UDK 519.7

Y. N. AMIRGALIYEV, A. K. NUSSIPBEKOV

(Al-Farabi kazakh national university, Almatym Kazakhstan,
Institute of information and computing technologies, Almaty, Kazakhstan)

BAYESIAN NETWORKS IN POSE RECOGNITION

Annotation. The Bayesian networks are widely used statistical models for different kind of purposes. However one of its advantages is that it can be use in human posture recognition problems. The reason is that they can describe the kinematic constraints of human body. Particularly they can describe physical dependencies among human body parts. In this paper we are going to discuss how to use Bayesian tree networks in pose recognition. We use skeleton joints position and orientation values as feature vector. The CPD values of the network are parameterized by Conditional Linear Gaussians. The learning procedure implemented by Expectation Maximization (EM) algorithm. Our experiment results demonstrate that Bayesian Networks are very efficient in posture recognition tasks. Particularly we have got 93% of accuracy while classifying basic human poses.

Keywords: pose recognition, classification, Bayesian network.

Тірек сөздер: дене күйін тану, классификациалау, Байес желісі.

Ключевые слова: распознавание поз, классификация, Байесовские сети.

Introduction. Bayesian network (BN) is probabilistic graphical model where each node in this graph is a random variable where the edges of graph represent conditional dependencies among nodes. They are used to describe some information in specific uncertain domain. It has been used widely used in many studies. For example proposed a Dynamic Bayesian Network model for upper body tracking [1]. They construct a Bayesian Network to represent the human upper body structure.

The good thing about Bayesian networks is that they can take into account the kinematic constraints of the human body structure. In other words, they capture dependencies among human body parts that are physically connected. Daniel P. Huttenlocher *et. al.* in their work show that by using BN we can capture additional important information such as coordination of the limbs [2]. In their work they create a model for 2D and 3D human pose recovery. Researchers from University of California demonstrated how efficiently BN can be used for tracking moving people in long sequence without manual initialization [3].

In this work we demonstrate a 2D posture recognition using Bayesian network that will capture the kinematic structure of human body. The proposed system is demonstrated in figure 1. Experimental results demonstrate the efficiency of BN in pose recognition.

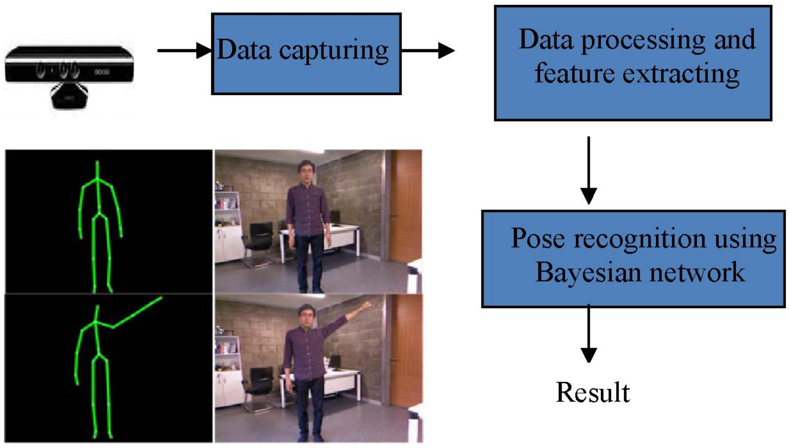
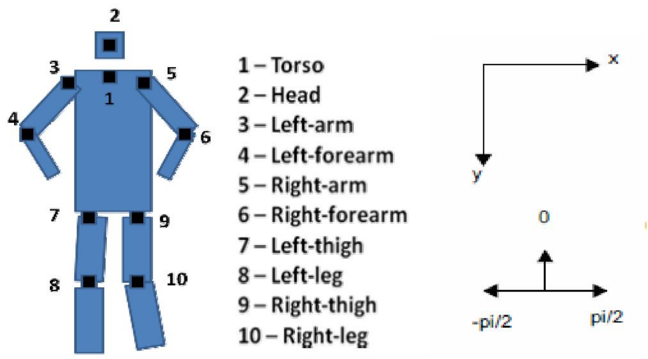


Figure 1 – System overview

Representation of body poses. As it was suggested in [6] we model human pose (see figure 2) using the feature set consisting of 3 elements: (y, x, α), where (y,x) is the position of specific human body part and α is the orientation of that part (inclination angle). There are 10 body parts such head, left-arm and etc. each having its coordinates and orientation.

Figure 2 – Human pose model



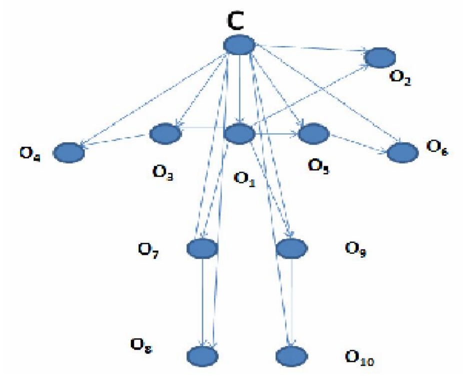


Figure 3 – Every body part has 2 parents: pose class C and physical parent $O_{(p)}$

