is demonstrated as a novel approach to recover 2D human joint locations and limb outlines from video sequences. Appearance models are discovered and employed through a novel use of the deformable organisms framework which has been extended to the temporal domain. Quantitative and qualitative results for a set of five test videos are provided, and the results demonstrate an overall improvement in tracking performance and that the method is relatively insensitive to initialization, an important consideration in gradient descent-style search algorithms.

Another work on human skeleton estimation in photographs is provided in [12], aiming to recover poses for the upper body only. In this work, the face, skin and torso are initially detected, and then the joints are properly initialized according with the observations and some heuristic restriction configuration. Finally, the Markov chain Monte Carlo (MCMC) method is employed to determine the final pose.

In this paper we propose a method to detect human posture in single images using neural networks. Our method can be viewed as a pipeline formed by 4 stages, as detailed in next sections.

III. The Model for Human Posture Detection

The process starts with a segmented image (illustrated in Figure 2(b)), where blue pixels represent skin pixels, red pixels are associated with the T-shirt (defined as the upper parts of the body that are not defined as skin colors), and green ones are associated with pants or shorts (a similar definition is used for the lower part of the body). This is the way we need to receive the segmented image that can be resulting of manual or automatic process of segmentation.



(a) (b)

Figure 2. (a) Original picture; (b) Segmented image obtained using a manual method for segmentation [4], [5].

The overall architecture is illustrated in Figure 3. The model is organized into 4 phases (see Figure 3). The first phase is responsible for generating the projection curves, which are computed based on the segmented image. The information obtained in this phase is important for other phases in the pipeline, as the ANNs and the method proposed to identity body parts (phases 2 and 3, respectively).

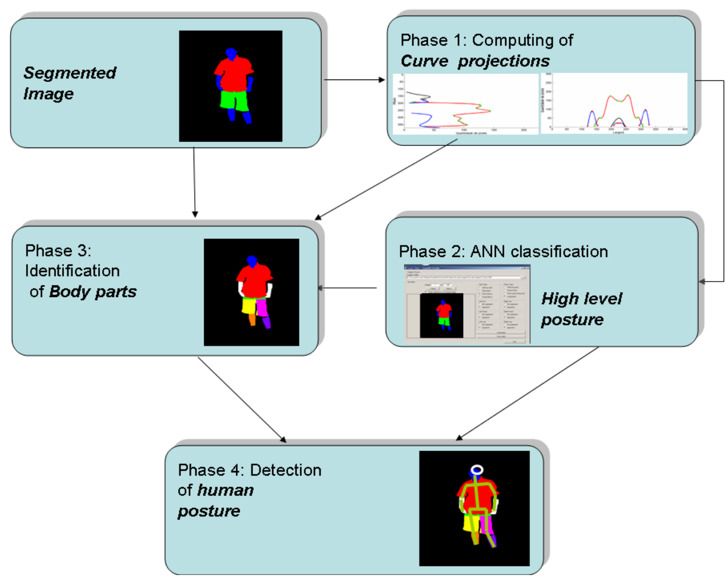


Figure 3. Overall architecture of our method.

Phase 2 is responsible for generating high level information of postures (HLP). In this paper, we propose that HLPs are related with global data related to the human body in the picture. The main motivation is to provide information for next phases (3 and 4) in the pipeline, which are more co-related with heuristics, based on curves (from phase 1) and HLP (from phase 2). Further details about the HLP are described in next sections.

In phase 3, the pixels in the image are identified as specific parts of the human body. Head, right and left arms as well as right and left legs should be identified (if they are visible in the picture, as informed by the HLP). Finally, the identified body parts as well as the HLP are used as input data to detect the human posture (phase 4), locating visible joints and bones.

The next sections describe each phase in details.

