

Markerless Vision-Based Skeleton Tracking in Therapy of Gross Motor Skill Disorders in Children

B. Karan¹, Š. Golubović², and M. Gnjatović³

¹*Institute of Technical Sciences of the Serbian Academy of Sciences and Arts, Serbia, e-mail: branko.karan@itn.sanu.ac.rs*

²*Faculty of Medicine, University of Novi Sad, Serbia, e-mail: spela@uns.ac.rs*

³*Faculty of Technical Sciences, University of Novi Sad, Serbia, e-mail: milangnjatovic@yahoo.com*

Abstract. This paper presents a research towards implementation of a computer vision system for markerless skeleton tracking in therapy of gross motor skill disorders in children suffering from mild cognitive impairment. The proposed system is based on a low-cost 3D sensor and a skeleton tracking software. The envisioned architecture is scalable in the sense that the system may be used as a stand-alone assistive tool for tracking the effects of therapy or it may be integrated with an advanced autonomous conversational agent to maintain the spatial attention of the child and to increase her motivation to undergo a long-term therapy.

Key words: 3D sensing, Skeleton tracking, Computer-aided therapy, Human-machine interaction.

1 Introduction

From the early stages of their development children have a desire to move and learn about the environment by exploring their physical surroundings. These experiences represent a foundation for future learning. Every learning process essentially includes the same stages: information acquisition, processing, association, cognition and response [4]. Learning and child development are based on development and regulation of basic mechanisms: sensor modulation and integration, motor planning and sequencing, visuospatial processing, auditory processing and emotional modulation. Children with motor development impairments can experience problems at any stage of the learning process or mechanism regulation, which affects their general functioning.

From a therapeutic point of view such a complex issue requires a range of interventions which should begin in early childhood. Early stimulations sometimes

need to start immediately after birth in order to efficiently harness high plasticity of the brain, thus encouraging reorganization of nerve structures, where intact areas of the brain take over the function of the affected areas [18]. When planning rehabilitation programs for children with motor disabilities, it seems wise to start the treatment by stimulating basic emotional reactions, perceptuo-motoric abilities and communication using movement and body as the basis for intervention. This is a specific type of rehabilitation which, through the use of structural motor activities, influences the development of functions and abilities that form the basis for learning. Movement therapy aims to remediate or improve incorrect motor activities, and develop compensatory abilities and functional cognitive mechanisms in order to improve the child's performance in the domains affected by impairment. Treatment strategies involve exercises through which children acquire new motor knowledge, develop motor skills [31] and develop perceptive and gnostic abilities, cognitive function and communication.

Treatment of motor and cognitive deficits requires long and painstaking work on the part of the child and his or her environment. To help direct inappropriate forms of behavior and movement into appropriate motor reactions we can use modern computer technology [6, 11, 25]. It provides opportunities for the child to be involved in treatment in a more interesting manner. Furthermore, it helps implement individual treatment plans with an emphasis on more frequent repetition of specific activities and movements, their correction, better record keeping and thus more precise evaluation of treatment outcomes.

Computer vision technologies may play an extremely important role in implementation of environments for computer-aided therapy. A notable example of a virtual reality rehabilitation system that uses the Microsoft Kinect to enable children with hemiparesis to improve their affected hand is reported in [35]. In the rehabilitation application, the patient controls the virtual avatar and attempts to touch a virtual object. A specialized algorithm is introduced to aid a Kinect sensor to recognize the paralyzed arm's limited field of movement, and to penalize the use of the unaffected arm [35]. In contrast to this research, we do not involve a prescribed motion-based video game tailored to a particular gross motor skill disorder, but propose an avatar-based conversational agent whose role is to facilitate the interaction between the child, the therapist and the animated (cf. Section 4).

The research presented in this paper aims at developing a scalable vision system that could provide assistance in different aspects of therapy of gross motor skills of pre-school to lower grade school children. From the technical aspect, the vision system could be used to make recordings of therapeutic sessions, and to extract clinically relevant quantitative measures related to the child movements, postures, etc. From the motivational aspect, it may provide feedback for a simple graphical system to present visual stimuli to engage children's attention. From the interactional aspect, the vision system may be incorporated with a more complex intelligent autonomous agent (e.g., a robotic system, a conversational agent, a computer game, etc.) whose role is either to be a child playmate or an independent actor in interaction between the child and the therapist.

