HUMANOID ROBOT MARKO - AN ASSISTANT IN THERAPY FOR CHILDREN

Branislav Borovac, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Vlado Delić, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Milan Gnjatović, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu, Megatrend Univerzitet
Branko Karan, Srpska Akademija Nauka i Umetnosti SANU
Mirko Raković, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Marko Penčić, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Srđan Savić, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Milutin Nikolić, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Jovica Tasevski, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu
Dragiša Mišković, Fakultet Tehničkih Nauka Univerziteta u Novom Sadu

Abstract: *This paper reports on work in progress towards development of a robot to be used as assistive technology in treatment of children with developmental disorders (cerebral palsy). This work integrates two activities. The first one is mechanical device design (humanoid robot) of sufficient capabilities for demonstration of therapeutical exercises for habilitation of gross and fine motor functions and for acquiring spatial relationships. The second one is design of appropriate communication capabilities of the robot. The basic therapeutical role of the robot is to motivate children to practice therapy harder and longer. To achieve this, robot must fulfil two requirements: it must have appropriate appearance to be able to establish affective attachment of the child to the robot, and must be able to communicate with children verbally (speech recognition and synthesis,) and non-verbally (facial expressions, gestures...). Thus, conversational abilities are unavoidable and among the most important capabilities. In short, robot should be able to manage three-party natural language conversation – between the child, the therapist and the robot – in clinical settings.*

Keywords: *Humanoid conversational robot, adaptive behavior, context awareness, cerebral palsy, robot-assisted therapy*

INTRODUCTION

The clinical descriptive term “cerebral palsy'” relates to permanent developmental disorders that occurred early in human biological development or in early childhood. These developmental disorders primarily refer to conditions of abnormal gross and fine motor functioning and organization that can lead to secondary problems with behavior and participation in society [1]. It also may happen that children do not have proper spatial relationships and notion of them – the child is not able to show where his head is, what is his left or right arm or leg, etc. Treatment is usually symptomatic and focuses on helping the person to develop as many motor skills as possible or to learn how to compensate for the lack of them.

Therapy of motor skill disorders in children is long-termed and should start as early as possible. Since cerebral palsy patients significantly differ among themselves, therapy should be adapted to a particular child. However, therapy may also be tiresome, uncomfortable or even painful for the child. Therefore, it is essential that the child is motivated to undergo therapy, which is a challenging therapeutic task.

This paper describes a work in progress towards a prototypical robotic system intended to be used as an assistive tool in therapy of children with cerebral palsy. Such a system must be able to demonstrate requested exercises, its appearance should be attractive to children, and it should be able to communicate (verbal and non-verbal) with patients and therapists. The expected benefit of the described robot is to support the therapist to conduct specific therapeutic exercises and to contribute to establish affective attachment of the child to conduct exercises.

**Branislav Borovac, Fakultet tehničkih nauka Univerziteta u Novom Sadu, Trg Dositeja Obradovića 6, Novi Sad, Srbija*
borovac@uns.ac.rs

CONVENTIONAL THERAPEUTIC PRACTICES AND ROBOT DESIGN

This robot is intended to be used in therapy for habilitation of gross motor functions and acquiring spatial relationships, and habilitation of fine motor functions. This robot is also planned to participate speech therapy, but this point will not be covered here.

The first questions to be answered are appearance and structure of the robot. According to testing performed by psychologists [2], kids prefer blue color and round shape. We accepted these suggestions.