A. Phase 1: Computing Projection Curves

The main goal of this phase is to generate projection curves of pixels based on a segmented image. At the beginning, the image is subdivided in two parts in order to separate the upper part of the body from the lower one. This process is important to train the ANNs with such 2 subdivisions (we also tried using the hole image, but classification results obtained with the ANNs were better when specific parts of the body were used in training phase, by subdividing the image in upper and lower part of the bodies). The subdivision process firstly estimates the height of lower part of the body (pants or shorts) based on green pixels in the segmented image. If the blob containing green pixels is significantly smaller than expected (using anthropometry [13]) and based on the face radius, there are possibly more skin pixels in the lower part of the body, or such part is occluded. In this case, we use anthropometry to estimate the probable height of lower body part. Using such data, the image can be subdivided horizontally in two parts.

For each part, we compute vertical and horizontal projection curves. These curves are generated based on equations 1 and 2. Let $I_c(x, y)$ be the color component of the image in a specific channel (R, G, B). Equation 1 describes the generation of the vertical projection curve, while 2 refers to the horizontal projection curve.

$$P_x x = \sum_{\boldsymbol{x}} I_c(x, y), \tag{1}$$

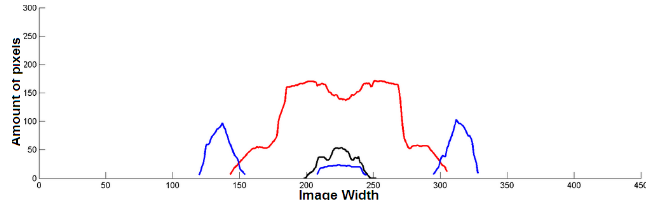
$$P_y y = \sum_{\boldsymbol{y}} I_c(x, y). \tag{2}$$

The vertical projection of upper part of the body is illustrated in Figure 4(a), and in Figure 4(b) we show the fitted polynomial curves. The polynomial degree depends on the adjusted curve, and they were adjusted experimentally: for the upper part of the body we used degree 15, and for the remaining parts, degree 5 showed to be enough to represent the curve. After finding the mathematical description of the curves, we are able to find the minimum and maximum peaks. Figure 5 shows the horizontal and vertical projection curves obtained for the image presented in Figure 2(b).

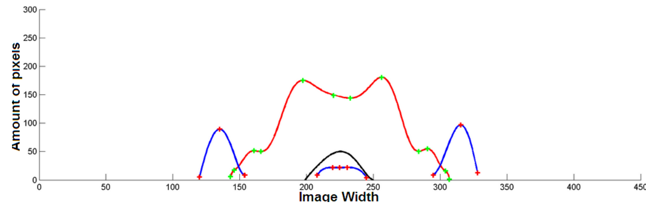
The output of this phase are the projection curves, as well as minimum and maximum peaks. Such data is passed on to phase 2, where the High Level Postures (HLPs) are obtained using ANNs.

B. Phase 2: Classification using ANNs

In order to determine some global characteristics of human postures in images, we propose the usage of ANNs to

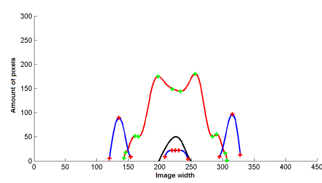


(a)

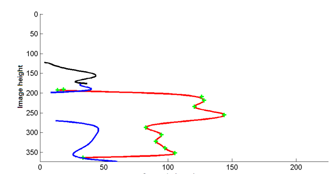


(b)

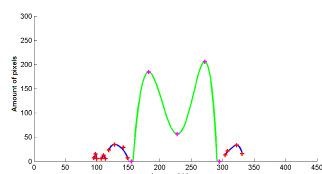
Figure 4. Phase 1: performs the projection curves in each channel (R, G, B). Black, red and blue curves correspond to head, upper part of the body and skin pixels presented in the segmented image. (a) the curves are not fitted yet; (b) image shows the curve after processing the peaks.