2 Therapy practices

Motor impairments usually emerge relatively early in childhood. These impairments appear as a result of brain damage caused by a set of factors that affect brain maturation. The symptoms manifest as deviations in the area of sensory, motor, verbal, cognitive and social development of a child. Impaired motor abilities and skills range from mild motor disability to severely impaired movement, balance and coordination. Impaired motor functioning is characterized by slower execution of movement, difficulties in planning and organizing movement, difficulties in learning new motor skills and combining basic motor activities, as well as in motor response itself.

Children with developmental disabilities experience problems when identifying parts of the body, difficulties with spatial and time orientation, rhythm, lateral awareness. This means that rehabilitation programs should structure activities which would integrate motor and cognitive aspects of learning, thereby developing sensory differentiation, perception and thinking, enhancing memory, attention, concentration, and promoting acquisition of new concepts and communication [2, 23]. Exercises should be structured to meet the needs and abilities of every individual child; starting from the level that presents no difficulties for the child, then gradually changing and combining different activities. Initially, the activities should have a relaxing influence on the child, creating the basis for fluid execution of new movements, correction of uncoordinated movements, execution of more complex movements, awareness of the body parts involved and the body as a whole.

When doing the exercises there is no predefined method or order. By using simple movements, experiences acquired earlier are revised, reintegrated, and communication between child and therapist is established. This enhances motivation and encourages the child to participate. If the child experiences feelings of anxiety, insecurity or is emotionally immature, this can hinder fluid execution of movements and interfere with communication. In such cases, the therapist needs to create situations and develop activities which would ensure the child receives more attention, and allow the child to actively participate in the activities, thus enhancing communication with the therapist.

Communication encouraged when doing the exercises can be verbal or non-verbal. The former includes identifying movements and body parts, describing activities, spatial and time relationships, expression of emotions; while the latter is achieved through gesture, mimicking, movement and images. The first communication is usually established through modeling movement, starting with the command: "Now, do this". Thus initiated dialogue continues with the child imitating the therapist, and therapist following the child's lead. Some behavioral studies have shown that observation of movements can affect their execution, because it involves motor observation, motor imagery and movement execution [3, 9].

Through movement children acquire awareness of themselves, their body and other individuals, which forms a basis for personal development. Becoming aware of its body, the child becomes aware of its own existence in physical environment, acquires and consolidates spatiotemporal relationships, develops a need for socializing and communication [2, 23]. These are simple activities which are easily recognized, and in which certain parts of the body are identified, shown and touched. “Let’s put our hands on our head, now on our stomach, now on our back, and now on our legs.” This is followed by movement of larger joints, which enhances awareness of own body; while using proprioceptive and vestibular systems allows control of movement. Exercises include movement in all directions: crawling, walking, running, moving upper part of the body, moving the head, arms, legs, hands and feet in all directions – stretching, bending, rotating, lifting, side bending, throwing and catching. The most common exercises are focused on: body scheme perception, differential laterality, equalizing muscle tone, achieving independent movement, coordination of movement to rhythm, spatial orientation and relaxation activities.

Both group and individual exercises take place in non-competitive atmosphere and without assessment according to the plans developed after the children’s abilities have been evaluated. Group exercises promote communication between children, whereby they establish social contact and are motivated for group work. Individual exercises implement plans tailored to meet individual abilities and needs. Irrespective of the type, group or individual, activities are carefully planned and aim to meet set objectives. Objective can be directed at mastering certain motor skills, or at learning the characteristics of a certain object, space, etc.

3 Vision-based skeleton tracking

The vision system is expected to achieve efficient tracking of full-body motion. Having in mind the actual functionality to be supported, we distinguish 3 different levels of functions, in increasing order of complexity:

- (1) Recognition of movement/gestures: at this level, the vision system should be able to answer whether the patient has made a move pattern belonging to a specific class (or at least that he or she attempted to make such a move). This information is a basic input to automatic interactive agent. Besides, from the standpoint of overall therapy tracking, this information provides a basis to build an indicator of number of times that the specific exercise has been repeated.
- (2) Estimation of quality of the movement, expressed e.g. as a function of differences of specific motion parameters with respect to the pre-specified “golden criterion.” This information can be used to improve system interaction by providing different levels of interactive feedback (e.g., audio/visual rewards or praises) in dependence to estimated level of patient efforts.

