Feature selection of motion capture data in gait identification challenge problem

Adam Świtoński[1], Henryk Josiński[2], Agnieszka Michalczuk[1], Przemysław Pruszowski[2], Konrad Wojciechowski[1]

[1] Polish-Japanese Institute of Information Technology,
Aleja Legionów 2, 41-902 Bytom, Poland
{aswitonski,amichalczuk,kwojciechowski}@pjwstk.edu.pl
[2]Silesian University of Technology,
ul. Akademicka 16, 41-100 Gliwice, Poland
{henryk.josinski,przemyslaw.pruszowski}@polsl.pl

Abstract . The method of discovering robust gait signatures containing strong discriminative properties is proposed. It is based on feature extraction and selection of motion capture data. Three different approaches of feature extraction applied to Euler angles and their first and second derivates are considered. The proper supervised classification is preceded by specified selection scenario. On the basis of the obtained precision of person gait identification, analyzed feature sets are assessed. To examine proposed method database containing 353 gaits of 25 different males is used. The results are satisfactory. In the best case the recognition accuracy of 97% is achieved. On the basis of classification which takes into consideration only the data of the specified segments, the ranking is constructed. It corresponds to the evaluation of individual features of the joint movements.

Keywords: gait identification, motion capture, supervised learning, feature selection, feature extraction, biometrics

1 Introduction

In a motion capture acquisition positions of attached markers on human body are tracked by calibrated multicamera system. Thus basic motion data representation contains time sequences of global 3D coordinates of the markers. It can be transformed to skeletal model representation with a kinematical chain of a tree-like structure. The root object of the tree usually corresponds to lower part of a spine and subsequent nodes point to following joints. Thus every pose is described by joints rotations and global translation of human body in respect to specified reference frame. The rotation are coded by three Euler angles or unit quaternions by default. .

There are two basic approaches of skeleton model estimation [1]. If the markers are located in specified anatomical points of a human body, the simple set of geometric transformations is sufficient. However such a method is very sensitive even to slight markers displacements. That is a reason why preliminary matching of markers and

adfa, p. 1, 2011.

skeleton segments usually is carried out by clustering techniques. It requires a special set of movements to be performed by a human, prior to proper acquisition, which should take into consideration all degrees of freedom of assumed skeleton model.

However direct applications of motion capture acquisition in deployed person gait identification systems is problematic. It removes one of the most important advantages of gait identification - non-awareness of recognized human and what is more the acquisition is time consuming - man has to put special suit with attached makers and perform prepared set of movements for the sake of skeleton calibration purposes. Thus, markerless motion capture can be utilized. The problem of skeleton estimation by such an acquisition in most cases leads to nonlinear high dimensional parametric optimization. The skeleton configuration which matches image data in best way has to be determined. Most often used technique to explore configuration spaces are particle filters. For instance in [2] particle swarm optimization is applied and the obtained mean markers displacement in respect to reference to Vicon motion capture is 50mm. The price which has to be paid for much convenient acquisition is worse precision of measurements. Thus application of motion capture in development phase of gait identification system is justified. It allows to focus on classification stage without influence of acquisition noise and to obtain best possible results. If the classification is satisfactory the second stage has to be completed - proper choice and parameters tuning of markerless motion capture.

. The paper presents the method of gait classification based on feature extraction and selection of motion capture data. The final recognition is carried out by supervised classification. To examine proposed approach gait data collected in human motion capture laboratory is utilized. On the basis of obtained classification accuracies features spaces are explored and their subsets are evaluated. It is the main contribution of the paper which relates to discovered features corresponding to the most valuable individual gait properties.

2 Related work

As described in previous section, motion data is represented as time sequence of pose parameters, markers positions or silhouettes. There are three basic approaches to classification of such a data: feature extraction, dynamic time warping and Hidden Markow Models.

.In the first approach, features of time sequences are calculated and motion descriptors are constructed. On the basis of obtained feature vectors, subsequent classification is carried out - for instance machine learning can be applied. In [3] generic extraction is proposed which utilizes tensor reduction technique by multilinear principle component analysis. The final recognition is performed by selected distance function. In [4] and [5] four types of features sets and supervised classification are used for motion capture data. In [6] all body parameters are transformed into the frequency domain and first two lowest Fourier components are chosen. Afterwards PCA reduction is applied.

