

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Human classification using gait features

### **This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/153211> since 2015-12-11T15:13:23Z

*Publisher:*

Springer Verlag

*Published version:*

DOI:10.1007/978-3-319-13386-7\_2

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



## UNIVERSITÀ DEGLI STUDI DI TORINO

**This is an author version of the contribution published on:**

Elena. Gianaria, Marco Grangetto, Maurizio Lucenteforte and  
Nello Balossino. Human classification using gait features. *1st  
International Workshop on Biometrics, BIOMET 2014*, 8897, ISBN  
9783319133850, 2014, 10.1007/978-3-319-13386-7\_2.

**The definitive version is available at:**

**http:**  
[//link.springer.com/chapter/10.1007%2F978-3-319-13386-7\\_2](http://link.springer.com/chapter/10.1007%2F978-3-319-13386-7_2)

# Human Classification Using Gait Features

Elena Gianaria, Marco Grangetto, Maurizio Lucenteforte, and Nello Balossino

Dipartimento di Informatica, Università degli Studi di Torino  
Corso Svizzera 185, 10149 Torino, Italy  
first.lastname@unito.it

**Abstract.** Gait exhibits several advantages with respect to other biometrics features: acquisition can be performed through cheap technology, at a distance and without people collaboration. In this paper we perform gait analysis using skeletal data provided by the Microsoft Kinect sensor. We defined a rich set of physical and behavioral features aiming at identifying the more relevant parameters for gait description. Using SVM we showed that a limited set of behavioral features related to the movements of head, elbows and knees is a very effective tool for gait characterization and people recognition. In particular, our experimental results shows that it is possible to achieve 96% classification accuracy when discriminating a group of 20 people.

**Keywords:** Gait characterization, Gait analysis, Kinect, Support Vector Machine

## 1 Introduction and Related Work

Biometrics is the science that studies the human characteristics for anthropometry research, people identification, access control and many more. Biometric features are measurable data classified as physical or behavioral [1]. The former are related to the body and its shape. Some examples are face, hand, iris, retina and fingerprint. Behavioral characteristics are associated to particular human action, for instance handwriting and walking. Automatic recognition systems are often expensive, intrusive and require the cooperation of the subject during the acquisition. The latter cannot be always guaranteed, for instance in a video surveillance context. In this case, it is useful to recognize people through biometric parameters that can be captured at a distance and without the collaboration of the person, such as gait [2].

Gait analysis finds interest in video surveillance systems [3, 4] and forensics science [5, 6]. Furthermore, many applications analyze the gait in order to discover pathologies of the body movement [7], rehabilitation therapy [8], identify the fall risk in elderly population in order to assess the frailty syndrome [9, 10]. All these applications are based on the analysis of video and 2D images. Images and videos are processed in order to collect gait parameters applying both model-based approaches, using the definition of a 3D model of the body in movement [11–13], or by model-free approaches, that process the silhouette of a walking person [14].

In this paper we implemented a model-based approach using the Microsoft Kinect sensor. Kinect is more than a simple RGB camera, since it is equipped with a depth sensor providing 3D information related to the movements of body joints. The 3D data are more precise compared to the 2D information extracted from images but, on the contrary, the depth sensor, based on the infrared rays, does not work in outside environment and the depth range is quite limited. In literature exists some applications that exploit Kinect 3D data for people recognition and classification. Preis *et al.* [15] used only anthropometric features, such as height, length of limbs, stride length and speed, for gait characterization. They tested few combination of features using three different classifiers: 1R, a C4.5 decision tree and a Naive Bayes classifier. Borrás *et al.* [16] extracted 2D and 3D gait features based on the body silhouettes, to achieve gender identification using a Kernel SVM. Satta *et al.* [17] combined the body silhouette, colors clothing and 3D skeleton data (torso and legs length) aiming at tracking people for video-surveillance application. Ball *et al.* [18] proposed several angular features related to the leg articulations and used  $K$ -means algorithm, with an Euclidean distance metric, for classification.

A preliminary work has been already presented in [19], where we only shown that dynamic gait features extracted from Kinect allow one to discriminate between two subjects with similar biometric features, a much simpler scenario as opposed to the classification task studied in this paper.