- (3) Estimation of improvements or generally changes in quality of performed movements. This level of functionality is similar to (2) and the difference stem from the amount of details to be recognized by the system. Whereas in (2) it is enough to provide course categorization, here it is necessary to sense more subtle differences. This level of details might be critical for successful tracking of long-term effects of the therapy.

Further, it is highly desirable that the system is implemented without wearable artificial markers so that it would minimally interfere with normal therapy exercises.

Bearing in mind the cost aspects, it is decided to build the system using a low-cost 3D sensor for recording therapy sessions. The text to follow contains details on the chosen sensing technology and initially obtained results.

3.1 3D sensing technology

Last few years witnessed appearance of a new generation of low cost 3D sensors. They all employ IR projectors to illuminate environment and therefore are effectively limited to indoor applications. According to involved measurement technology, two types of sensors are dominant: structured light sensors and time of flight sensors (Fig. 1). With structured light sensors, a pattern of pixels is projected on external objects and the sensor infers the distance on the basis of disparity between reflected and reference patterns. With time of flight sensors, depth is determined on the basis of measurement of the total travelling time for a light pulse emitted from the IR laser projector, reflected from an external object, and returned back to the sensor. The time of flight method is considered more advantageous, because in principle it allows achieving a more compact design and faster response. However, the actual implementations of both technologies share the similar characteristics: the depth sensor is accompanied by a visible light camera, field of view is on the order of 60 degrees, and working distances are on the order of 1-4m (some sensors allow switching the operating range to near-range distances of 0.15-1m). Thanks to mass production, price is low, on the order of 250€.



Fig. 1 Sample 3D sensors for indoor applications: left – Microsoft Kinect for Windows [24], right – Softkinetic DepthSense 311 [32].

Accuracy and precision are generally low [12, 20] and the errors increase with distance. In the middle of the working range (2m), typical errors are on the order

of 35mm with standard deviation of 10mm, but they can be reduced by calibration [19, 20, 22]. Such level of accuracy can be regarded as satisfactory from the standpoint of whole body motion tracking [7, 26].

3.2 Skeleton tracking data set

To verify ability of chosen sensing technology to record 3D motion, track trajectories, and classify/evaluate performed exercises, a set of measurements has been organized in Milan Petrović School for children with special needs, Novi Sad, and University Children's Hospital Novi Sad. The measurements encompassed acquisition of depth maps and motion sequences for therapy sessions involving nine types of activities performed by children between the ages of 5 and 8 that were affected by mild motor disorders, such as mild intellectual disabilities, hemiparetic form of cerebral palsy and autism. All children understood the instructions and were motivated for work.



Fig. 2 Therapy session recording: left – therapy setting; right – recorded depth map with annotated tracked joints.

Measurements were made using Microsoft Kinect for Xbox 360 sensor and OpenNI/NiTE software [27,28]. The obtained data set contains video clips, depth maps, and 3D motion sequences for fifteen tracked points (see Fig. 2). Collected sequences contain nine types of exercises:

1. Standing at attention;
2. Lateral weight transfer: (a) shifting body weight laterally from left to right and vice versa; right leg lunge with weight shift to the right leg (repeated afterward with left leg);
3. Step forward with weight transfer: right leg step back with lunge with weight shift on the right foot (repeated afterward with left leg);
4. Step backward with weight transfer: right leg step back with lunge with weight shift on the right foot (repeated afterward with left leg);
5. Standing on one foot: One legged stand 3-5 seconds;
6. Crouch;
7. Horizontal walk: forward gait on a horizontal flat surface;

8. Walking up/down stairs;
9. Shoulder movements: arms in front, arms up, arms down, arms back, arms to the sides.

In all cases the collected depth maps appeared satisfactory. However, the inferred trajectories were of low quality.

An illustration is given in Fig. 3 that shows a part of measurements recorded for an exercise involving lateral steps with weight transfer performed by two patients: one (Fig 3a) with very light motion disorders and the other (Fig 3b) having significant problems to keep balance during motion. The first row for each patient contains an outline of a sequence of depth maps. The sequence is accompanied below by corresponding trajectories of the left (red lines) and right shoulder (green lines). The trajectories were obtained in two manners:

- (1) Manually, by selecting joint positions from depth maps (solid lines) and
- (2) Automatically extracted by NiTE skeleton tracking software (dashed lines).

