Visualization System as a Support Tool for Mapping Human Motion Patterns to Humanoid Robot

[1]*Martina ČIEFOVÁ,* [2]*Mária VIRČÍKOVÁ,* [3]*Martin ČERTICKÝ*

Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

[1]tina.ciefova@gmail.com, [2]maria.vircikova@tuke.sk, [3]martin.certicky@tuke.sk

***Abstract* — There is a need for the field of Human–robot interaction to be natural for its users to accept a robot as a cooperative partner. Therefore, it is necessary to develop a system providing human–like robot behavior. This idea brings a lot of problems due to the human communication style complexity and the surrounding environment. A combination of social learning and human–robot interaction can represent a solution for this. Visualization system is introduced in this paper. This system was designed as a part of the cloud-based technology for the human-robot interaction. Its aim is to support robot learning by demonstration involved in mentioned technology. Thanks to that, new motion is learned by robot without requirement of additional programming.**

***Keywords* — Cloud, human motion data, Microsoft Kinect, robot learning by demonstration, visualization**

I. Introduction

Development of the robot which is able to operate in a natural human environment is a new field of study in robotics. The aim is to create an equal robot partner of human, which can work by his side and immediately respond to his requests. An acceptance of robot by human partner is a basic task in human-robot interaction. The robot has to be a natural part of human life. It means, that interaction must be acceptable, fast and satisfying for a human user.

However, this brings lots of basic problems needed to be solved. The natural human environment is specific to every user and every human is a unique person. It is impossible to develop a universal interaction system fit to everyone. So that, to create a robot able to obtain information directly from the environment itself is a suitable solution. This can be achieved through interaction with a human partner, what is the fundamental of social robotics. Furthermore, robot could learn a specific task by observation of the human partner. In this way, if a new motion has to be performed by robot, it isn't necessary to program it manually. Motion is performed by demonstrator and consequently it is learned by robot through observation.

Center of Intelligent Technology (CIT) [1] in cooperation with the Kitakyushu Institute of Technology (Kyutech) [2] are developing a framework [3] for learning new motions and gestures based on a human-robot interaction and Cloud. The main idea of the project is to create public cloud-based database of motion samples (Fig. 1). The database is available for people from the whole world, so they can get motion samples as well as add new samples to the database. The framework employing advantages of Microsoft Azure [4] – a hybrid cloud platform, which offers various integrated services.

Motion samples are achieved through web-based application using Microsoft Kinect. The motion database is loaded in the part of web-based application called Explore and motions are displayed as buttons with the names of the motions. Demonstrator can't see how the sample was recorded and decides, if the motion was recorded correctly. Also, he doesn't know, which motion is represented with a specific name. Visualization system was designed to solve this problem. This system is described in this paper.

The remainder of this paper is organized as follows. In Section II there is state of art. Section III describes mentioned framework and proposed visualization system as a support tool for this framework. Section IV contains experiments and their results. Section V is formed with future work and conclusion.

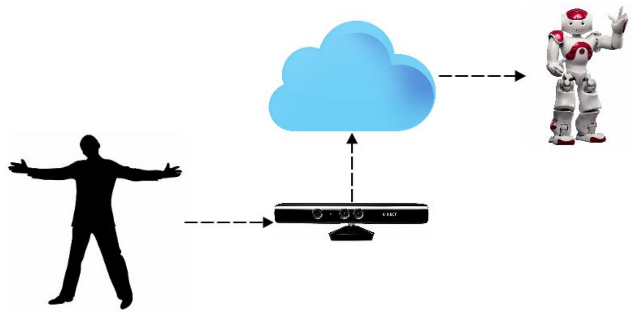

Fig. 1. Scheme of framework without visualization

II. STATE OF ART

A. Robot learning from demonstration, human – robot interaction

Due to the idea of Human-robot Interaction forum [5], human-robot interaction is a field of study focused on understanding, design and robot evaluation set to use by or with people. Interaction demands the presence of a certain type of communication between robot and human. This communication can be remote, immediate and social [6]. Social interaction is an area including social, emotional and cognitive interaction factors. It is a process in which people act and react to impulse from other people in their environment. This process represents dynamic social activities sequence between individuals. They modify their actions and reactions due to partners' behavior in interaction. Robot and human interact as peer-to-peer in case of social human-robot interaction [7][8].