The major contributions of this paper are:

- exploitation of cheap and widespread Kinect sensor for joint acquisition of static biometric features, e.g. height, leg length, etc., and dynamic parameters related to gait, e.g. knees movement, head oscillation;
- analysis and selection of the biometric and gait features that are the most effective for people identification;
- experimental campaign worked out on a set of 20 subjects showing that the proposed set of features can be profitably exploited to classify people using SVM.

The paper is organized as follows. In Sect. 2 is explained our method to extract gait features while the clustering approach is presented in Sect. 3. The experiments done are shown in Sect. 4 together with the discussion of achieved results. Conclusion and future work are outlined in Sect. 5.

## 2 Proposed Method

In this work we propose a model-based approach for gait characterization using dynamic skeleton acquisition. In particular, we have used the widespread diffuse Microsoft Kinect sensor for acquisition of gait parameters. Kinect is a popular gaming device, that is able to capture body motion and gestures based on camera and depth sensor. Kinect is able to track in real-time a skeleton model, composed of 20 body joints  $J_i$  shown in Fig. 1. The skeleton can be used to describe body movements in real-time and in 3D space [20].

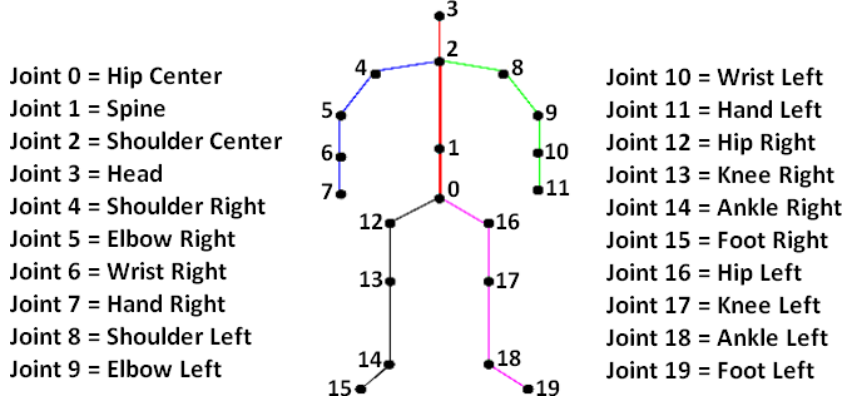
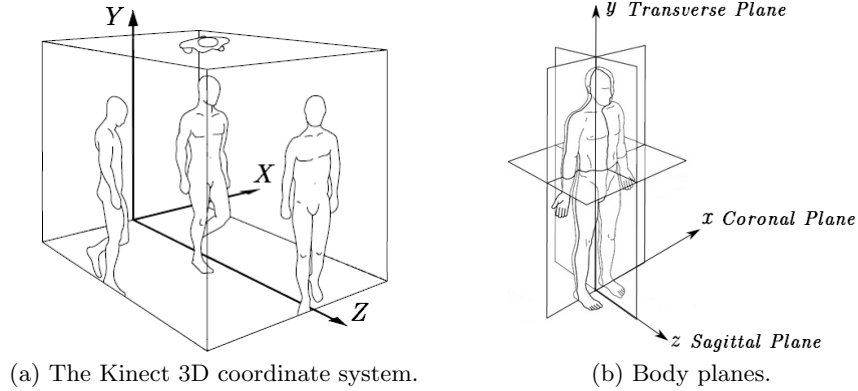


Fig. 1. Kinect skeleton model.

The skeleton data provided by Kinect consist in the coordinates of the 20 joints in its local reference system as shown in Fig. 2-(a) where  $(X, Y, Z)$  are the horizontal, vertical and depth direction, respectively. Joint coordinates are provided at a rate of 30 Hz along with the estimated floor clipping plane, that is the plane where the user is walking. The floor plane equation is derived as  $AX + BY + CZ + D = 0$ , where  $(A, B, C)$  is the normal vector to the plane and  $D$  is the height of the camera center with respect to the floor. As opposed to the Kinect reference system, human biometric parameters are usually measured with respect to the body planes shown in Fig. 2-(b), where coronal, sagittal and transverse plane are represented. The latter is clearly parallel to the floor clipping plane provide by Kinect, whereas the sagittal plane can be determined if we infer the walking direction. Since the Kinect depth range is between 80 centimeters and 4 meters, we can easily assume that the observed subject follows a straight path within a single gait acquisition. As a consequence, we have estimated the walking direction as the line connecting the initial and final coordinates of the center of mass (Joint 0). To make the estimation more robust, the initial and final coordinates of the center of mass are averaged at the beginning and at the end of the acquisition. Finally, the novel reference system  $(x, y, z)$  is constructed as shown in Fig. 2-(b), considering the floor clipping plane, its normal and the walking direction as the novel  $z$  axis. In the following all the joints coordinates will be expressed according to  $(x, y, z)$ .