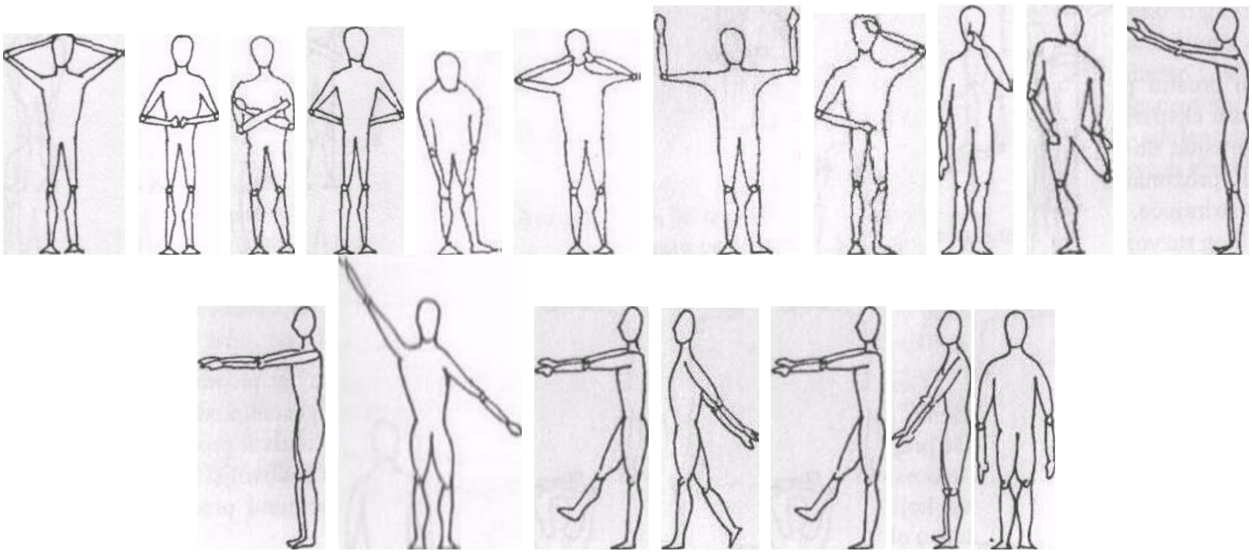

Figure 1: Examples of gross motion exercises used in conventional therapy

The second basic decision is structure of the robot to be able to demonstrate exercises. Examples of gross motion exercises used in conventional therapy are given in Figure 1. It is obvious that robot must have body, arms, legs and head (Figure 2). Motions expected to be performed by arms are complex and robot's arm should have mobility similar to human's arm. However, movements required by legs are just abduction, adduction and knee bending. Having in mind that it is not justified to design biped walking robot just for therapeutic purposes[1] we decided to design robot which have arms and legs, but sitting on a platform. Thus, robot will be able to demonstrate most of requested exercises[2], but, dynamic balance will not be endangered.

For fine motion exercises, the child is sitting at table and in front of him is a set of wooden objects of different shapes and colors. The child is asked by the therapist to collect or move objects having certain characteristics (for example, blue triangles). In the case that the child cannot fulfill the task, the therapist will ask the robot to do the same. When the robot moves the object, the child can try to repeat this[3].

The robot, also, has vision system, implemented by two cameras placed at his eyes[4]. The vision has two purposes: to detect and recognize objects and their positions on the table for fine motion exercises, and to detect where the child is. Child's position is important for its safety: if any part of the child's body is too close to the robot, robot have to stop its motion immediately and switch to compliant mode. Such a situation may happen if the robot was asked to move object on table, but suddenly the child decides to move the same object at same time. The Robot should allow the

---

[1] The robot has to perform just a motion by legs and arms. Biped locomotion is too complex task (and possibly too dangerous for child in case of biped fall) to be performed as a side effect during exercise demonstration.

[2] For therapies using this robot, the list of exercises should be reviewed and updated.

[3] To see an example of this exercise, please see the video clip (https://www.youtube.com/watch?v=e6LpaQ3Es3A) that we made before the robot we are designing was available. Industrial robot ABB Irb 140 was used instead.