Looking at the red-green pair of solid lines at the diagram of lateral movement for the first patient, one could easily see the part between 1s and 2.5s, when the patient performed a lateral movement of shoulders, and the second part, beginning at 4.5s, when he started the second movement. Moreover, the trajectories are smooth, and the diagram of vertical movements clearly shows that the line between the shoulders is always maintained horizontal, implying a good balance. Trajectories obtained by the skeleton tracking software display similar behavior. However, their dynamics is not appropriate. With fast position changes of tracked joints, direction of registered movement may even become opposite with respect to the actual movement (see crossings of solid and dashed lines in the upper diagram at Fig. 3a). The trajectories obtained for the second patient (Fig. 3b) reflect a rather different behavior: here, the lateral transfer of weight is performed between 1s and 5s and it is characterized by zigzagging with one large overshoot at 4s. Additionally, the diagram of vertical movements shows the characteristic time-varying slope of the line between the shoulders. Here, an important feature to emphasize is that this motion is left unseen by the tracking software: the dashed trajectories are always smooth and they do not show the disparity between the vertical positions of the shoulders.

Such behavior of the trajectory tracking software can be explained by the fact that it is optimized for recognition of the motion and not for evaluation of motion parameters. Furthermore, the software is optimized to run on hardware with modest abilities (alternative Microsoft Kinect SDK software is built using the same motion recognition principles, so that it is expected to display the similar performance).

In conclusion, the tested automatic tracking system could only be used in registering the number of conducted exercises and possibly in very coarse categorization of movements. A more subtle parameterization would require better tracking performances and possibly ability for fine-tuning of the tracking algorithm to the set of tracked exercises. Works on such software exist, e.g. [5,30], and it is worth to explore them.

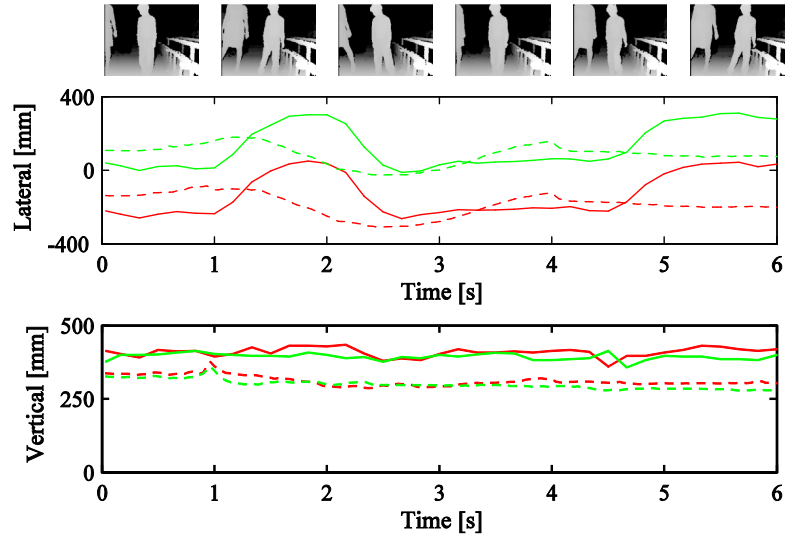


Fig. 3a Trajectories recorded for a patient with light balance disorders: red lines – left, green lines – right shoulder; solid lines – manually extracted from recorded depth maps, dashed lines – automatically extracted by the skeleton tracking software.