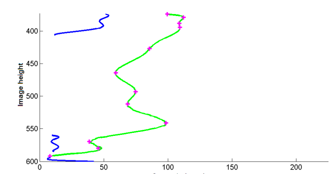
(a)



(b)



(c)



(d)

Figure 5. Generated curves based on image 2. The black, red, green and blue curves correspond respectively to head and projection of channels (R, G, B). (a) Vertical projection - upper part of the body; (b) Horizontal projection - upper part of the body; (c) Vertical projection - lower part of the body; (d) Horizontal projection - lower part of the body.

generate HLPs, as previously mentioned. The input of our ANNs are the result of equations 1 and 2, i.e. the sum of pixels for each channel. In addition, peaks of each curve (for each channel) can also be used in phase 4, in order to locate the skeleton joints. The ANNs should classify the following characteristics in the images, we called HLP criteria:

- **Type of T-shirt?** 0- without T-shirt, 1- without sleeves, 2- short sleeves, 3- long sleeves.
- **Right arm is visible?** Yes/No
- **Left arm is visible?** Yes/No
- **Right hand is visible?** Yes/No
- **Left hand is visible?** Yes/No

- **Type of pants?** 0- without pants, 1- short pants and/or short skirt, 2- medium pants and/or medium skirt, 3- long pants and/or long skirt.
- **Right leg is visible?** Yes/No
- **Left leg is visible?** Yes/No

One of the most used neural learning methods is the Multi Layer Perceptron (MLP), which employs the back-propagation learning algorithm [14], [15]. In this work, we used the method proposed by Levenberg-Marquardt [16] and [17]. The ANN method is proposed to be solved into two stages. Firstly, a vector formed by result of Equations 1 and 2 is defined, as illustrated in the following algorithm (1). Then, this vector is used for training and classification using ANNs.

Algorithm 1 Algorithm used for generation of input layer in neural network

repeat
  For each image
  1. Subdivide image in upper and lower part of the body
  2. Compute Vertical Projection Curves:
  2.1 Upper part of the body in channel R
  2.2 Upper part of the body in channel B
  2.3 Lower part of the body in channel B
  2.4 Lower part of the body in channel G
  3. Compute Horizontal Projection Curves:
  3.1 Upper part of the body in channel R
  3.2 Upper part of the body in channel B
  3.3 Lower part of the body in channel B
  3.4 Lower part of the body in channel G
  4. Generate a vector concatenating 8 vectors and normalizing them for having a standard resolution. We used images with resolution of 600x800 pixels.
until End of images to be used in training phase

We have chosen to use a single neural network for each criteria to be classified in the HLP. Our tests describe better performances when all projection curves were informed to all 8 criteria (each one determined in one ANN). For all neural networks we used vectors explained in the algorithm in the input layer and 10 hidden neurons (which presented better results in our case). The output is dependent of number of answers for each network. For instance, visible right hand requires only one output, referring to YES or NO.

In order to learn a specific task, a database containing examples including input patterns and expected answers is required. The training phase is performed as illustrated in Figure 6. Firstly, we have a database containing 164 images which are subdivided in training (85%) and testing (15%) groups. Each one of the 139 images (training set) were manually classified by a specialist based on a supervised algorithm, where 8 HLP criteria are informed. Each image is then scaled 10 times and rotated vertically, generating a training set of 2780 images and their included HLP. Finally, these 2780 classified images are used to train the ANNs.

Once the training phase is performed, no more database or training process are required. In the classification procedure, the neural network generates the HLP for test images.

The next section discusses phase 3, responsible for the identification of body parts in the image, using heuristics and HLP classification for each test image.

C. Phase 3: Identification of Body Parts

This phase is responsible for identifying the body parts of a segmented person in the image. It uses the information processed in last phases:

- segmented image
- Projection curves
- HLP classified as a function of ANN´s

The identification of body parts is also performed using heuristics and anthropometry, and it is organized into the two following steps:

- Step 1: Face detection is performed in the original image. Pixels inside the face region are firstly identified as part of the face.
- Step 2: Depending on the HLP, expected areas for blobs related to different body parts are defined. For instance, if it is defined in the HLP that the right arm is visible, but the person has short sleeves, then skin pixels can have the expected area of approximately half arm (based on anthropometry). If the right arm is not visible, this specific body part is not searched in the image.
- Step 3: After having the expected size and channels of each body part (T-shirt in red, skin pixels in blue and pants in green), the algorithm tries to detect the most probable projection curve that should represent each body part. However, this process includes a special treatment for visible arms and legs, as detailed next.