Dynamic time warping (DTW), originally introduced to spoken word recognition [7], is normalization technique which matches poses of compared motions on the basis of monotonic transformation. It makes them faster or slower in subsequent time instants to obtain the lowest total distance of corresponding poses. DTW is utilized to estimate motion dissimilarity which is a crucial challenge in nearest neighbors classification scheme. In [8] DTW is used for binary relational features, in [9] different distance metrics of Euler angles and unit quaternions spaces are examined for DTW classification of gait motion capture data and in [10] DTW is based on linearly and nonlinearly reduced silhouettes.

In Hidden Markow Model (HMM) approaches motion is modeled as Markow chain. Models are hidden, because it is stated assumption that poses depend on states which are unknown for an observer. A single model usually is taken for every class and its parameters contains probability transition matrix between states in subsequent time moments and probability distribution of poses in specified states. In the training phase parameters of HMM for every class are calculated and in classification stage model with the greatest probability is determined. Exemplary HMM based methods are presented in [11], [12] and [13].

3 Collected database

To validate proposed method and to discover most valuable individual features, motion database in PJWSTK Human Motion Laboratory[1] was collected. It contains 353 gaits of males at the age from 25 to 35 years old - students and staff of PJWSTK university. The gait route was specified to be a straight line of about 5 meters long. Example data is visualized in Fig. 1.

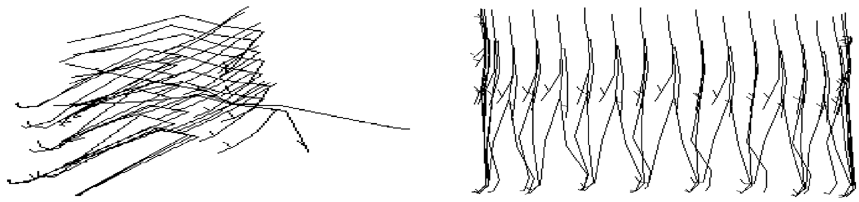

Fig. 1. Example collected gait

In acquisition Vicon Blade software was applied with default skeleton model containing 22 bone segments as presented in Fig. **2**a. Complete pose description is defined in 72 dimensional space. There are 23 rotations coded by three Euler angles - additional one is associated with global skeleton rotation The pose space also contains a three dimensional vector related to global translation.

Gait can be defined as coordinated cyclic combination of movements which results in human locomotion [14]. What is more, because requirement of calibration process of Vicon Blade software, every gait starts and terminates with T-pose, which is quite untypical during normal gait. Thus to remove such a data and to detect representative

[1] http://hml.pjwstk.edu.pl/en/

main cycle containing two adjacent steps distance tracking between two feet is carried out, as is presented in [5].

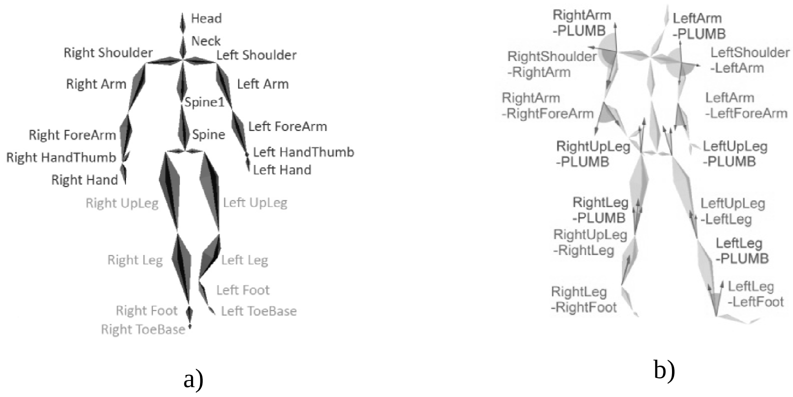

Fig. 2. Skeleton model a) upper and lower body segment names, b) custom angles

To avoid learning and recognition on the basis of gait location, which seems to be strongly related to place of acquisition, the global translation vector is removed from motion data and it is not taken into consideration by subsequent feature extraction procedures.

4 Feature extraction

The proper classification is based on the sequences of Euler angles triples, representing joint rotation in subsequent time instants. What is more, considering the results obtained in gait paths classification [5], the first and the second derivates are calculated by simple differential filtering. It gives in total 207 separate time sequences - 22 segments and global rotation, each of them represented by three Euler angles and their first two derivates.