Since Kinect depth sensor exhibits limited resolution and precision, all joint estimates can be considered as noisy acquisitions. To limit this effect we propose to use median filter on all acquired data before performing any further processing. Such refined estimates are then exploited to define our gaits feature vector that comprises both physical and behavioral parameters, as detailed in the following. All the collected features are summarized in Table 1.



**Fig. 2.** Kinect reference system and coordinate system according to the walking direction.

## 2.1 Physical features

Using joint coordinates we have estimated some physical biometric features that are almost constant during the walk, namely the height and the length of arms and legs, respectively. The height is defined as:

$$\gamma = \sum_{i \in \tau} \sqrt{(J_{i,x} - J_{i+1,x})^2 + (J_{i,y} - J_{i+1,y})^2 + (J_{i,z} - J_{i+1,z})^2}$$

where  $\tau = [3, 2, 1, 0, 12, 13, 14, 15]$  or  $\tau = [3, 2, 1, 0, 16, 17, 18, 19]$ , i.e. the joints going from the head to the right (or left) foot. All the estimates of  $\gamma$  acquired during a given gait acquisition are averaged to get a single average value. Similarly, we compute the length of the left (and right) arm as the overall length of the skeleton segments from the shoulder to the hand. A single parameter  $\alpha$  is obtained by averaging both left and right arm lengths along all the acquisition period. The same approach is used to get a single feature  $\beta$  by averaging the lengths of the skeleton segments from the hip to the feet.

## 2.2 Behavioral features

A clear advantage of using Kinect is the possibility to devise a rich set of behavioral gait features derived by the continuous monitoring of joint positions. Usually gait behavior is characterized in terms of stride length and walking speed. The former is the distance between two stationary position of the same foot while walking. The stride is detected using the same technique we already presented in [19]. The latter is the ratio between the distance covered by the center of mass and its duration in a single stride. In this study we estimate such features in real time and all their measurements are averaged to get the gait parameters  $\delta$  and  $v$  shown in Table 1. We complement these standard gait

**Table 1.** Gait features list.

Label	Features
$\alpha$	arms length
$\beta$	legs length
$\gamma$	height
$\delta$	stride length
$v$	walking speed
$\epsilon$	elbows distance
$\kappa$	knees distance
$\eta$	hands distance
$\lambda$	ankles distance
$\mu_{J_{3,x}} - \sigma_{J_{3,x}}^2$	mean/variance of head (along $x$ )
$\mu_{J_{3,y}} - \sigma_{J_{3,y}}^2$	mean/variance of head (along $y$ )
$\mu_{J_{4,x}} - \sigma_{J_{4,x}}^2$	mean/variance of left shoulder (along $x$ )
$\mu_{J_{8,x}} - \sigma_{J_{8,x}}^2$	mean/variance of right shoulder (along $x$ )
$\mu_{J_{13,y}} - \sigma_{J_{13,y}}^2$	mean/variance of left knee (along $y$ )
$\mu_{J_{17,y}} - \sigma_{J_{17,y}}^2$	mean/variance of right knee (along $y$ )

parameters with other behavioral features that are related to the gait dynamic, in particular the movement of head, shoulders, elbows, hands, knees and ankles. To this end, we track the movement of the corresponding joints along the  $x$  and  $y$  axes, i.e. we measure the vertical and lateral oscillations of the corresponding skeleton joints. Then the dynamic of their trajectory during the acquisition is represented in terms of mean value  $\mu$  and variance  $\sigma^2$ . These features are denoted as  $\mu_{J_{i,x}}, \mu_{J_{i,y}}, \sigma_{J_{i,x}}^2, \sigma_{J_{i,y}}^2$  where  $i$  indexes the relevant joint and  $x$  (or  $y$ ) refers to the considered coordinate component (see Table 1).

