Multisensory Robotic Therapy through Motion Capture and Imitation for Children with ASD

Rachael Bevill[1], Srineil Nizambad[1], Chung Hyuk Park[1] (PI)
[1]Department of Biomedical Engineering, School of Engineering and Applied Science,
The George Washington University, DC
{rbevill, skn5, chpark}@gwu.edu

Myounghoon Jeon[2] (PI), Ayanna M. Howard[3](Senior Consultant)
[2]Department of Cognitive and Learning Science, School of Sciences and Arts,
Michigan Technological University
[3]School of Electrical and Computer Science, College of Engineering, Georgia Institute of Technology

Abstract— **It is known that children with autism have difficulty with emotional communication. As the population of children with autism increases, it is crucial we create effective therapeutic programs that will improve their communication skills. We present an interactive robotic system that delivers emotional and social behaviors for multi-sensory therapy for children with autism spectrum disorders. Our framework includes emotion-based robotic gestures and facial expressions, as well as tracking and understanding the child's responses through Kinect motion capture.**

I. Introduction

Over the years, there has been a significant increase in the population of children diagnosed with autism spectrum disorder (ASD). Using robotic system as a form of therapy for children with autism is a strong topic of interest for the robotics community. However, the autism spectrum presents such a diversity of cases that there are still many challenging and important issues to pursue. One area of therapeutic research in which research is critical is to provide new approaches and therapeutic tools in the domains of emotional and social interaction for children with autism.

In our work, we propose using autonomous robotic systems to stimulate the emotional and social interactivity of children. Children with autism sometimes have difficulty expressing emotions, as well as understanding the emotions of others. Sometimes, people can present confusing or mixed emotional signals. The use of robots in therapy, instead of focusing on human-only interaction, is ideal for multiple reasons. Robots are simple -- they avoid the detail and complexity of a human's body language and facial expressions. In our framework, we use both humanoid and non-humanoid robots. Humanoid robots are also useful in that they provide greater potential for generalization activities, such as imitation games. Finally, robots are also often more interesting and engaging to children with autism than a human therapist. Therapy can feel more like play than work. [9]

For our study, the robots can be modified to therapeutically interact with subjects by learning and analysing human movement within emotional and social contexts. This is done by utilizing a joint trajectory-based device -- we chose to work with the Microsoft Kinect. We aim to study the relationship between emotions and motion movements, and how those correlations may vary between neurotypical children and children with autism.

II. The Set-up

For our system, we are currently using two robots. The first is a modified Romotive, referred to hereinafter as Romo. Romo is an iPhone rover type robot which focuses on demonstrating emotion through facial expressions. The Romo character was modified from a bright blue monster to a baby penguin character, in order to be more appealing to children. Both Romo's expression and the background color can be changed to reflect the appropriate emotion. Children with autism often see colors at a higher intensity than neurotypical children [5]. For this reason, all emotions are associated with pastel colors.

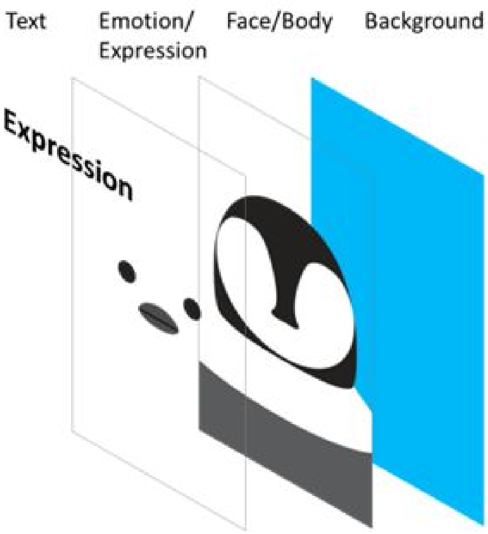
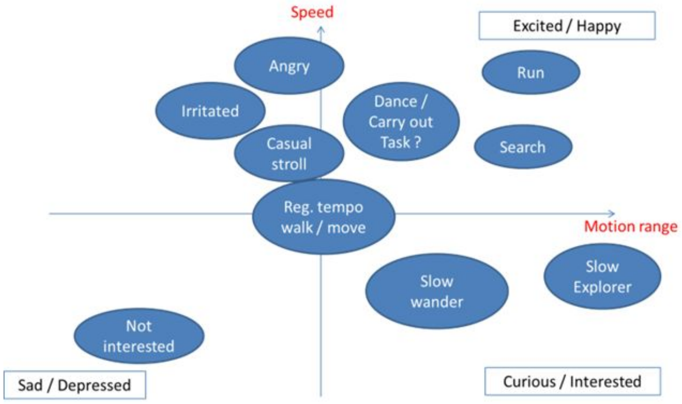


Figure 1 (left). The emotional color wheel is used to help young children identify distinct emotions [3].
Figure 2 (right). Romo's separate expression layers allow for independent adjustment of the expression and background color.