Social robot research bears on social learning principles because a lot of reasons and advantages emerged from social learning. The system able to adapt to a specific environment and user condition could be created thanks to this type of learning [9]. The robot is competent to achieve new knowledge and information automatically without the need of inventor additional program. Social learning includes learning how to perform special task or behavior and learning when realize this task or behavior based on the social environment impulse [10].

Meltzoff (1996) expressed a hypothesis about children capability to transform their observation of other people's actions to creating their own actions. This represents the fundamentals of learning the similarities between the person themselves and the others and fundamentals of learning observable behavior and mental states created by this behavior. Learning by observation competence and ability of imitation observed actions is the base of appropriate social behavior development and of developing the capability of logic consideration about thoughts, intents and desires of another person. The ideas mentioned above are forming robot learning from demonstration postulates. Robot learning in this case is represented by performing a specific task in addition to generalizing from several perceptions of the demonstrator. The demonstrator performs this task in the role of teacher [11][12].

B. Microsoft Kinect Skeletal Tracking

Microsoft Kinect represents a cheap and easy tool for detecting people and action tracking and analysis. Its great advantage is the ability to isolate the human body from environmental noise. It is a combination of video and infrared measuring, which creates a set of joints with their 3D coordination. Human video frame can be reduced as skeleton [13]. The skeleton is created with the set of 20 joints.

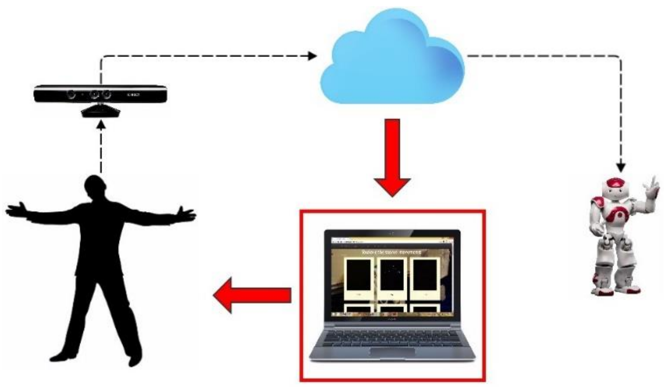

Fig. 2. Scheme of framework with visualization

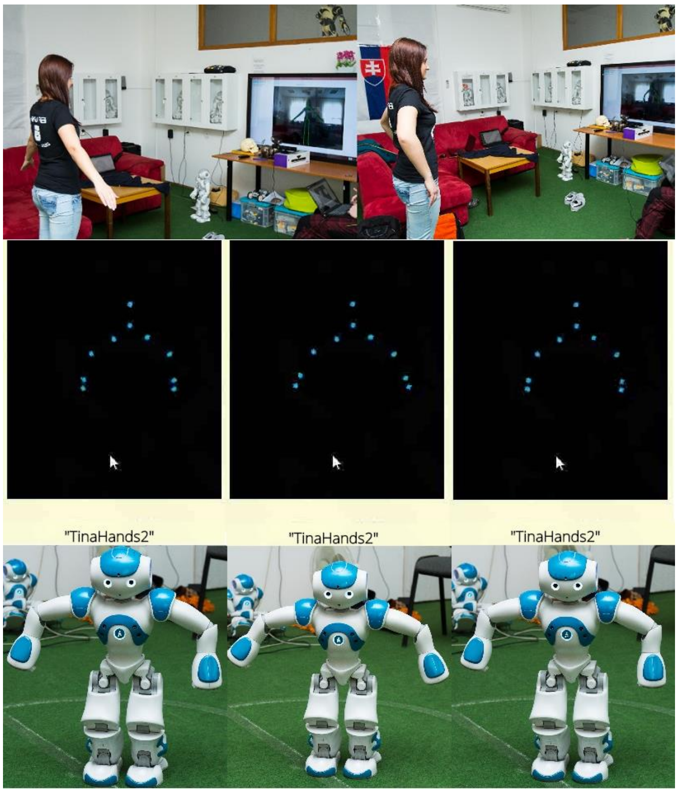

Fig. 3 Rendering of new motion sample, its visualization, the motion performed by Nao

C. Visualization system survey

Correctness of motion synthesis and analysis systems are mostly evaluated visually. Visualization in [14] represents the engine for displaying the state of systematic learning for the demonstrator. This means that the demonstrator knows when learning phase is ended to perform a new task. Work [15] describes several problems to emphasize necessary of visualization in this field of study. In [16] Human Motion Database creation is described. This database represents an ambition to create a base camp for motion recognize approach. The skeleton is represented by a tree hierarchy of 21 bones included a root. Each bone has its own coordination system calculated due to direct parent. Work [17] contains a proposition of web-based 3D animated system able to visualize real-time ballet scene. The skeleton is also represented by hierarchical structure and figure description is segmented into 5 parts – pas, motion, right hand, left hand, and head. System architecture consists of virtual environment based on VRML and Java applet running in a web browser. The user can choose from 3 human model representation types – skeleton, see-through and tutu. Motion history image (MHI) is another human motion visualization type. It represents a visual motion display during the time. The silhouette is obtained from the video sequence and compressed into a gray-scale image. Relevant motion information is conserved. MHI indicates motion progress by using of intensity of each pixel due to time, which present density function at given space [18].