Example time sequences of randomly selected five persons, reduced to detected main cycle, are visualized in Fig. **3** and in Fig. **4**. The charts present Euler angles and their first and second derivates in respect to local coordinate system RX-RY-RZ. What is more rotational data is transformed into representation of angles between two adjacent segments by calculating arc cosine of dot products of unit vectors related to segment orientation in 3D space. For instance in case of left knee joint it is the angle between LeftUpLeg and LeftLeg segments from Fig. **2**. It allows to visualize aggregated dependency among joint orientation, its angular velocity and acceleration. as shown in Fig. **3**d and Fig. **4**. Different scales on charts in Fig. **3**d and Fig. **3**a are caused by opposite meaning of presented data. In case of raw Euler angles supplied by motion capture acquisition, the pose is specified in respect to given reference frame and after above described transformation, angles correspond to direct rotation which has to be performed to match both adjacent segments.

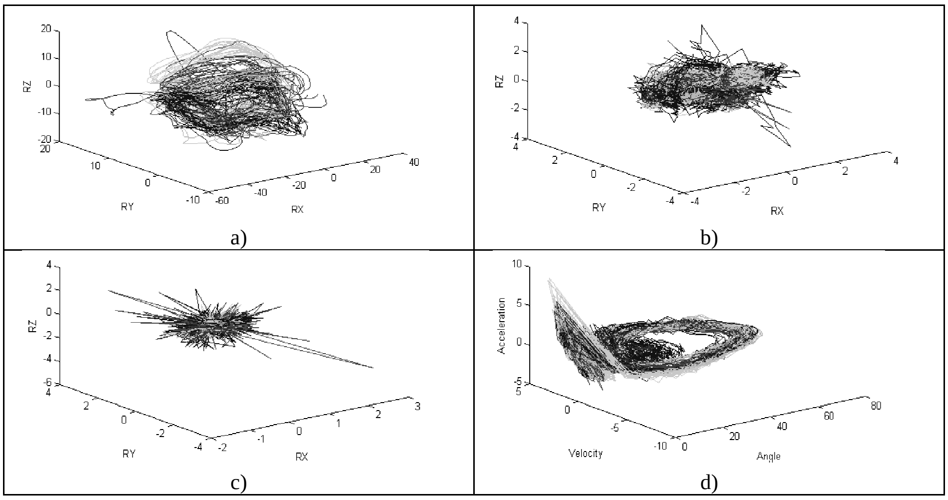

Fig. 3. Left hip joint time sequences: a) Euler angles, b) angular velocities, c) angular accelerations, d) single angle of rotation

Data of different humans in Fig. **3** and in Fig. **4** are marked by separate colors. It is very difficult to manually point discriminative properties of presented time series. What is more problem is probably more challenging and efficient classification requires simultaneous analysis of multiple joints movements.

Similar to methods presented in [5] and [4], three types of feature sets are applied in extraction:

- Statistical
- Fourier transform
- TimeLine

In the first approach two most basic statistics - mean value and variance of every considered pose attribute are calculated. Such a simple feature extraction does not take into consideration dependencies in time domain and analyzes only the emerging values of sequences.

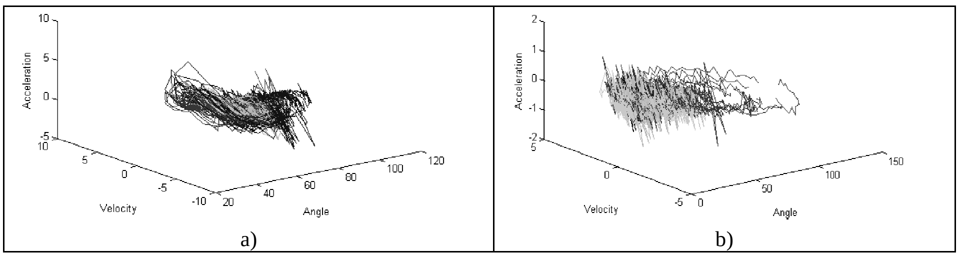

Fig. 4. Left knee and left arm joints time sequences a) left knee, b) left arm

In Fourier approach a motion is transformed into frequency domain and first twenty components with the lowest frequencies are taken. The feature set includes the module of the complex number, which gives information of the total intensity of a given frequency and the phase that points its time shift. Similar like in [5] additional representation is constructed by applying linear scaling of the time domain to the equal number of 100 frames. In such a case frequency components are more compatible across different recordings, which may be crucial for their efficient comparison and classification.