[4] The second vision system is based on a 3D sensor [3]. It is designed to detect the child's movements and postures in therapy of gross motor skill disorders, and extract therapy-relevant measures of the child's nonverbal behavior.

child to have priority. Another such situation is if child, before the session start, want to approach the robot close and touch him. In spite that it was not supposed that the child and robot will come into direct contact, such a situation is possible and child's safety is of highest priority.

Because the robot has to encourage the child, it must be able to express (positive) emotions. There are four ways to do this: using eyes[5], by using eyebrows (rising or inclining), by voice modulation, and by different lightening of the transparent ears.

Having in mind that the most natural way of communication for humans is speech, verbal communication is unavoidable during therapy. However, it is not possible to have formalized speech (using predefined set of sentences in strictly formal way) with children. It regularly happens that child's sentences are not grammatically correct and semantically complete. System for speech recognition and synthesis should be able to overcome this.

To summarize, in our approach, a therapeutic scenario includes three-party speech-based interaction between the child, the therapist and the robotic system. In the two dedicated therapeutic exercises, the therapist verbally instructs the child to move different parts of his body (e.g. to perform gross motor movements or to maintain a posture) or to manipulate an objects placed on the table (e.g., to perform fine motor movements), cf. [4]. The first role of the robot is to support the child to perform the therapist's instructions. If the child does not understand the therapist's dialogue act or cannot perform the instructed movement, the robot may (in accordance with its dialogue strategy) demonstrate the instructed movement. Thus, the robot has an anthropomorphic appearance, cf. Figure 2. It has a head, two arms and two legs. The robot is mounted on a chassis to make the legs available for the therapeutic purposes (e.g., demonstration of gross motor movements) and to minimize the risk of overturning. However, it is not sufficient just to enable the robot to appropriately demonstrate gross motor movements and manipulate objects positioned on a table. It is also important that its parts are easily perceivable by, and appealing to, the child.

This leads us to the second role of the robot – to establish affective attachment of the child to the robot and to trigger social interaction between the child and the therapist. Two important motivational factors relate to visual appearance of the robot and its dialogue behavior.

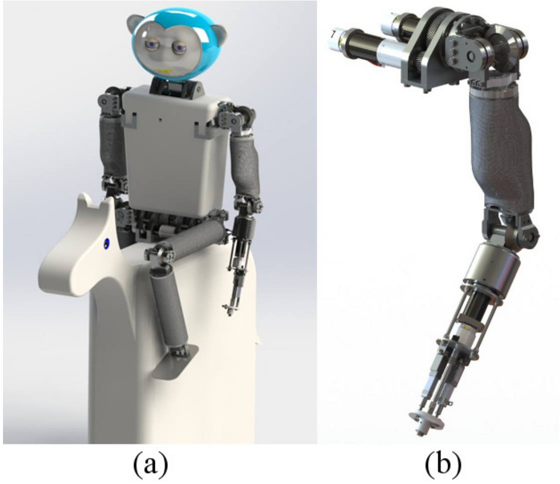

Figure 2: Models of (a) the robot, and (b) the arm.

Especially important for the visual appearance of the robot is its head, cf. Figure 3. It is designed to be appealing to the child, and can mimic facial expressions of the basic emotional states. In addition, the integrated vision system enables the robot's head to direct its eyes towards, and to track, the interlocutor. Another motivation factor relates to the robot's capacity to engage in and manage natural language dialogue. For example, the therapist may verbally instruct the robot to demonstrate a movement to the child, or the robot may take initiative and utter a comment to

---

[5] The video https://www.youtube.com/watch?v=vdQmshozC6Y shows the first attempt of using eyes and lightening of ears to express some basic emotions of the robot MARKO.

encourage and motivate the child. To achieve this advanced behavior, the robot is integrated with a conversational agent.