In our work, the skeleton model is composed by nineteen bones and twenty joints, as illustrated in Figure 7. All these bones have initial 3D lengths and widths, both parameterized as a function of the height h of an average person based on anthropometric values [13]. More precisely, for a certain body part with label i, the corresponding length l_i and width w_i are given by

$$l_i = hf_{li}, \quad w_i = hf_{wi}, \tag{3}$$

where the proportionality factors f_{li} and f_{wi} are derived from [13]. Table I presents all body parts used in this work, along with the corresponding values for f_{li} and f_{wi}.

The main idea here is to identify the projection curves as a specific part of the body. The legs can be represented using only one projection curve or more. In order to help the identification of body parts, we compute the torso direction, which is a vector formed by the center of the detected face

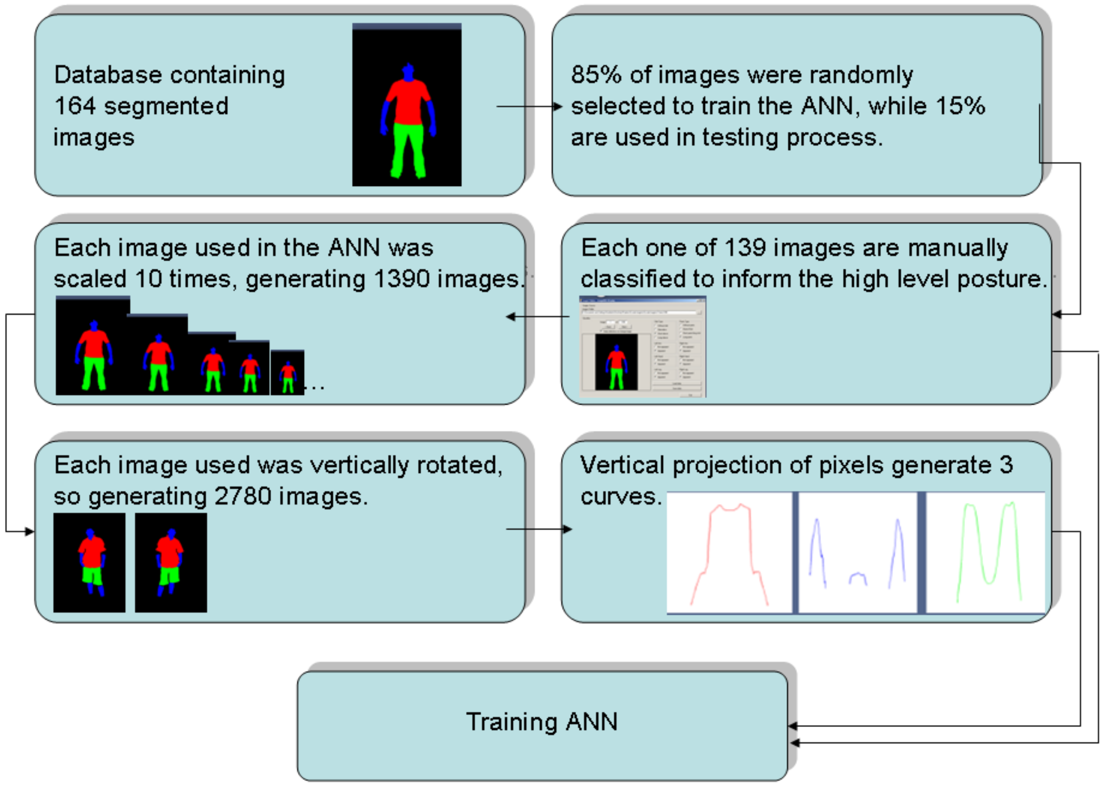


Figure 6. Training process of the ANNs, which includes supervised classification using projection curves.