Finally, our experiments showed that other important gait features are the distance between the left and right elbows  $\epsilon$ , hands  $\eta$ , knees  $\kappa$  and ankles  $\lambda$ . Also in this case we average all the estimates during the acquired walking period. All the collected features are summarized in Table 1.

### 3 Classification

In this paper we used Support Vector Machine (SVM), a widely-used and robust clustering algorithm, to classify people based on both physical and behavioral gait biometric feature. SVM is a supervised learning method for data classification. The main idea is to find the pair of parallel hyperplane that separates between a set of features having different class memberships. Each pair of parallel hyperplanes is characterized by specific sets of feature points, the so-called support vector. In a  $2D$  feature space the planes are fully defined by three support vectors, so in  $ND$  the support vector examples should have  $N + 1$  points, in order to avoid overfitting. Since SVM is a supervised algorithm, the classification task involves separating data into training and testing sets. Each instance in the training set contains one “target value” (i.e. the class labels) and several

“attributes” (i.e. the features or observed variables). The goal of SVM is to produce a model (based on the training data) which predicts the target values of the test data.

The hyperplane algorithm proposed by Vapnik in 1963 was a linear classifier. The nonlinear classification, proposed in 1992 [21], is possible by applying the so-called kernel trick, originally proposed by Aizerman *et al.* [22]: the dimensional space of the function is mapped into an higher (maybe infinite) dimensional space, so that it is possible to find a linear separating hyperplane with the maximal margin in that higher dimensional space. The advantage of the kernel function is that it does not require explicit calculation of the data coordinates in the new space. It is done by simply computing the inner products between the images of all pairs of data in the feature space: this is a computationally cheaper operation respect to the explicit computation of the coordinates. To construct an optimal hyperplane, SVM employs an iterative training algorithm, which is used to minimize an error function. According to the form of the error function, in multi-class classification, SVM models can be classified into two distinct groups: classification SVM Type 1 (also known as  $C$ -SVM classification) and classification SVM Type 2 (also known as  $\nu$ -SVM classification).

The  $C$ -SVM training involves the minimization of the error function:

$$\frac{1}{2}w^T w + C \sum_{i=1}^N \xi_i, \quad C > 0$$

subject to the constraints:

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0, \quad i = 1, \dots, N$$

where  $C$  is the capacity constant,  $w$  is the vector of coefficients and  $\xi$  represents parameters for handling nonseparable data (input).  $x_i$  is the training vector, so the index  $i$  labels the  $N$  training samples, that are mapped into an higher dimensional space by the function  $\phi$ .  $b$  is a constant.  $C$  should be chosen with care to avoid overfitting.

The  $\nu$ -SVM model minimizes the error function:

$$\frac{1}{2}w^T w - \nu\rho + \frac{1}{N} \sum_{i=1}^N \xi_i, \quad 0 \leq \nu \leq 1$$

subject to the constraints:

$$y_i(w^T \phi(x_i) + b) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, N \quad \text{and} \quad \rho \geq 0$$

The kernel functions that can be used in SVM are:

- Linear:  $K(x_i, x_j) = x_i^T x_j$ .
- Polynomial:  $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$ .
- Radial Basis Function (RBF):  $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$ .
- Sigmoid:  $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$ .



where  $\gamma$ ,  $r$  and  $d$  are kernel parameters.

The effectiveness of SVM depends on the selection of kernel and the kernel's parameters.

## 4 Experimental Results

The goal of our experimentation is, first to find out which features collected in Sect. 2 are more relevant for gait characterization, and then to evaluate the achievable classification accuracy.

Since a standard gait dataset acquired with the Kinect sensor is not available, we have collected a set of gait samples. To this end we acquired gait samples with Kinect for Windows, recording 20 subjects that were asked to walk naturally along a corridor; each subject is acquired 10 times for a total of 200 gait samples. The software model to estimate the features proposed in Sect. 2 has been developed using SDK 1.7.