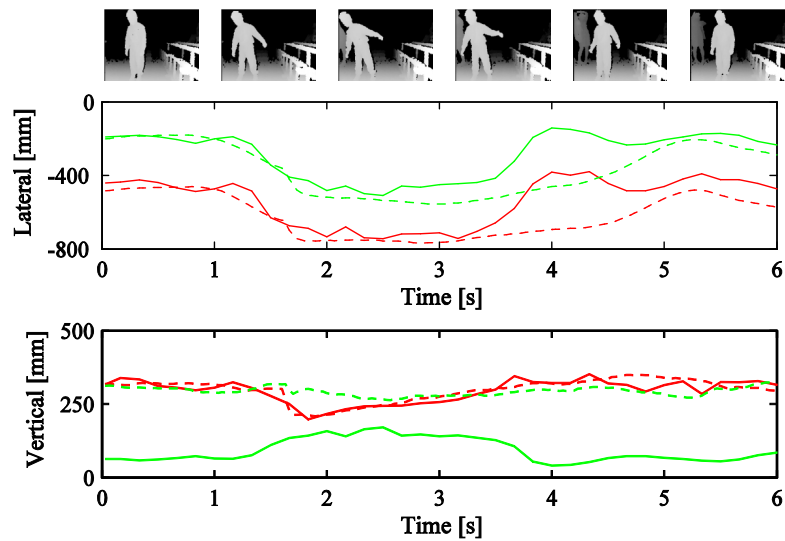


Fig. 3b Trajectories recorded for a patient with heavy balance disorders: red lines – left, green lines – right shoulder; solid lines – manually extracted from recorded depth maps, dashed lines – automatically extracted by the skeleton tracking software.

4 Motivational and interaction aspects

As explained in Section 2, the goal of the therapy is to improve children's awareness and functional control of their own bodies. During the conventional two-party interaction between the therapist and the child, the therapist instructs and guides the child to perform different movements, e.g., to raise an arm, to make a step forward, etc. Related to the types of the exchange commodity, the interaction includes both verbal (e.g., repetition or reformulation of the command, encouragement, etc.) and nonverbal (e.g., physical demonstrations, gestures, etc.) means of communication. In our envisioned system, this general therapeutic scenario is extended to a three-party interaction between the child, the therapist and the animated avatar-based conversational agent, whose role is to facilitate the interaction, cf. [15]. To achieve this, the system must implement two separate but related functionalities:

- *Dialogue competence*: the system should be able to engage in and manage natural language dialogue with the therapist, e.g., to recognize and interpret the therapist's verbal instructions, apply adaptive dialogue strategies, generate appropriate dialogue acts, etc.
- *Spatial awareness*: the system should be able to identify the relevant entities in the spatial context and to determine their positions. For this interaction domain, it means that the system should be able to detect the therapist's and the child's movements and postures.

The research question of the system's dialogue competence is already addressed in previous work in detail [13, 14, 15, 16]. Special attention was devoted to the research problems related to the given computer-aided therapeutic settings: robust processing of spontaneously produced therapist's linguistic inputs (e.g., elliptical or context-dependent instructions that instantiate different frames of spatial reference, etc.), and design of (emotion-)adaptive dialogue strategies for therapeutic purposes. The introduced approaches to these research problems are demonstrated to be appropriate through several prototypical dialogue systems with diverse interaction domains, such as: supporting users while they solve problems in a graphics system, a verbal user interface for the visually impaired, spatial context-aware interaction with the robot, etc.

In this paper, the focus is on the spatial awareness in the context of computer-aided therapy of gross motor skill disorders in children. Thus, the discussion is intentionally dedicated to the vision system, abstracting away from details of the conversational agent. In the text to follow, we elaborate details of engaging the markerless vision-based tracker of human skeleton: (i) to maintain the spatial attention of the child and to increase her motivation to undergo a long-term therapy, and (ii) to support the therapist to keep track of the child's progress.

4.1 Therapeutic settings

The dedicated therapeutic settings are given in Fig. 4. The motion sensor uses skeletal tracking to detect the therapist's and the child's movements and postures. Therefore, both the child and the therapist are expected to be within the range of the motion sensor (this area is marked in gray in Fig. 4), and they may take different spatial perspectives with respect to each other. It should be noted that this technical requirement is not restrictive for the purpose of the therapy. From the child's point of view, both the therapist and the monitor are situated in her social space and may easily come in the focus of attention.

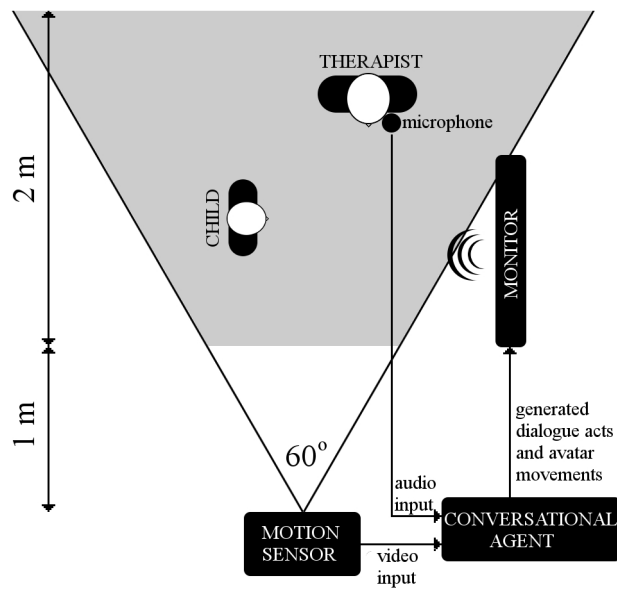


Fig. 4 Therapeutic settings.