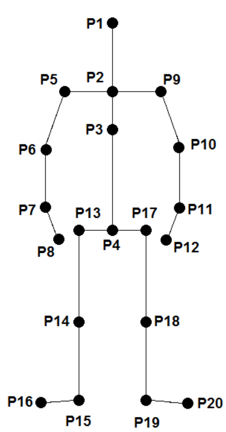


Figure 7. The adopted skeleton model.

and the division point between two legs. Figure 5(c) shows a projection curve (green) in which there is only one curve for the both legs, and the separation is clearly visible in the minimum peak. On the other hand, Figure 8(a) shows separated projection curves. In this case, we compute the mean point between two or more curves and connect it with the center of the face, resulting in the torso direction vector. This procedure allows the detection of both legs and avoid problems occurred as a function of people orientation in the picture, as illustrated in Figures 8(b) and 8(c). Since we know the probable curve for each leg, the HLP is used in order to inform if the leg should have more blue pixels (skin) or green ones (pants).

After having identified the face, the legs and pants (based on HLP), the algorithm will detect the arms and hands. T-shirts could be detected based on the HLP and color pixels, as performed for legs. The challenge with arms and hands happens when they are visible but connected in some way. In this case, the area of the curve should be large enough in order to contain both arms/hands. This process can present some misclassification, as discussed briefly in Section IV, but the number of correct detections can argue in favor of simple heuristics and interactive frame-rates.

Figure 9 illustrates a correct classification result using our algorithm. Each leg is detected and visualized in a different color (orange and purple). The two arms are visualized in white, each leg of the pants is visualized in a different color (yellow and pink), while pixels related to the T-shirt are shown in red.

Table I
IN THE FIRST COLUMN: THE BODY PART INDEX; IN THE SECOND COLUMN: THE BODY PART (BONE); IN THE THIRD COLUMN: THE TWO JOINTS THAT FORM EACH BONE; IN THE FOURTH COLUMN: THE WEIGHTS USED TO COMPUTE EACH BONE LENGTH; AND IN THE FIFTH COLUMN: THE WEIGHTS USED TO COMPUTE EACH BONE WIDTH.

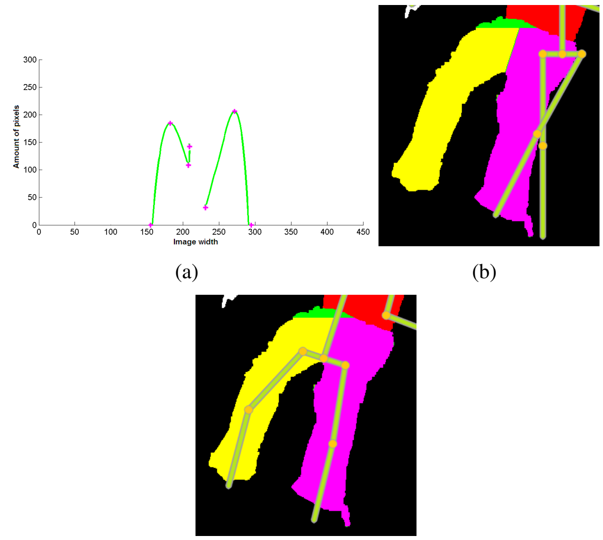
i	Bone	Joints	f_{li}	f_{wi}
0	Head	(P1 - P2)	0.20	0.0883
1	Chest	(P2 - P3)	0.098	0.1751
2	Abdomen	(P3 - P4)	0.172	0.1751
3	Right Shoulder	(P2 - P5)	0.102	not used
4	Right Arm	(P5 - P6)	0.159	0.0608
5	Right Foreman	(P6 - P7)	0.146	0.0492
6	Right Hand	(P7 - P8)	0.108	0.0593
7	Left Shoulder	(P2 - P9)	0.102	not used
8	Left Arm	(P9 - P10)	0.159	0.0608
9	Left Foreman	(P10 - P11)	0.146	0.0492
10	Left Hand	(P11 - P12)	0.108	0.0593
11	Right Hip	(P4 - P13)	0.050	not used
12	Right Thigh	(P13 - P14)	0.241	0.0912
13	Right Calf	(P14 - P15)	0.240	0.0608
14	Right Feet	(P15 - P16)	0.123	0.0564
15	LeftHip	(P4 - P17)	0.050	not used
16	LeftThigh	(P17 - P18)	0.241	0.0912
17	LeftCalf	(P18 - P19)	0.240	0.0608
18	LeftFeet	(P19 - P20)	0.123	0.0564