Gunnar Johansson [19] approved that several light points which are strategically placed on the human body are immediately organized to create a coherent human body perception. To create an effect of human being is sufficient less than isolated points with display time 200 ms. Approach [20] utilizes a set of 15 virtual marks with their 3D coordinates. Some experiments were oriented to achieve various conclusions of perception of different information processing style, where information is obtained from biological motion perception. Here, web-based application was designed to display walking samples using light points. Visualization contains from 15 white light points representing derived virtual marks on the black background.

III. Visualization System

A. Cloud-based technology for Human-Robot Interaction

The proposed visualization system was purposed as a part of the framework available on [3] (Fig. 2). Web-based application of this framework was created using ASP.NET technology. It can be used in combination with Kinect sensor to record a new motion and add this motion to the database and also get the chosen motion from the database to perform it by Nao, a robot humanoid. The web-based application is divided into several sub-parts. One of them, Explore, is aimed to display the whole motion database. Design of Explore was created from buttons with

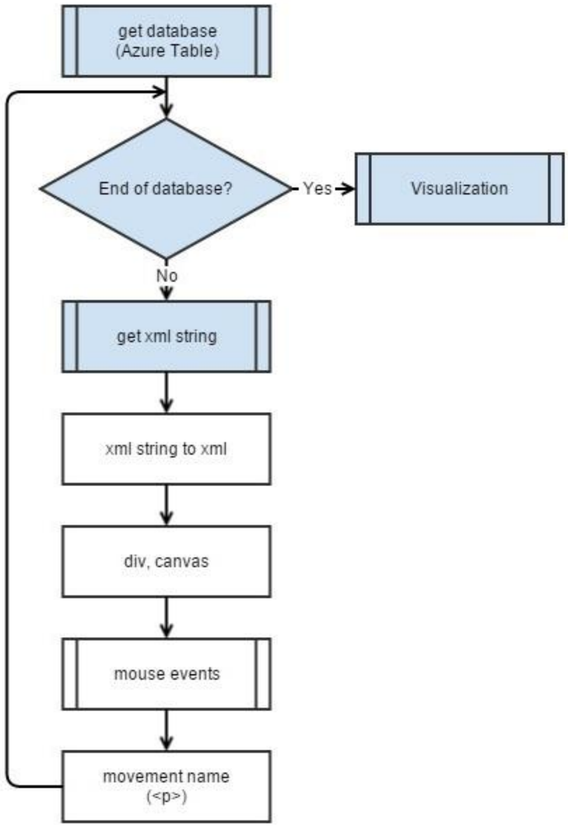

Fig. 4 Algorithm of main system

motion name firstly.

The visualization system described in this paper displays recorded motions from cloud-based database and offers to demonstrator a support tool for future experiments. The database is loaded in masonry style view. Each brick represents one motion from the database and consists of a window for display visualization and a motion name. Visualization starts after moving the mouse over the window and it is composed of light points standing for individual skeletal joints (Fig. 3). Visualization is stopped after going out from the window.

B. Visualization system proposition

The style of visualization was conditioned by representation of motion samples. This is composed with 3D coordination of 15 skeletal joints in each frame of recorded sequences. Chosen skeletal joints represent the upper part of the body. Each motion in cloud-based database is represented by xml file format, the nodes of which are hierarchically organized. System design

was inspired by [20] visualization style and decided to use light points to visualize individual skeletal joints. JavaScript language and library three.js were used, which represent a simple solution for creation of 3D animated computer graphics by using of WebGL technology. Renderer, scene and camera are necessary to display the visualization on the page. WebGL renderer was chosen which uses canvas to draw a scene and PerspectiveCamera.

The parameters of this components were customized according to several tests. The mentioned framework is created in Visual Studio by using of ASP.NET technology, which offers classes and functions to work with Microsoft Azure. Cross connection between languages is ensured by Razor. For motion visualization is created one renderer with a specific canvas, one scene and one camera. Renderer canvas is redrawn into brick window to display visualization.