4.2 Motivating the child

One of the main goals of the system is to maintain the spatial attention of the child and to increase her motivation to undergo a long-term therapy. To achieve this, the system includes an animated avatar-based conversational agent. The vision system uses skeletal tracking to detect the therapist's movements and postures, and to map them in real-time onto the companion avatar displayed on the screen. The intention is to shift the spatial attention of the child from the therapist to the avatar

that reflects the movements of the therapist. From the child's point of view, the avatar should be perceived as a companion. We expect that an appropriately designed avatar (both in terms of visual appearance and dialogue behavior) may be a strong motivational factor for the child to engage in the interaction and try to mimic the avatar's movements. This is in line with research in the field of robot-aided therapy for children with autism [1, 8, 10, 29, 33, 34] and cerebral palsy [6, 17] showing that the robot may induce positive social behavior in children, such as imitation, eye gaze, joint attention and increased motivation.

Avatar's appearance. Children's preferences for the avatar's appearance may vary depending on various factors such as their age, gender, current emotional state, etc. Therefore, a set of interchangeable, visually different, but functionally equivalent avatars have to be designed. There are two general requirements for the visual appearance of an avatar. First, it should be appealing to the child in order to enable the development of emotional attachment. Second, the avatar should have an anthropomorphic appearance, and the therapy-relevant parts of its body (i.e., arms, legs, head, etc.) should be easily perceivable by the child. In other words, the level of visual details must not be distracting for the child [34]. Fig. 5 provides two prototypical avatars.

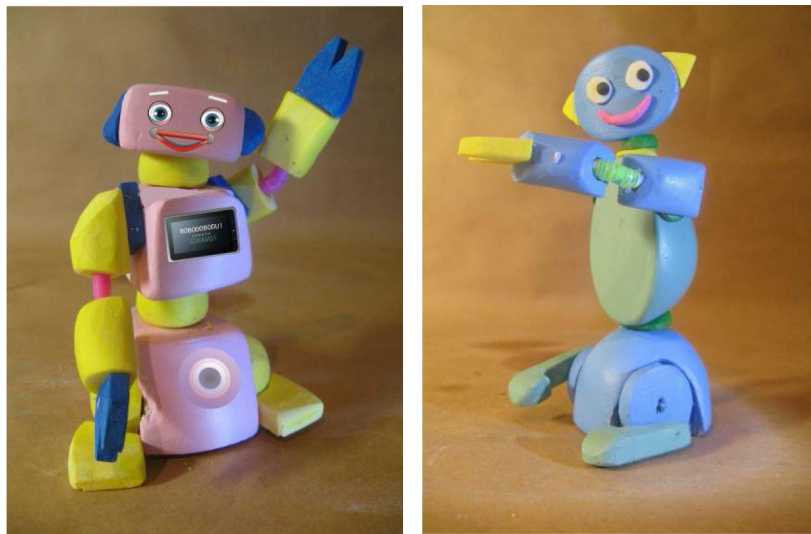


Fig. 5 Examples of female (left) and male (right) prototypical avatars (courtesy of D. Matic and D. Živančević).

Avatar's behavior. One of the possible approaches to developing virtual-reality rehabilitation systems for children with gross motor skill disorders (e.g., cerebral palsy) is to implement a motion-based video game aimed to motivate and encourage the child to use, e.g., the affected limb [21, 35]. However, children with cerebral palsy are very different, and the therapy is always adapted to the particular

child. In contrast to this fact, therapeutic games address only a specific motion and can be applied only to a small subset of children [21]. Therefore, in our approach, we do not presuppose a preset interaction scenario or a therapeutic game. Instead, the therapist is free to organize the interaction according to his own discretion, and may interact with the avatar in two ways:

- Nonverbal interaction: The therapist is demonstrating moves that the child should perform, and the avatar reflects these moves in real-time.
- Verbal interaction: The therapist utters a command to the avatar to perform a non-verbal action (e.g., to raise an arm, etc.). The conversational agent recognizes and interprets the given verbal instruction, and instructs the avatar to perform the requested move. Please note that the conversational agent is discussed in [15] and further references given there.