D. Phase 4: Detection of Human Posture

This phase starts when the projection curves and blobs of pixels are already identified. The goal here is to generate a skeleton automatically, i.e. to estimate the position of joints in the human body. The skeleton starts in the center of the face, and another useful information to help in the skeleton detection is the torso orientation vector. The pipeline for skeleton detection after face and torso orientation detection is organized as follows:

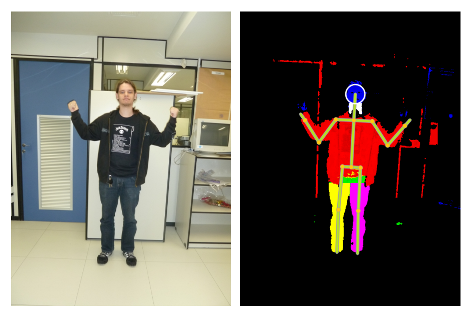
- Neck, torso and hips are easily estimated based on the expected size of body parts and torso orientation.
- Clavicle detection is based on neck detection. A simple rule is to put clavicles perpendicular to the torso, and then use validation procedure in order to localize joints into the upper part of the body (avoiding clavicles in the boundary of upper part of the body). To do this, we used a distance threshold from the boundary to the clavicles to keep the joints inside the upper part of the body, as illustrated in Figure 10(a).
- Arms detection is based on anthropometry. Figure 10(b) illustrates the bone estimation (starting in the clavicles) and elbows, and then the fist. The process includes rotation of estimated size of the bones and an error measured for each elbow. Since, the HLP informs if arms have skin pixels or long sleeves, the heuristics are able to estimate the position of joints.
- Legs detection is very similar to the arms detection. The only difference is concerned with the range of angles used for searching the joints, as illustrated in Figure 10(c) and (d).

At the end of this process, joints are estimated and skeleton is defined, as illustrated in Figure 9.



(a) (b) (c)

Figure 8. (a) Two Curves projection representing both legs. (b) Wrong skeleton detection. (c) Corrected skeleton based on torso orientation.

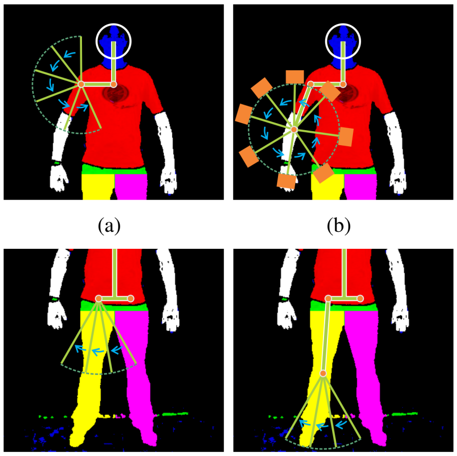


(a) (b)

Figure 9. (a) Original image. (b) Automatic segmentation and skeleton detection.

IV. RESULTS

This section presents results obtained using our method. Firstly, we discuss some results obtained using manual segmentation and then, other results using an automatic segmentation technique [18]. In this recent work, a method to automatically segment people in images was proposed, and then we estimate the human pose in static images. The



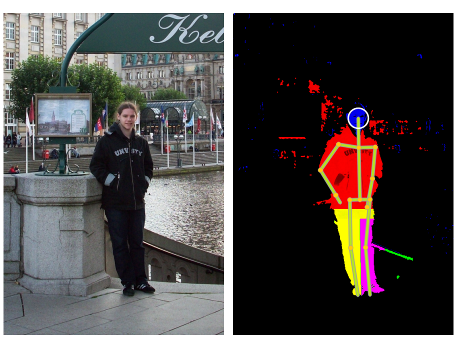
(a) (b)

(c) (d)

Figure 10. Phases of skeleton detection.

whole process is based on dominant colors, and given the face captured by a face detector. The posture is estimated using a 2D model combined with anthropometric data. Experimental results showed that this technique works well in frontal views images and for the upper part of the body. The current model improves the method proposing the ANN-based classification, where higher level information about the posture is provided. One advantage of our method is the possibility of finding events based on the HLP, for instance, finding in a dataset pictures where people have long T-shirts.