C. Logic of proposed system

The whole database is loaded after calling sub-part Explore. Then renderer, scene and the camera are initialized. This is followed by the cycle, in which each movement in the database is passed and a string containing xml file with information about the joint is loaded into a specific variable. Brick represents a specific motion created with brick window for displaying visualization and motion name. Also, brick mouse events are generated (Fig. 4).

D. System requirements

WebGL library used in proposed system is supported by the most modern browsers and it doesn't require any plug-ins. It is integrated into all web standards of the browser, allowing GPU accelerated usage. WebGL programs are designed with usage of JavaScript code and shader code processed by the computer graphic processor.

WebGL browser support [21]:

1. Internet Explorer - partially supports in Internet Explorer 11
2. Google Chrome - enable on all platform with capable graphics card with upstaged drivers since version 9
3. Mozilla Firefox - enable on all platform with capable graphics card with upstaged drivers since version 4.0
4. Opera - implemented in Opera 11 and Opera 12, disabled by default

IV. Experiment

A questionnaire was used to evaluate system function and effectiveness. This questionnaire was responded by people working at CIT. This target group was chosen because of their practical experiences with the framework. That means their answers should be the most directive.

A. Experiment proposition

The proposed questionnaire is divided into 2 parts. Questions in the first part discuss about system functionality, type of chosen browser, respondents' suggestions and reminders. The second part contains various visualization layouts and types of human body displaying. They are rated by respondents.

The questionnaire was created by Google Docs and it is available on [22].

B. Experiment results

The style of visualization was conditioned by representation of motion samples. This is composed with 3D coordination of 15 skeletal joints in each frame of the recorded sequence. Chosen skeletal joints represent the upper part of the body.

The questionnaire was completed by 6 people. All of them recorded movement to the database before visualization testing according to their answers. During visualization testing, 3 of them were picking out movement from the database and 1 was recording movement. Most of them, 4 concretely, decided to use Google Chrome, Mozilla Firefox was used once. One of the respondents using Internet Explorer confirms that visualization doesn't work correctly on this browser.

Visualization works without problems in one case, in other cases several problems were revealed. Responders annotated that some of the movements weren't centered, visualization of some movements was strange and masonry layout wasn't displaying correctly in case of high definition. Their suggestions to improve the system discussed to add an option to display the duration of movement, to stop movement or to choose from more types of human body visualization (avatar). Suggestions also discuss about loading movements in sequence, not loading the whole database at once.

Layout with red point represent skeletal joint on white background was chosen as the most informative. The system's effectiveness rate was 4.17 points on average, which approves that visualization is a useful tool for robot learning of new motion.

V. Future work

Future work could focus on the solution of respondents' reminders. Masonry layout could be replaced by another way of database displaying. Also, the style of a human body representation could be placed by an avatar or single units connected. Next work could be about adding functionality to label individual motion by emotion tags, such as in [23]. Set of tags could be represented by tag cloud, what is a type of visual representation of text data according to their rating. This tag cloud could contain the set of several emotion tags different by their context. The user can label given motion after motion visualization according to emotion which, given motion involved with him.

The meaning of social learning in this framework could be broadened by this way.

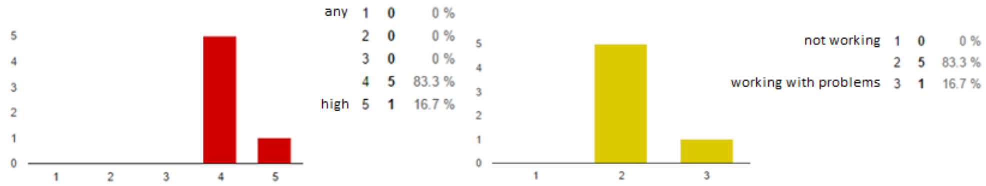

Fig. 5 Questionnaire results about system effectiveness (left) and system functionality (right)

VI. Conclusion

This paper was focused on social learning in human-robot interaction. A framework developed by CIT and Kyutech was described. This framework represents an ambition to create a public cloud-database containing human motion samples. Samples are achieved by Kinect Sensor and web-based application of mentioned framework. Missing tool for recorded motion evaluation is solved by designed visualization system described in this paper. Visualization is available in web-based application sub-site, Explore.

The questionnaire described in this paper serves as an evaluation of system functionality and effectiveness. It revealed some system deficiency that could be subject of future work. It also affirmed visualization system's potential in this field of study.