Finally, for classification, we resort to the LIBSVM tool [23], an SVM free library. We used both the  $C$ -SVM type and the  $\nu$ -SVM type, with linear and RBF kernel function. The dataset is divided in two subset: 60% of samples are used for training and the remaining for testing. Before applying SVM we normalized our features matrix to the range  $[0, 1]$ . Scaling data is very important in order to avoid attributes in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to simplify the calculations of the kernel function. As said in Sect. 3 the parameters  $C$  and  $\nu$  should be chosen with care. The  $C$  parameter is computed using the  $K$ -fold cross-validation procedure [23]. The training set is divided into  $K$  subsets of equal size. A single subset is tested using the classifier trained on the remaining  $K - 1$  subsets. The process is repeated  $K$  times for each of the  $K$  subset. The  $K$  results are then averaged to produce a single estimation. The  $\nu$  parameter is calculated as suggested in [24]:

$$\nu = 2 * \min(|SVs^+|, |SVs^-|) / |SVs|$$

where  $SVs$  is the vector containing the support vector elements (the feature points of each hyperplane), while  $SVs^+$  and  $SVs^-$  are the vectors containing the positive and negative support vectors, respectively.

As an objective metric for performance evaluation we used the classification accuracy, defined as:

$$accuracy = \frac{\psi}{\Omega} * 100\%$$

where  $\psi$  is the number of correctly classified samples and  $\Omega$  is the total number of samples.

### 4.1 Discussion

To investigate the effectiveness of the different features that we proposed, we analyze the classification accuracy obtained with several sets of gait parameters.

**Table 2.** Classification accuracy of different features sets.

Set	Features	C-SVM		$\nu$ -SVM	
		Linear	RBF	Linear	RBF
$S_1$	$\alpha, \beta, \gamma$	37.50	25.00	62.50	58.75
$S_2$	$\delta, v$	23.75	23.75	13.75	17.50
$S_3$	$\alpha, \beta, \gamma, \delta, v$	46.25	41.25	48.75	37.50
$S_4$	$\epsilon, \kappa,$	37.50	38.75	52.50	46.25
$S_5$	$\epsilon, \kappa, \lambda$	42.50	37.50	55.00	53.75
$S_6$	$\epsilon, \kappa, \eta$	41.25	38.75	32.50	35.00
$S_7$	$\epsilon, \kappa, \lambda, \eta$	45.00	43.75	56.25	55.00
$S_8$	$\lambda, \eta$	18.75	17.50	15.00	10.00
$S_9$	$\alpha, \beta, \gamma, \epsilon, \kappa$	72.50	55.00	75.00	76.25
$S_{10}$	$S_1, S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2$	70.00	56.25	61.25	46.25
$S_{11}$	$S_1, S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}$	86.25	85.00	86.25	85.00
$S_{12}$	$S_1, S_4, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	75.00	52.50	63.75	51.25
$S_{13}$	$S_1, S_4, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	81.25	65.00	81.25	81.25
$S_{14}$	$S_1, S_4, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	66.25	55.00	55.00	45.00
$S_{15}$	$S_1, S_4, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	77.50	71.25	77.50	77.50
$S_{16}$	$S_1, S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	65.00	57.50	57.50	41.25
$S_{17}$	$S_1, S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	83.75	82.50	86.25	83.75
$S_{18}$	$S_1, S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	70.00	53.75	55.00	46.25
$S_{19}$	$S_1, S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	<b>88.75</b>	<b>82.50</b>	<b>90.00</b>	<b>90.00</b>
$S_{20}$	$S_1, S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	66.25	55.00	58.75	42.50
$S_{21}$	$S_1, S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	87.50	82.50	90.00	88.75
$S_{22}$	$S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2$	48.75	41.25	46.25	43.75
$S_{23}$	$S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}$	90.00	80.00	92.50	91.25
$S_{24}$	$S_4, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	36.25	35.00	46.25	48.75
$S_{25}$	$S_4, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	58.75	57.50	57.50	45.00
$S_{26}$	$S_4, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	47.50	38.75	55.00	51.25
$S_{27}$	$S_4, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	61.25	50.00	57.50	57.50
$S_{28}$	$S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	48.75	37.50	45.00	35.00
$S_{29}$	$S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	86.25	76.25	87.50	88.75
$S_{30}$	$S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	45.00	32.50	32.50	32.50
$S_{31}$	$S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	<b>92.50</b>	<b>81.25</b>	<b>96.25</b>	<b>96.25</b>
$S_{32}$	$S_4, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	48.75	32.50	41.25	30.00
$S_{33}$	$S_4, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	92.50	78.75	76.25	65.00
$S_{34}$	$S_1, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2$	45.00	42.50	43.75	37.50
$S_{35}$	$S_1, \mu_{J_{3,x}}, \mu_{J_{3,y}}$	71.25	65.00	73.75	73.75
$S_{36}$	$S_1, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	40.00	27.50	46.25	33.75
$S_{37}$	$S_1, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	61.25	45.00	67.50	70.00
$S_{38}$	$S_1, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	45.00	43.75	37.50	31.25
$S_{39}$	$S_1, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	55.00	43.75	62.50	62.50
$S_{40}$	$S_1, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2$	48.75	51.25	47.50	36.25
$S_{41}$	$S_1, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}$	68.75	58.75	75.00	76.25
$S_{42}$	$S_1, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	46.25	46.25	41.25	33.75
$S_{43}$	$S_1, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	<b>77.50</b>	<b>60.00</b>	<b>75.00</b>	<b>78.75</b>
$S_{44}$	$S_1, \sigma_{J_{3,x}}^2, \sigma_{J_{3,y}}^2, \sigma_{J_{4,x}}^2, \sigma_{J_{8,x}}^2, \sigma_{J_{13,y}}^2, \sigma_{J_{17,y}}^2$	50.00	51.25	42.50	35.00
$S_{45}$	$S_1, \mu_{J_{3,x}}, \mu_{J_{3,y}}, \mu_{J_{4,x}}, \mu_{J_{8,x}}, \mu_{J_{13,y}}, \mu_{J_{17,y}}$	78.75	58.75	77.50	80.00