Figures 11, 12 and 13 illustrate 2 good classifications and a misclassification obtained in our work. The segmentation was performed using an automatic method for people segmentation in single images [18]. In the 2 correct cases, the HLPs were correctly classified. However, in Figure 13, generated HLP classified as not visible the right arm, and also phase 3 failed to identify the body parts.



(a) (b)

Figure 11. (a) Original image. (b) Results of human posture detection correctly classified using an automatic algorithm for segmentation.

After evaluating the test images (25 pictures containing automatically segmented pictures) we compute the following mean errors (answer incorrect according to specialists) for

(a) (b)

Figure 12. (a) Original image. (b) Results of human posture detection correctly classified using an automatic algorithm for segmentation.

(a) (b)

Figure 13. (a) Original image. (b) Results of human posture detection misclassified due to the segmentation problem and ANN classification.

each HLP criteria:

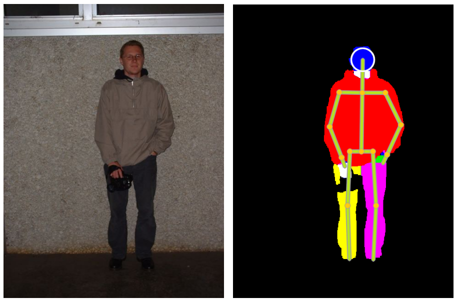
- **Type of t-shirt**$= 30.7\%$
- **Right arm is visible**$= 3.85\%$
- **Left arm is visible**$= 0\%$
- **Right hand is visible**$= 7.69\%$
- **Left hand is visible?**$= 30.77\%$
- **Type of pants?**$= 7.69\%$
- **Right leg is visible?**$= 3.85\%$
- **Left leg is visible?**$= 0\%$

Figures 14, 15 and 16 illustrate 2 good classifications and a misclassification obtained in our work. The segmentation was performed using an interactive approach (based on grab cuts [4]) for people segmentation in single images. In the 2 correct cases, the HLPs were correctly classified, but the generated HLP in Figure 16 contains errors concerning hands detection (informed YES when correct answer was NOT) and detected long sleeves in T-shirt type. Consequently, phase 4 detected the skeleton in wrong position.

After evaluating the test images (25 pictures containing only manually segmented pictures) we compute the following mean errors (answer incorrect if compared to specialists' answers) for each HLP criteria:

- **Type of t-shirt**$= 4\%$
- **Right arm is visible**$= 12\%$
- **Left arm is visible**$= 4\%$
- **Right hand is visible**$= 36\%$
- **Left hand is visible?**$= 24\%$
- **Type of pants?**$= 0\%$
- **Right leg is visible?**$= 0\%$
- **Left leg is visible?**$= 4\%$

As it can be seen in discussed results, the correct seg-



(a) (b)

Figure 14. a) Original image. (b) Results of human posture detection correctly classified using a manual algorithm for segmentation.

(a) (b)

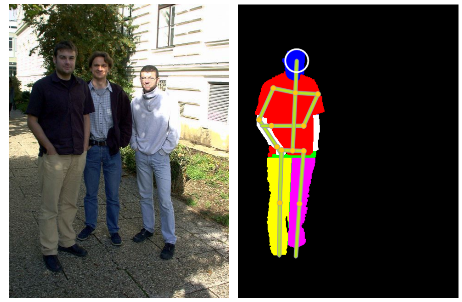
Figure 15. a) Original image. (b) Results of human posture detection correctly classified using a manual algorithm for segmentation.

mentation can impact the quality of classification performed by the ANN. When segmentation is manually produced, the errors were more focused on the hands, which generally do not impact the final posture. Yet, recognizing the presence of hands is not trivial, since people have hands in their pockets in most pictures. It is also important to mention that our pipeline produces results in interactive frame-rates, once the the model has been trained.

In addition, we performed an evaluation with a public dataset [19]. We manually segmented the 100 images, but it is important to mention that the face detector worked well in a good part of them, approximately in 80% of dataset. Then, we applied the proposed approach, which produced a coherent human pose in 75% of the images based on a visual inspection. In most of wrong cases, the ANN method detected wrong classification for type of paints and visibility of legs or arms and hands. Consequently, the obtained postures were wrong.