References

[1] Center for Intelligent Technology, 'Center for Intelligent Technology', 2015. [Online]. Available: http://www.ai-cit.sk/. [Accessed: 19-May-2015].

[2] Kyushu Institute of Technology, 'Kyushu Institute of Technology.' [Online]. Available: https://www.kyutech.ac.jp/english/. [Accessed: 19-May-2015].

[3] CIT and Kyutech, 'Cloud-based technology for Human-Robot Interaction.' [Online]. Available: http://jpsk.cloudapp.net:8080. [Accessed: 19-May-2015].

[4] Microsoft, 'Microsoft Azure.' [Online]. Available: https://account.windowsazure.com/Home/Index. [Accessed: 27-May-2015].

[5] Human-Robot Interaction, 'Human - Robot Interaction.' [Online]. Available: http://humanrobotinteraction.org/. [Accessed: 10-Dec-2014].

[6] M. Virčíková, 'Machine Empathy: Towards Artificial Emotional Intelligence with Active Personalization in Social Human-Robot Interaction', Technická univerzita v Košiciach, 2014.

[7] Study.com, 'Study.com.' [Online]. Available: http://education-portal.com/academy/lesson/social-interactions-definition-types-quiz.html. [Accessed: 13-Feb-2015].

[8] Boundless, 'Understanding Social Interaction', 03-Jul-2014. [Online]. Available: https://www.boundless.com/sociology/textbooks/boundless-sociology-textbook/social-interaction-5/understanding-social-interaction-50/understanding-social-interaction-314-5912/. [Accessed: 13-Feb-2015].

[9] C. Breazeal and B. Scassellati, 'Robots that imitate humans', *Trends in Cognitive Sciences*, vol. 6, no. 11. pp. 481–487, 2002.

[10] M. J. Mataric, 'Learning to behave socially', in *Third international conference on simulation of adaptive behavior*, 1994, pp. 453–462.

[11] A. Billard and D. Grollman, 'Robot learning by demonstration', *Scholarpedia*, vol. 8, no. 12, p. 3824, Dec. 2013.

[12] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumberg, 'Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots', *Artif. Life*, vol. 11, no. 1–2, pp. 31–62, 2005.

[13] L. Xia, C.-C. Chen, and J. K. Aggarwal, 'Human detection using depth information by Kinect', *CVPR 2011 Work.*, pp. 15–22, 2011.

[14] D. Kulic, C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura, 'Incremental learning of full body motion primitives and their sequencing through human motion observation', *The International Journal of Robotics Research*, vol. 31, no. 3. pp. 330–345, 2012.

[15] J. Bernard, N. Wilhelm, B. Krüger, T. May, T. Schreck, and J. Kohlhammer, 'MotionExplorer: exploratory search in human motion capture data based on hierarchical aggregation.', *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2257–66, 2013.

[16] G. Guerra-Filho and A. Biswas, 'The human motion database: A cognitive and parametric sampling of human motion', in *Image and Vision Computing*, 2012, vol. 30, no. 3, pp. 251–261.

[17] A. Soga, M. Endo, and T. Yasuda, 'Motion description and composing system for classic ballet animation on the web', in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 2001, pp. 134–139.

[18] M. A. R. Ahad, J. K. Tan, H. Kim, and S. Ishikawa, 'Motion history image: Its variants and applications', *Machine Vision and Applications*, vol. 23, no. 2. pp. 255–281, 2012.

[19] G. Johansson, 'Visual perception of biological motion and a model for its analysis', *Perception & Psychophysics*, vol. 14, no. 2. pp. 201–211, 1973.

[20] N. F. Troje, 'Retrieving information from human movement patterns', in *Understanding events: How humans see, represent, and act on events*, 2008, pp. 308–334.

[21] 'WebGL - 3D Canvas graphics.' [Online]. Available: http://caniuse.com/#feat=webgl. [Accessed: 22-May-2015].

[22] 'No Title.' [Online]. Available: https://docs.google.com/forms/d/1jeVTei1DOIJA-KSEjzMFMalVqhzww4B9jHy_XtVqTXs/viewform?usp=send_form.

[23] R. Hasebe, R. Kouda, K. Ohnishi, and M. Munetomo, 'Human-based genetic algorithm for facilitating practical use of data in the internet', in *Soft Computing and Intelligent Systems (SCIS), 2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS), 15th International Symposium on*, 2014, pp. 1327–1332.