In the last extraction approach called timeline, the feature set stores information of every attribute values as time sequence. The moments in which attribute values are taken into the set are determined by the division of the motion into the specified number of intervals. Timeline motion representation is calculated with sequence of five, ten, twenty, fifty and one hundred different time moments

5 Feature selection

The proposed feature spaces are high dimensional. In case of most simple statistical extraction feature space contains 414 attributes and for the most complex timeline approach with 100 intervals 20700 separate features are determined. The hypothesis of useless attributes without discriminative, individual properties can be stated in such features spaces. It is proven that irrelevant or distracting attributes to a dataset often confuses machine learning systems [15], which may result in worse classification accuracy.

. To verify the hypothesis of useless features and to discover the most valuable ones, attribute selection is carried out prior to proper classification. Ii is based on the manually proposed scenarios with feature subsets evaluated on the basis of obtained identification efficiency. At the current stage, we have not used automatic selection techniques because of the complexity of the problem and to get more interpretable results.

The prepared selection scenarios are as follows. In all cases we selected every possible combination of attributes associated with:

- axis RX, RY and RZ of the local coordinate system.
- rotation, angular velocity and acceleration
- statistical feature sets: mean and variance
- Fourier feature sets
  - first n Fourier components - low pass filtering, n in the range (1,20),
  - absolute value, phase, real and imaginary parts of complex number

The extra attribute selection takes into consideration joints of applied skeleton model. There is no possibility to examine all their combinations because of limited computational power, thus experiments were iterated independently for pose described by all joints, reduced to only single one joint and containing lower and upper body parts. Such a division relates to root element location and it is shown in Fig. **2**a - the lower segments are labeled by green color and upper ones by red. What is more the custom combination of angles, containing selected data of direct relationships between adjacent segments and their orientation to vertical axis is investigated as presented in Fig. **2**b. The combination was prepared according to our subjective expectations to keep individual gait features. Motion data in this case is aggregated into absolute values of angles without distinction of three basic rotations.

6 Experiments, results and conclusions

The experiments were iterated for all previously described feature sets and their selection scenarios. What is more additional variants are generated by applied prefiltrations containing main cycle detection, linear scaling of the time domain into given number of frames and also by (α, β, χ) Euler angles transformation which determines aggregated total angle between two adjacent segments according to equation (1). The number 100 of frames used in linear scaling was adjusted to roughly approximated mean duration of the main cycles.

$$angle(\alpha, \beta, \chi) = \arccos(\cos\alpha \cdot \cos\beta \cdot \cos\chi - \sin\alpha \cdot \sin\beta \cdot \sin\chi) \qquad (1)$$

Because of numerous experiments were required, applied classifiers have to be characterized by low computational requirements. That is a reason why k nearest neighbors (kNN) [15] and naive Bayes [15] classifiers are chosen. The number k of considered neighbors for kNN is in the range <1,10> and for Naive Bayes parametrical estimation of normal distribution and non parametrical kernel based one are used. To split collected gait instances into training and testing parts, leave one out validation is utilized [15].

Obtained results are shown in Fig. **5**, **6**, **7**, **8** and **9**.which present best achieved accuracy of identification expressed by percent of correctly classified gaits in respect to specified feature extraction and selection approaches and applied prefiltration.

The results are satisfactory. Best obtained precision of recognition is 97.1%, which means only 10 misclassified gaits of 353. It belongs to first five Fourier components calculated for whole set of Euler angles and their first derivates of gaits reduced to detected main cycle.

According to expectations, both main cycle detection and linear scaling improve performance of feature extraction and subsequent classification – the former noticeable, the latter slightly, as shown in Fig . **5** and **6**. To achieve maximum possible accuracy of identification. analysis of complete joint description containing three Euler angles is necessary. Its simplification to single rotational angle removes some individual features. This probably explains unexpected worse results obtained by custom angles from Fig **2**b.

The most precise classification is carried out on the basis of Fourier extraction. Beside the case of custom angles. timeline extraction is much worse. Simple statistical features are suitable for recognition with 95% accuracy, which is even better in comparing to much more complex timeline features.

First and second derivates corresponding to angular velocity and acceleration still represent strong individual features. They allow to recognize gaits with 95.4% and 95.9% precision, respectively, which is only less than 1% worse in comparing to raw

Euler angles. If the derivates are combined together in a single gait signature accuracy increases to 96.3%.