V. Conclusions

This paper presented a novel method for detecting human posture in a single images. The main contributions are the use of ANNs to classify the HLP in the picture, as well as the generation of human skeleton without user intervention and only based in a single image. Our method can present some classification errors, but we were able to achieve



(a) (b)

Figure 16. a) Original image. (b) Results of human posture detection misclassified due to the segmentation problem and ANN classification.

a maximum error rate of 16% in a training database of almost 100 images. When we use the public dataset [19], the achieved error rate was 25%. The wrong detections happen mainly when the ANN misclassified the HLP. However, more testsare necessary in order to evaluate wach phase of our pipeline. Future work is concerned with more tests and detection of 3D posture, based on currently available methods.

References

[1] Y. S. H. D. Cheng, X. H. Jiang and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 2259–2281, 2001.

[2] X. L. H. Y. Zhilan Hu, Guijin Wang, "Recovery of upper body poses in static images based on joints detection," *Pattern Recognition Letters*, vol. 30, pp. 503–512, 2009.

[3] A. Agarwal and B. Triggs, "A local basis representation for estimating human pose from cluttered images," in *Asian Conference on Computer Vision*, January 2006, pp. 50–59.

[4] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, pp. 309–314, August 2004. [Online]. Available: http://doi.acm.org/10.1145/1015706.1015720

[5] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images," *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, pp. 105–112 vol.1, 2001.

[6] P. Guan, A. Weiss, A. Balan, and M. Black, "Estimating human shape and pose from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2010, pp. 1381–1388.

[7] J.-M. Hasenfratz, M. Lapierre, J.-D. Gascuel, and E. Boyer, "Real-time capture, reconstruction and insertion into virtual world of human actors," in *Vision, Video and Graphics*, Eurographics. Elsevier, 2003, pp. 49–56. [Online]. Available: http://artis.imag.fr/Publications/2003/HLGB03

[8] G. Mori, X. Ren, A. A. Efros, and J. Malik, "Recovering human body configurations: combining segmentation and recognition," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, 2004, pp. II–326–II–333 Vol.2. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2004.1315182

[9] G. Mori, "Recovering 3d human body configurations using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1052–1062, 2006, senior Member-Malik, Jitendra.

[10] M. W. Lee and I. Cohen, "A model-based approach for estimating human 3d poses in static images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 6, pp. 905–916, 2006, student Member-Lee, Mun Wai and Member-Cohen, Isaac.

[11] C. McIntosh, G. Hamarneh, and G. Mori, "Human limb delineation and joint position recovery using localized boundary models," in *WMVC '07: Proceedings of the IEEE Workshop on Motion and Video Computing*. Washington, DC, USA: IEEE Computer Society, 2007, p. 31.

[12] Z. Hu, G. Wang, X. Lin, and H. Yan, "Recovery of upper body poses in static images based on joints detection," *Pattern Recogn. Lett.*, vol. 30, no. 5, pp. 503–512, 2009.

[13] A. R. Tilley, *The measure of man and woman - Human factors in design*. John Wiley & Sons, inc, 2002.

[14] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1998.

[15] ——, *Neural Networks and Learning Machines*, 3rd ed. Prentice Hall, November 2008.

[16] J. yan Fan and Y. xiang Yuan, "On the convergence of a new levenberg-marquardt method," in *Technical Report, AMSS, Chinese Academy of Sciences*, 2001.

[17] L. wan Chan, "Levenberg-marquardt learning and regularization," in *in Progress in Neural Information Processing*. Springer Verlag, 1996, pp. 139–144.

[18] J. C. Jacques, L. Dihl, C. Jung, M. Thielo, R. Keshet, and S. Musse, "Human upper body identification from images," in *ICIP*. IEEE, 2010, pp. 1717–1720.

[19] B. Yao, X. Yang, and S. C. Zhu, "Introduction to a large-scale general purpose ground truth database: Methodology, annotation tool and benchmarks," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2007, pp. 169–183.