Darwin Mini, our second robot, is a humanoid that expresses emotions through body language. Each robot is currently programmed with the same fourteen emotions and actions, a sample of which can be seen below in Figure 3. Nine more emotions are currently being created, to give our robots a full range of the emotional spectrum. In order to effectively use the robots to teach emotions, it is important a wide variety of emotions are displayed.

Figure 3. The current emotions available for Romo. Darwin Mini expresses the same emotions through body language.

Our emotions were chosen based on Russell's Circumplex Model of Affect [8]. This model breaks emotions down into two axes: arousal, which refers to the energy level present with the emotion, and valence, which indicates whether the motion is a positive or negative experience. Between these two axes, all levels of emotion can be represented. Romo can express a limited degree of emotion by driving and tilting forward and backward. Darwin Mini is more adept in expression through movement. Each robot's mobility was taken into account. Along with Russell's Circumplex Model of Affect, each robot uses a variety of speeds and motion ranges to express emotion.



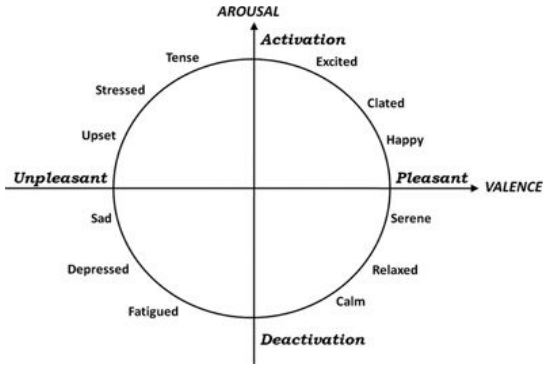
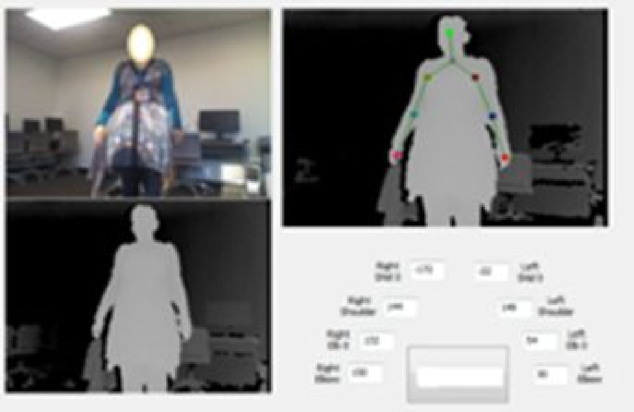


Figure 4 (left). Spatial mapping of motion primitives with emotional mapping.
Figure 5 (right). A version of Russell's Circumplex Model of Affect.

Our study starts by engaging the child and robots in fun, demonstration-based behavioral sessions, to help the child learn to overcome sensory overload. Once the child has gained empathy for the robots, they will progress to interactive dance activities. The child's physical responses will be tracked by the Kinect. The child can mimic the robot, and the robot can track and mimic the child.

III. Tracking Emotional States via Body Language

Microsoft Kinect is an RGB camera that can motion capture the changes in the dynamics at joints, tracking the positions of about 30 joints in a human subject. Figure 6 depicts the user interface for the device and Figure 7 depicts a sample movement sequence of joints. We utilize Microsoft Kinect as a tool to track the engagement of the children with the "social robots" employed. It is also used to provide live feedback within our robotic framework, to make Darwin and Romo more suitable to the children's emotions and desire. For instance, if the robots are too overbearing and the children start becoming anxious, that information will be fed back into the robots' functionality using motion dense sequences tracked by Kinect. The robots' functionality will "learn" and adjust accordingly.



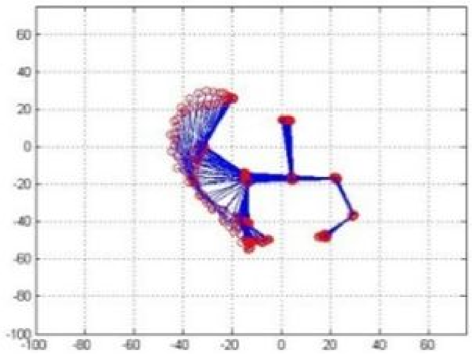


Figure 6 (left). Graphic user interface of the Kinect system.
Figure 7 (right). Captured trajectories of upper-torso movements.

We will calculate changes in motional features, such as velocity, acceleration, and torque, and partly draw on Laban Movement Analysis to understand the body language of subjects as they interact with the social robots. As shown in Figure 8 below, basic emotion emotional states, such as anger, fear, disgust, surprise, happiness, sadness, tend to involve distinct expressions in body positions, which can be tracked by Kinect. Laban Movement Analysis (LMA) is one generally well-known method for categorizing and interpreting the differences in human movement, based on 4 categories: body, space, effort, and shape. Because of the practicality in quantifying motion via the effort category, in our paradigm we only take account of effort.