Bayesian model. A Bayesian network (see figure 3) is a network, where each node represents human body part (e.g. Head, Right-arm). Every body part variable $\{O_i\}_{i=1}^{10}$ has continuous value $O_i = (\alpha_i, x_i, y_i)$. As it was said above, we use BN because they can capture the kinematic structure of human body. Every node has it's physical parent node and class node to which pose class it belongs to. Therefore all variables in BN can be parameterized by Conditional Linear Gaussian equation:

$$\alpha_i \mid O_{p(i)}, C = k \approx N(\beta_{ik}^1 \alpha_{p(i)} + \beta_{ik}^2 x_{p(i)} + \beta_{ik}^3 y_{p(i)} + \beta_{ik}^4, \sigma_{ik}^{\alpha 2}), \quad (1)$$

$$x_i \mid O_{p(i)}, C = k \approx N(\beta_{ik}^5 \alpha_{p(i)} + \beta_{ik}^6 x_{p(i)} + \beta_{ik}^7 y_{p(i)} + \beta_{ik}^8, \sigma_{ik}^{x 2}), \quad (2)$$

$$y_i \mid O_{p(i)}, C = k \approx N(\beta_{ik}^{9}\alpha_{p(i)} + \beta_{ik}^{10}x_{p(i)} + \beta_{ik}^{11}y_{p(i)} + \beta_{ik}^{12}, \sigma_{ik}^{y2}), \quad (3)$$

where i = 1,2,…,10 are body part indexes, k is the pose class number, $p(i)$ denotes the parent of $i'th$ node.

Model learning. A model can be learned using Expectation Maximization algorithm. Roman Filipovich *et. al.* [4] in their study demonstrated the efficiency of using EM algorithm in estimating the parameters of Bayesian trees. The idea of EM algorithm is iterating between doing expectation by generating approximate current assignments for hidden variables (E step) and updating parameters of the model by using these assignments (M-step). These iterations last until reaching local maxima.

E-step: for each pose calculating conditional probabilities which calculated from computing joint probability of the class and pose:

$$P(C = k, O_1, ..., O_{10}) = P(C = k)\prod_{i=1}^{10} P(O_i \mid C = k, O_{pa(i)}), \quad (4)$$

then we compute conditional class probability:

$$P(C = k \mid O_1, ..., O_{10}), \quad (5)$$

M-step: for each pose class we fit CLG parameters.

Experimental results. In order to evaluate our method we collected the database consisting of 300 postures of 10 people related to 5 posture classes. About 70% of our database was used for learning a model and remain 30% for evaluation purpose. In order to capture poses we used Microsoft Kinect camera. Kinect camera is able to capture human skeleton images using its infrared technology. The program was written on. NET C# language. The captured poses then learned and evaluated by program written on Octave (Matlab). Finally we have got the following experiment results demonstrated in table.

Pose recognition accuracy results

	Pose	Accuracy
1	Standing	95%
2	Sitting	89%
3	Star pose	97%
4	Waist bow	92%
5	Kicking	91%
	Average accuracy	**93%**

Accuracy is calculated from following equation:

$$Accuracy = (N_c \times N_T^{-1}) \times 100, \quad (6)$$

where N_c – number of right poses, N_T – total number of poses.

The demonstrated above results are high enough. The average recognition accuracy is 93% which can be comparable to the state of the art works. For example [5] have got also the same result using their method based on multilinear analysis.

Conclusion. In this paper we wanted to demonstrate that Bayesian Networks are very efficient tool in human pose recognition problems. And this is because they can handle the constraints among human body parts. CPD values of BN nodes can be parameterized by conditional linear Gaussians of body part position and coordination. Model learning can be implemented by EM algorithm. The demonstrated method is comparable with state of the art works in this domain. Pose recognition may be used for other recognition tasks like gesture recognition, surveillance problems and etc.

REFERENCE

1 L. Zhang, J. Chen, Z. Zeng, and Q. Ji, "2D and 3D upper body tracking with one framework," *2008 19th International Conference on Pattern Recognition*, pp. 1–4, Dec. 2008.

2 D. P. Huttenlocher, "Beyond trees: common-factor models for 2D human pose recovery," *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pp. 470–477 Vol. 1, 2005.

3 D. Ramanan and D. a. Forsyth, "Finding and tracking people from the bottom up," *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, vol. 2, pp. II–467–II–474.

4 R. Filipovych and E. Ribeiro, "Discovering Constrained Substructures in Bayesian Trees Using the E . M . Algorithm," 2008.

5 B. Peng and G. Qian, "Binocular dance pose recognition and body orientation estimation via multilinear analysis," *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, Jun. 2008.
6 Probabilistic Graphical Models, available at https://www.coursera.org/course/pgm, December 8, 2013

Резюме

Е. Н. Әмірғалиев, А. К.Нүсіпбеков

(Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан)

БАЙЕСТІК ЖҮЙЕЛЕРМЕН ДЕНЕ КҮЙІН ТАНУ

Мақалада Байестік желілер арқылы дене күйін тану әдісі ұсынылған. Байестік желілер дене бөліктерінің физикалық байланыстарын ескеруге қабылетті. Таратылудың шарты ықтималдылықтарының мәндері дене бөліктерінің көлбеу бұрыштары мен кординаталарының ескерілуімен Гаусс таратулары арқылы желінің түйіндерінде есептелінеді. Модельдің оқытылуы максималды ұқсастық негізінде ЕМ алгоритмі арқылы іске асырылады.

Тірек сөздер: дене күйін тану, классификациалау, Байес желісі.

Резюме

Е. Н. Амиргалиев, А. К. Нусипбеков

(Казахский национальный университет им. аль-Фараби, Алматы, Казахстан,
Институт информационных и вычислительных технологии, Алматы, Казахстан)

БАЙЕСОВСКИЕ СЕТИ В РАСПОЗНАВАНИИ ПОЗ

В работе предложен метод использования Байесовских сетей для распознавании поз. Байесовские сети способны учитывать физическую взаимосвязь между частями тела. Значения условных вероятностных распределений на узлах сети расчитываются распределением Гаусса с учетом координат и угла наклона частей тела. Обучение модели происходит на основе максимального правдоподобия с ЕМ алгоритмом.

Ключевые слова: распознавание поз, классификация, Байесовские сети.

Поступила 14.01.2014 г.