As a benchmark for classification accuracy, we run a first experiment where we discriminate subjects considering only physical features, in particular their height ( $\gamma$ ) and arms ( $\alpha$ ) and legs ( $\beta$ ) lengths; this is representative of usual classification based only on simple anthropometric parameters. The classification results using this set of features ( $S_1$ ) are shown in Table 2, where it can be noted that  $\nu$ -SVM yields an accuracy of about 60%. As a second trial we used the set  $S_2$  that includes two standard gait parameters, namely stride length ( $\delta$ ) and walking speed ( $v$ ) but, as shown in Table 2, the achieved classification accuracy turns to be rather poor. Also, jointly using sets  $S_1$  and  $S_2$  (set  $S_3$  in Table 2) we cannot improve performance, significantly. The limited performance obtained using gait parameters in  $S_2$  can be explained by considering the limited depth range of the Kinect, that allows one to acquire only three/four strides, leading to poor estimates of  $\delta$  and  $v$ .

As a consequence, we move our investigation towards other gait features that can be estimated more reliably by Kinect, i.e. the inter-distance between corresponding left and right body joints. A subset of the obtained classification results is reported in Table 2 (see sets  $S_4$ - $S_9$ ). Inspecting these results one can notice that, using physical features and the distance between elbows and knees (set  $S_9$ ) we achieve about 75% accuracy, with an improvement of about 15% with respect to the  $S_1$  benchmark.