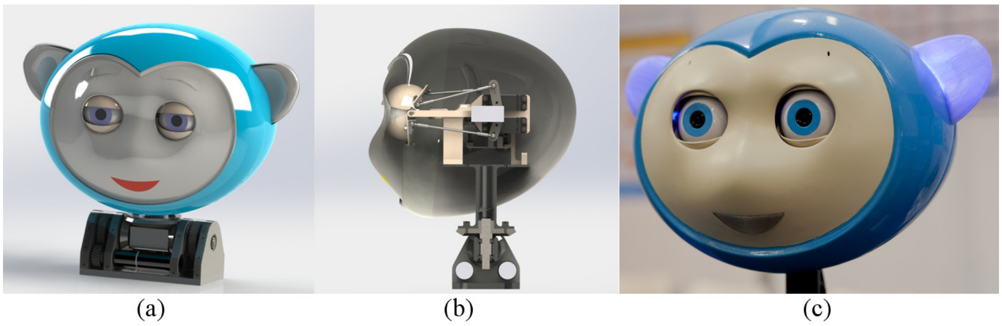

Figure 3: (a) A model of the robot's head; (b) a model of the first vision system; (c) the actual robot's head captured while mimicking the facial expression of surprise. As part of future work, eyebrows will be added to the robot's head.

COGNITIVE SUBSYSTEM WITH AUDIO INTERFACE

The cognitive subsystem implements the functionality of adaptive dialogue management and natural language processing. This module is based on attentional state model used in the system introduced in our previous work [5, 6]. The model integrates neurocognitive understanding of the focus of attention in computational and corpora linguistic. This paper expands upon previous work on adaptive multimodal interaction and linguistic encoding of motion events [7]. Previously we introduced a spoken natural language dialogue system that manages the interaction between the user and the anthropomorphic robotic arm. In that work it is allowed to the user to utter instructions of different syntactic form that relate to the object manipulation. The task of the system is to interpret the user‘s commands and to perform the requested operations. When needed, the system applies an adaptive dialogue strategy and initiates a conversation in order to help the user to specify the required information.

The audio subsystem, as a part of the cognitive system, implements the functionalities of automatic speech recognition and text-to-speech synthesis (ASR and TTS) for the Serbian language, developed at the Faculty of Technical Sciences, University of Novi Sad, Serbia [6].

The Text-to-Speech engine is provided as a standalone SAPI 5 (Speech Application Programming Interface) speech synthesizer, called anReader. Having several male and female voices, it allows various options and settings related to pronunciation.

For speech recognition, we use engine based on Alfanum ASR, small and medium-sized vocabulary continuous recognizer. The system is speaker independent, phoneme-based with a number of HMM (Hidden Markov Model) states which is proportional to phoneme duration. A vocabulary (an arbitrary set of words) is defined by a grammar that describes all utterances that may be produced by the user in the given interaction domain. In order to make the system able to recognize specific words, a grammar must be defined. The grammars are defined using Backus-Naur form.

The engine is integrated into the ASR IP-based server. This enables a remote access to the server and speech engine relocation from the robot’s head.

Communication with ASR server is provided by specialized library intended for tasks that are essential for ASR part of human-robot interaction. The three main roles are:

- To collect samples from the microphone;
- To detect efficiently voice activity (VAD); and
- To initiate recognition and obtain result for cognitive system.

To determine active speech period is important task in application of speech processing, especially for unsupervised interaction. During the samples processing, adaptive thresholds are calculated using minimum and maximum values of a short-term energy estimation. Considering the discontinuous signal as a speech characteristic and presence of real environments noise, effective

VAD enables non-speech frame dropping to reduce the amount of insertion errors. The second task of our algorithm is silence detection (i.e., the moment when the user has already finished his utterance). That is a time when the library sends a recognition request to the server and waits for a response. The server's response comprises the recognized words and also numeric values, each of which represents the recognition reliability for a corresponding word. This information (i.e., the recognized command) is then forwarded to the cognitive subsystem for an interpretation.