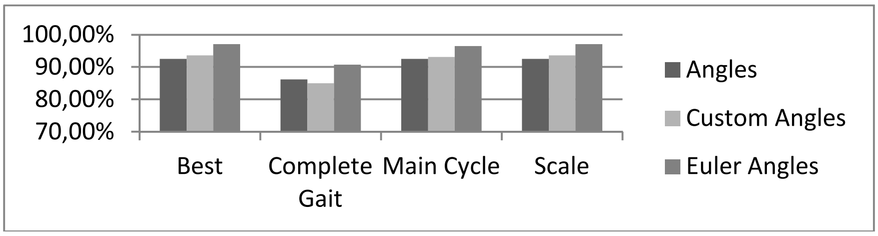

Fig. 5. Classification results in respect to different prefiltrations and parameters of a skeleton data

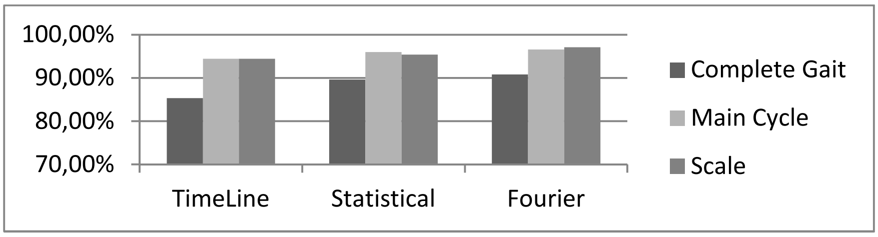

Fig. 6. Classification results in respect to a different extraction approaches and prefiltrations.

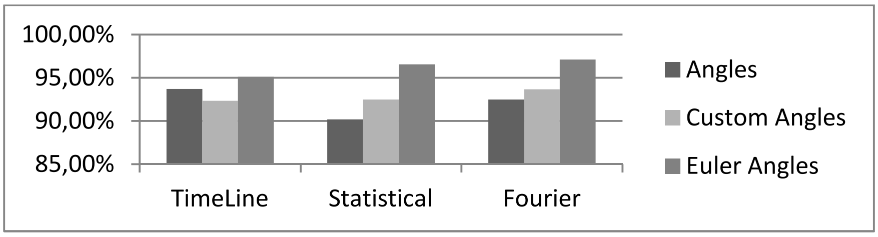

Fig. 7. Classification results in respect to different extraction approaches and parameters of a skeleton data

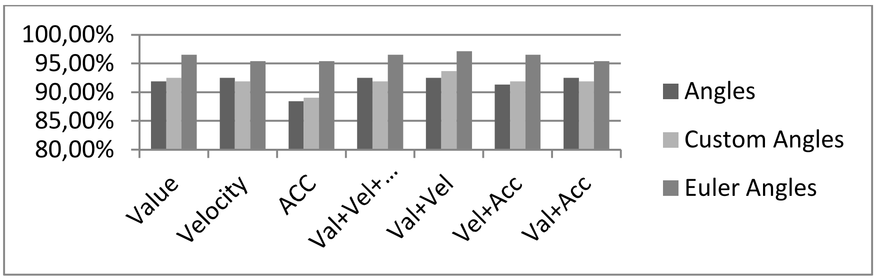

Fig. 8. Classification results in respect to derivates and parameters of a skeleton data

It is sufficient to persist only first five Fourier components to obtain maximum accuracy. Subsequent ones are more influenced by the acquisition noise, which slightly worsens the results. The single first Fourier component corresponding to mean value is efficient only for Euler angles and not for their derivates. It can be explained by individual features of a human posture which are reflected by skeleton Euler angles. while derivates correspond to the joint movements.