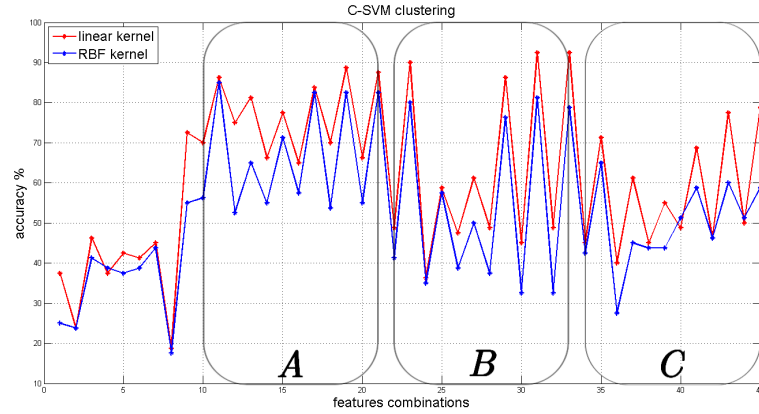
Then, we try to further improve the classification accuracy including in the features vector the remaining dynamic parameters defined in Sect. 2, namely the mean and variance of the position of head, shoulders, elbow and knees. In sets  $S_{10}$ - $S_{21}$  these latter are used along with  $S_1$  and  $S_4$ . It can be noted that a few parameter sets yield accuracy larger than 80%. In particular, the best classification is obtained by  $S_{19}$  that yields 90% accuracy with  $\nu$ -SVM. In sets  $S_{22}$ - $S_{33}$  we perform the same experiments removing  $S_1$  from the feature vector, i.e. without using the standard physical features. In this case a few combinations reach an accuracy around 90%. In particular, the features set  $S_{31}$  exhibits an accuracy of 96%, that represents the best result we obtained. Finally, the remaining experiments  $S_{34}$ - $S_{45}$  use  $S_1$  along with the proposed dynamic gait parameters, but the obtained accuracy does not achieves 80%.

To better appreciate all our experimental results in Fig. 3 we plot the accuracy as a function of the features sets for  $C$ -SVM (a) and  $\nu$ -SVM (b), respectively. To improve the readability of the graph the last three groups of parameters sets, are highlighted by boxes labeled by  $A$ ,  $B$  and  $C$  respectively.

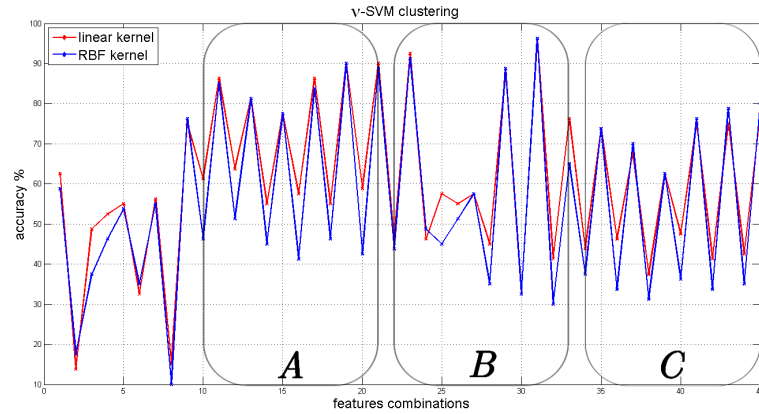
In conclusion, our experiments show that gait parameters extracted using Kinect, can be used as a powerful biometric feature. In particular, we can argue that simple statistics estimated from a set of skeletal joints are very effective for people classification based on SVM.

## 5 Conclusions and Future Work

Relying on the information provided by the Kinect sensor, an analysis system able to classify people based on their gait has been proposed. We have especially



(a)



(b)

**Fig. 3.** Classification accuracy of  $C$ -SVM (a) and  $\nu$ -SVM (b) for different combinations of gait features.

pointed out that the movement of elbows, knees and head are of great importance for diversifying gait. The achieved results obtained by using the robust SVM clustering algorithm show an accuracy classification equal to 96.25% when a suitable parameter set is chosen. Future works include the exploitation of other statistical procedures, such as the principal component analysis, to better rank the importance of gait parameters in terms of classification accuracy. Moreover, we are planning an extensive gait acquisition campaign to better validate the classification results from a statistical viewpoint. From the applicative point of view we will consider other scenarios that go beyond classification, e.g. forensic and health/well-being contexts.