The cognitive model is extended with a dialogue strategy. The dialogue strategy may be described as follows: if the user does not completely formulate a command (and the system cannot recover the omitted information from the context) or utters an ambivalent command (i.e., it can be assigned more than one interpretation in the given context), the system takes the initiative in the interaction and guides the user to complete the started command, or to resolve the ambiguity, by stating iterative questions. The dialogue system is adaptive, it takes the verbal and spatial contexts into account in order to adapt its dialogue behavior and to process spontaneously formulated user commands of different syntactic forms without explicit syntactic expectations.

CONCLUSION

This paper reported on work in progress towards a prototypical conversational robot intended to be used as an assistive tool in therapy for children with cerebral palsy. We integrate two lines of research: designing humanoid robots, and designing context-aware and adaptive conversational agents.

The presented approach introduces capability that the robotic assistive system is designed and developed to manage multi-party natural language conversation in clinical settings. The idea of conversational assistive tools *per se* is not novel – currently there are few social agents intended to assist individuals in health-related tasks [8, 9]. However, relatively little attention has been devoted to design and development of robotic assistive tools capable to manage conversation in clinical settings. This capacity of the robot to engage in dialogue is intended not only to support the therapist to conduct specific therapeutic exercises, but also to contribute establishing affective attachment of the child to the robot. The appealing appearance of the robot and its advanced dialogue competence is expected to significantly contribute to the child's perception of the robot as a “companion” with its own intentionality, rather than just as a therapeutic tool or a toy. This would be a strong and sustainable motivation for children. Future work will include testing the reported robotic system in realistic settings with healthy children and in real clinical settings with children with cerebral palsy.

ACKNOWLEDGMENTS

The presented study is performed as part of the projects “Design of Robots as Assistive Technology for the Treatment of Children with Developmental Disorders” (III44008) and “Development of Dialogue Systems for Serbian and Other South Slavic Language'” (TR32035), funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia. The responsibility for the content of this paper lies with the authors.

REFERENCES

1. Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., Dan, B., Jacobsson, B. A report: the definition and classification of cerebral palsy April 2006, Dev Med Child Neurol Suppl. 109, pp. 8-14 (2007)
2. Oros M., Nikolic M., Borovac B., Jerkovic I.: Childrens Preference of Appearance and Parents Attitudes Towards Assistive Robots, 2014 IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS 2014), November 18-20th 2014, Madrid, Spain
3. Karan, B., Golubović, S., Gnjatović, M.: Markerless Vision-Based Skeleton Tracking in Therapy of Gross Motor Skill Disorders in Children, in New Trends on Medical and Service Robotics: Challenges and Solutions, Springer monograph series Mechanisms and Machine

Science, vol. 20, pp. 79-93, Proc. of the 2nd International Workshop New Trends on Medical and Service Robotics, MESROB 2013, (2014)

4. Gnjatović, M.: Therapist-Centered Design of a Robot's Dialogue Behavior, Cognitive Computation, in press, available online (2014)
5. Milan Gnjatović: "Adaptive Dialogue Management in Human-Machine Interaction", Verlag Dr Hut, 2007
6. V. Delic, D. Pekar, R. Obradovic, N. Jakovljevic, D. Miskovic : "A Review of AlfaNum Continuous Automatic Speech Recognition System", XII international conference "Speech and Computer" (SPECOM'2007), Moscow, Russia, 15-18. october, 2007
7. Gnjatović M., Tasevski J., Nikolić M., Mišković D., B. Borovac., Delić V., "Linguistic encoding of motion events in robotic system" 6th PSU-UNS International Conference on Engineering and Technology ICET 2013., Novi Sad, Serbia, 2013.
8. Bickmore, T.W., Picard, R.W.: Establishing and maintaining long-term human–computer relationships. ACM Trans Comput Human Interact. 12(2):293–327 (2005)
9. Fasola, J., Matarić, M.J.: Using socially assistive human–robot interaction to motivate physical exercise for older adults. Proc IEEE. 100(8):2512–26 (2012)