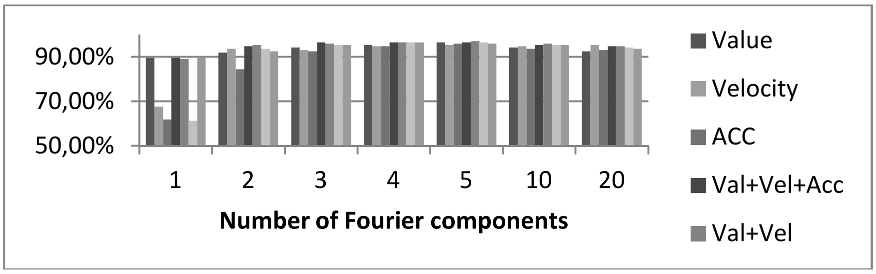

Fig. 9. Classification results in respect to selected number of Fourier components and derivates

On the basis of classification performed by taking into consideration only the data of specified joints, the ranking is constructed as shown in Table 1. Very similar high precision of identification corresponds to complete set of lower and upper body segments. According to expectations, the joints which are active during gait as for instance, hip, ankle, and shoulder are top ranked. Bit surprisingly classification based on knee joints is much less efficient It also partially explains why results obtained for the custom angles feature extraction, which considers knee movements instead of hip ones, were much worse than expected. High positions of spine segments are probably caused by a human posture abnormalities.

Table 1. Joint ranking.

Prec.	Segment
92,49%	UP
91,91%	LeftUpLeg
91,33%	DOWN
90,75%	Spine
88,44%	RightUpLeg
86,13%	Spine1
81,50%	RightFoot
79,77%	LeftShoulder

Prec.	Segment
78,61%	LeftFoot
76,88%	RightShoulder
75,72%	root
71,68%	LeftArm
70,52%	RightArm
61,85%	LeftLeg
57,80%	RightLeg
57,23%	Neck

Prec.	Segment
57,23%	LeftHand
56,07%	Head
53,18%	RightHand
52,60%	LeftForeArm
47,98%	RightForeArm
45,09%	LeftToeBase
38,15%	RightToeBase

Acknowledgement

This work was supported by project NN 516475740 from the Polish National Science Centre.

References

1. Janik Ł., Polański K, Wojciechowski K., Models and algorithms for human skeleton estimation form 3D marker trajectories, Machine Graphics and Visiuon 2013
2. Krzeszowski T., Kwolek B., Michalczuk A., Świtoński A., Josiński H., View independent human gait recognition using markerless 3d human motion capture, International Conference on Computer Vision and Graphics, Warszawa 2012
3. Plataniotis Lu, H, Venetsanopoulos K. N., , A. N.: MPCA: Multilinear Principal Component Analysis of Tensor Objects, IEEE Transactions on Neural Networks, Vol. 19, No. 1, 18-39, 2008
4. Świtoński A. Mucha, R, Danowski D, Mucha, M, Cieślar G., Wojciechowski K, Sieroń A., Human identification based on a kinematical data of a gait, Electrical Review, Volume 87, Issue 12 B, 2011, Pages 169-172
5. Świtoński A., Polański A., Wojciechowski K., Human identification based on gait paths, International Conference on Advanced Concepts for Intelligent Vision Systems,; Ghent, 2011
6. Zhang Z., Troje N.: View-independent person identification from human gait, Neurocomputing 69, 2005
7. Sakoe H., Chiba S., Dynamic programming algorithm optimization for spoken word recognition, IEEE Transaction on Acoustics, Speech and Signal Processing, 1978
8. Müller M., Röder T., Clausen M., Efficient content-based retrieval of motion capture data, ACM Transactions on Graphics, Volume 24, Issue 3, 2003
9. Świtoński A., Michalczuk A., Josiński H., Wojciechowski K.,, Dynamic Time warping in gait classification of motion capture data, International Conference on SIgnal Processing, Pattern Recognition and Applications, Venice, 2012
10. Świtoński A., Michalczuk A., Josiński H., Wojciechowski K., Gait human identification by Dynamic Time Warping applied to reduced video recordings datra, International Conference on Computer Science and Engineering, Kuala Lumpur, 2013
11. Cheng M., Ho M., Huang C., Gait analysis for human identification through manifold learning and HMM, Pattern recognition, vol. 41, 2008
12. Josiński H., Świtoński A., Michalczuk A., Kostrzewa D., Feature extraction and HMM-based classification of gait video sequences for the purpose of human identification, Visioan Based Systems for UAV Applications, 2013
13. Aqmar M.R.,Shinoda K.,Furui S., Efficient model training for HMM-based person identification by gait, 2012 4th Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, 2012
14. Boyd J. E., Little J., Biometric Gait Identification, Advanced Studies in Biometrics, Berlin, 2005
15. Witten I. H., Frank E., Hall M. A., Data mining - Practical Machine Learning Tools and Technique, Morgan Kaufman, Burlington, 2011

adfa, p. 10, 2011.