## References

1. Julian Ashbourn. *Biometrics - advanced identity verification: the complete guide*. Springer, 2002.
2. Chiraz BenAbdelkader, Ross Cutler, and Larry S. Davis. View-invariant estimation of height and stride for gait recognition. In Massimo Tistarelli, Josef Bign, and Anil K. Jain, editors, *Biometric Authentication*, volume 2359 of *Lecture Notes in Computer Science*, pages 155–167. Springer, 2002.
3. R. Cucchiara, C. Grana, A. Prati, and R. Vezzani. Computer vision system for in-house video surveillance. In *IEE Proceedings of Vision, Image and Signal Processing*, 2005.
4. Michela Goffredo, Imed Bouchrika, John N. Carter, and Mark S. Nixon. Performance analysis for automated gait extraction and recognition in multi-camera surveillance. *Multimedia Tools Appl.*, 50(1):75–94, October 2010.
5. P. K. Larsen, E. B. Simonsen, and N. Lynnerup. Gait analysis in forensic medicine. In *Proc. Videometrics IX.*, volume 6491, January 2007.
6. Imed Bouchrika, Michaela Goffredo, John Carter, and Mark Nixon. On using gait in forensic biometrics. *Journal of Forensic Sciences*, 56(4):882–889, July 2011.
7. D.A. Winter. *The biomechanics and motor control of human gait: normal, elderly and pathological*. University of Waterloo Press, 1991.
8. M E Eastlack, J Arvidson, L Snyder-Mackler, J V Danoff, and C L McGarvey. Interrater reliability of videotaped observational gait-analysis assessments. *Phys Ther*, 71(6):465–72, 1991.
9. RetoW. Kressig and Olivier Beauchet. Guidelines for clinical applications of spatio-temporal gait analysis in older adults. *Aging Clinical and Experimental Research*, 18(2):174–176, 2006.
10. Marybeth Brown, David R. Sinacore, Ellen F. Binder, and Wendy M. Kohrt. Physical and performance measures for the identification of mild to moderate frailty. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 55(6):M350–M355, 2000.
11. Rong Zhang, Christian Vogler, and Dimitris Metaxas. Human gait recognition. In *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 1 - Volume 01*, CVPRW '04, pages 18–25, Washington, DC, USA, 2004. IEEE Computer Society.
12. Raquel Urtasun and Pascal Fua. 3d tracking for gait characterization and recognition. In *In: Proc. IEEE Automatic Face and Gesture Recognition, Seoul, Korea*, pages 17–22, 2004.
13. Jang-Hee Yoo and Mark Nixon. Automated markerless analysis of human gait motion for recognition and classification. *ETRI Journal*, 33(3):259–266, April 2011.
14. Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. Silhouette analysis-based gait recognition for human identification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(12):1505–1518, December 2003.
15. Johannes Preis, Moritz Kessel, Martin Werner, and Claudia Linnhoff-Popien. Gait recognition with kinect. In *Proceedings of the First Workshop on Kinect in Pervasive Computing*, 2012.
16. Ricard Borrs, gata Lapedriza, and Laura Igual. Depth information in human gait analysis: An experimental study on gender recognition. In Aurlio Campilho and Mohamed Kamel, editors, *Image Analysis and Recognition*, volume 7325 of *Lecture Notes in Computer Science*, pages 98–105. Springer Berlin Heidelberg, 2012.

17. Riccardo Satta, Federico Pala, Giorgio Fumera, and Fabio Roli. Real-time appearance-based person re-identification over multiple kinect cameras. In *8th International Conference on Computer Vision Theory and Applications (VISAPP 2013)*, Barcelona, Spain, 21/02/2013 2013.
18. Adrian Ball, David Rye, Fabio Ramos, and Mari Velonaki. Unsupervised clustering of people from 'skeleton' data. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, HRI '12*, pages 225–226, New York, NY, USA, 2012. ACM.
19. E. Gianaria, N. Balossino, M. Grangetto, and M. Lucenteforte. Gait characterization using dynamic skeleton acquisition. In *Multimedia Signal Processing (MMSP), 2013 IEEE 15th International Workshop on*, pages 440–445, Sept 2013.
20. Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Commun. ACM*, 56(1):116–124, January 2013.
21. Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT '92*, pages 144–152, New York, NY, USA, 1992. ACM.
22. M. A. Aizerman, E. A. Braverman, and L. Rozonoer. Theoretical foundations of the potential function method in pattern recognition learning. In *Automation and Remote Control*,, number 25, pages 821–837, 1964.
23. Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.
24. David J. Crisp and Christopher J. C. Burges. A geometric interpretation of v-svm classifiers. In S.A. Solla, T.K. Leen, and K. Müller, editors, *Advances in Neural Information Processing Systems 12*, pages 244–250. MIT Press, 2000.