In both cases, the avatar occasionally displays simulated manifestations of its own intentionality according to its adaptive dialogue strategy (e.g., it makes an encouraging comment to the child, or complains that the exercise is hard, etc.)

4.3 Supporting the therapist

Additional purpose of the overall system is to support the therapist to keep track of the child's progress. The conversational agent can interpret the therapist's instruction, i.e., it extracts information on the movement instructed by the therapist. On the other hand, the vision system can provide information on the child's actual movements and postures. Thus, by fusing these sources, the overall system may infer, in real-time, whether the child performed the therapist's instruction correctly. In addition, it may keep track of the child's performance over a number of therapeutic sessions, which supports the therapist to access the long-term progress of the child.

5 Conclusions

This paper has presented initial steps conducted towards implementation of an advanced system for computer-aided therapy of children. Its most critical component is a system for markerless visual tracking of motion performed by children and it is envisioned as a scalable system built upon a low-cost 3D sensor. To provide a proof of concept, a set of recordings of therapy sessions has been made and the results have confirmed the feasibility of the approach. Additionally, scenarios of possible operation of the system within the framework of an autonomous conversational agent have been elaborated thus setting the directions for further work.

Immediate steps following this study should be focused on specific algorithms for full human body tracking in the context of not only the recognition, but also

the evaluation of conducted moves. Here, two main and interrelated areas of research are tracking algorithms and motion evaluation algorithms.

A complementary task is related to motivational aspect of the therapeutic interaction. In this field, our future work will be focused on the integration of the vision system with the existing conversational agents, and the design of the agent's adaptive dialogue strategies for therapeutic purposes, cf. [16].

Acknowledgments This research is funded in part by the Ministry of Education, Science, and Technological Development of the Republic of Serbia under the contracts III44008 and TR32035. The research is complementary supported by the project *Inclusive physical education in Vojvodina schools: challenges and perspectives*, co-financed by the Provincial Secretariat for Science and Technological Development of AP Vojvodina. The pictures given in Fig. 5 are produced by Dragan Matić and Dragan Živančević, affiliated with the Academy of Arts at the University of Novi Sad, and are not published before. The authors thank D. Matić and D. Živančević for allowing them to include these pictures in the paper.

References

1. Belokopytov, M. and Fridin, M.: Motivation of children with cerebral palsy during motor involvement by RAC-CP fun. In: Proc. of the Workshop on Motivational Aspects of Robotics in Physical Therapy, IEEE/RSJ Int. Conf. Intelligent Robots and Systems, (2012)
2. Bojanin, S.: Reeducacija psihomotorike ili tretman pokretom. Psihijatrija danas, 38(1), 11-27 (2006)
3. Brass, M. et al.: Compatibility between observed and executed finger movements: comparing symbolic, spatial and imitative cues. Brain and Cognition, 44, 124-143 (2000)
4. Buccino, G. and Riggio, L.: The role of the mirror neuron system in motor learning. Kinesiology, 38(1), 5-15 (2006)
5. Buys, K. et al.: An adaptable system for RGB-D based human body detection and pose estimation. Journal of Visual Communication and Image Representation (in press).
6. Calderita, L. et al.: THERAPIST: Towards an autonomous socially interactive robot for motor and neurorehabilitation therapies for children. In: Proc. 2013 7th Int. Conf. Pervasive Computing Technologies for Healthcare, PervasiveHealth 2013, pp. 374-377 (2013)
7. Clark, R. A. et al.: Validity of the Microsoft Kinect for assessment of postural control. Gait & Posture, 36, 372-377 (2012)
8. Colton, M.B. et al.: Toward therapist-in-the-loop assistive robotics for children with autism and specific language impairment. In: Proc. of the AISB 2009 Symposium on New Frontiers in Human-Robot Interaction, Edinburgh, Scotland (2009)
9. Craighero, L., Bello, A., Fadiga, L., and Rizzolatti, G.: Hand action preparation influences the responses to hand pictures. Neuropsychologia, 40, 492-502 (2002)
10. Dautenhahn, K. et al.: KASPAR – A minimally expressive humanoid robot for human-robot interaction research. Applied Bionics and Biomechanics 6(3-4), 369–397 (2009)
11. Dias, O.P., Amaral, T.G.B., and Fernão Pires, V.: Computer assisted learning in manipulative therapy education. Manual Therapy, 16(3), 270-272 (2011)
12. Dutta, T.: Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Applied Ergonomics, 43(4), 645–649 (2012)
13. Gnjatović, M., and Rösner, D.: Inducing genuine emotions in simulated speech-based human-machine interaction: The NIMITEK corpus. IEEE Trans. Affect. Comput. 1, 132–144 (2010)

14. Gnjatović, M., Janev, M., and Delić, V.: Focus tree: Modeling attentional information in task-oriented human-machine interaction. *Applied Intelligence*, 37(3), 305–320 (2012)
15. Gnjatović, M. et al.: Linguistic encoding of motion events in robotic system. In: Proc. 6th PSU-UNS International Conference on Engineering and Technology (2013)
16. Gnjatović, M. and Delić, V.: End-user design of emotion-adaptive dialogue strategies for therapeutic purposes. In: Proc. 23rd Italian Workshop Neural Networks, WIRN 2013 (Series: Smart Innovation, Systems and Technologies). Springer (in press)
17. Huber, M. et al.: PlayStation 3-based telerehabilitation for children with hemiplegia. In: Proc. Virtual Rehabilitation Conference, Vancouver, 2008, pp. 105–112 (2008)
18. Joković-Turalija, I., Ivkić, D., and Oberman-Babić, M.: Neki aspekti rane dijagnostike i terapije djece sa cerebralnom paralizom. *Hrvatska revija za rehabilitacijska istraživanja*, 38(1), 121-126 (2002)
19. Karan, B.: Calibration of depth measurement model for Kinect-type 3D vision sensors. In: Proc. 21st Int. Conf. Central Europe on Computer Graphics, Visualization, and Computer Vision, Plzen, pp. 61-74 (2013)
20. Karan, B.: Accuracy improvements of consumer-grade 3D sensors for robotic applications. In: Proc. IEEE 11th Int. Symp. Intelligent Systems and Informatics (2013)
21. Kelleher, C., et al.: Towards a therapist-centered programming environment for creating rehabilitation games. *Proc. IEEE Computer Games*, 2011, pp. 240–247 (2011)
22. Khoshelham, K. and Elberink, S.O.: Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors*, 12, 1437–1454 (2012)
23. Krstić, N.: Reeducacija psihomotorike kao oblik neuropsihološke rehabilitacije. *Psijhijatrija danas*, 38(1), 51-66 (2006)
24. Microsoft Corporation: Kinect for Windows. Online: <http://www.microsoft.com/en-us/kinectforwindows/> (2013)
25. Moreno, J.C. et al.: Effects of robotic guidance on the coordination of locomotion. *Journal of NeuroEngineering and Rehabilitation*, 10, 79 (2013)
26. Obdržálek, Š. et al.: Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. In: Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS), 2012, pp. 1188-1193 (2012)
27. OpenNI Consortium: OpenNI, the standard framework for 3D sensing. Online: <http://www.openni.org/> (2013)
28. PrimeSense Ltd.: NiTE 2 API programmer tutorial guide. PrimeSense (2013)
29. Ricks, D.J. and Colton, M.B.: Trends and considerations in robot-assisted autism therapy. In: Proc. of the IEEE International Conference on Robotics and Automation (ICRA), Anchorage, AK 2010, pp. 4354–4359 (2010)
30. Rocha, J.: Skeltrack - A free software skeleton tracking library. Online: <http://github.com/joaquimrocha/Skeltrack> (2013)
31. Schmidt, R.A.: Motor learning and performance: from principle to practice. Human Kinetics Pub. (1991)
32. SoftKinetic: DepthSense cameras. Online: <http://www.softkinetic.com/products/depthsense-cameras.aspx> (2013)
33. Thill, S. et al.: Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn. Journal of Behavioral Robotics* 3(4), 209–217 (2012)
34. Vanderborght, B. et al.: Using the social robot probot as a social story telling agent for children with ASD. *Interaction Studies* 13(3), 348–372 (2012)
35. Yamaguchi, T. et al.: Upper-body interactive rehabilitation system for children with cerebral palsy: the effect of control/display ratios. In: Proc. 9th Int. Conf. Disability, Virtual Reality & Associated Technologies, Laval, France, 2012, pp.181–186 (2012)