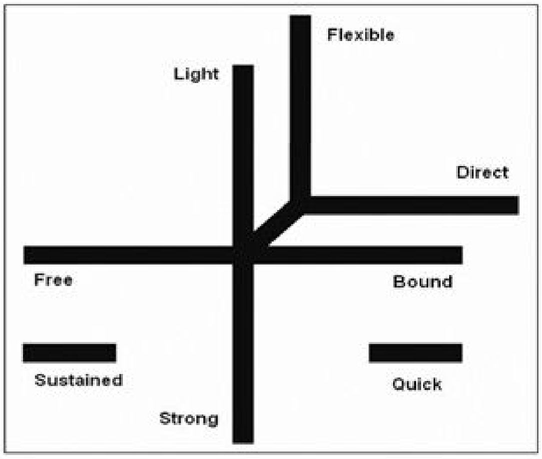
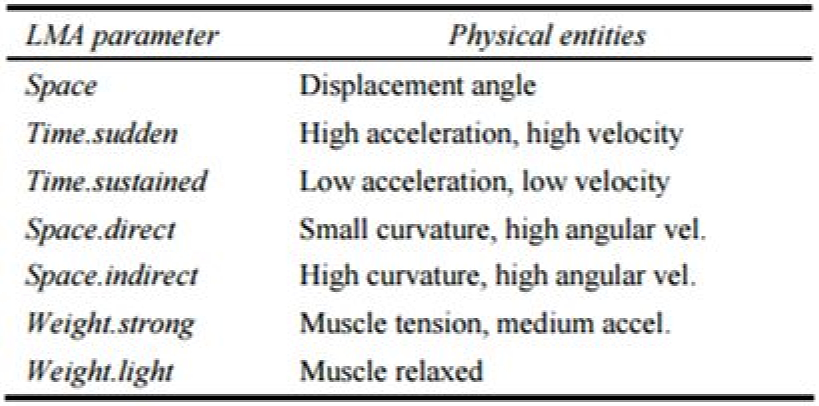


Figure 8 (left). Sample correspondence of emotions to posture and body language [1].
Figure 9 (right). Laban effort graph with subcategory extremes.

Effort is generally divided into four subcategories, which to some extent, can be quantitative for our purposes in analyzing emotions and relating to emotions: weight, space, time, and flow. Weight focuses on the amount of force exerted by the body (strong vs light). Space focuses on the attitude towards a movement pathway (direct vs indirect or flexible). Time focuses on how urgent a movement is (sudden or quick vs sustained) and can be measured by tracking acceleration. Flow focuses on how controlled/intentional a movement sequence is (bound vs free) and can be related to the amount of muscle contraction. Figure 9, above, shows the extremes of the effort subcategories. A "strong, bound, quick, and direct" movement is known as of "fighting form" while a "weak, free, sustained, and inflexible" movement is known as of "indulging form."

Initially, we will involve measure weight and time factors, as those are the most practical and feasible parameters for which quantitative information about movement can be attained. Additionally, we will also create a third parameter, "motion unit" parameter to quantify the number of continuous movements a subject does before pausing. For instance, if an arm is moved back and forth five times in one single time span with no pause in the arm movement, it would be measured as "one motion unit;" if on the other hand, there was five pauses between the movement back and forth, it would be measured as "five motion units."

Time and weight features can be calculated by quantifying the kinematic features of the joints Kinect tracks. As shown by Table 1 below, both high acceleration and high velocity of joint movement both indicate that a movement is sudden, whereas low acceleration and low velocity indicate that a movement is sustained. Knowing whether an action is indirect vs direct, for instance, can give us information about the inner attitude towards a particular movement. If it is direct, we could say that the subject showed confidence towards that movement. Curvature can be approximated by finding the change of the displacement angles (angular velocity).



LMA parameter	*Physical entities*
Space	Displacement angle
Time.sudden	High acceleration, high velocity
Time.sustained	Low acceleration, low velocity
Space.direct	Small curvature, high angular vel.
Space.indirect	High curvature, high angular vel.
Weight.strong	Muscle tension, medium accel.
Weight.light	Muscle relaxed

Table 1. Initial hypotheses of correspondences between LMA parameters and physical entities [7].

IV. Relations between Body Movement and Emotions

A number of rising studies indicate strong correlations between kinematic features, including LMA features, and emotional states. In one study [6], ten people were put in one room hearing "pleasant sounds" and ten others in another room hearing "unpleasant sounds." Researchers rated the persons' emotional states after hearing the sounds & their movements were recorded by video cameras while in the room. Laban Movement Analysis was carried out on their movements and indicated that unpleasant sounds induced active, directed, and sudden movements, that were thus part of "fighting form" whereas pleasant sounds induced unsteady, weak movements leaning towards "indulging form" [6]. It was also revealed that tension, depression, fatigue, and confusion related to active, directed, and sudden movements while fatigue was exhibited in particular in strong and rapid movements. Moreover, the unpleasant sounds induced tension, and increasing tension coincided with an increase in the time feature. This makes sense given that time measures the degree of hurriedness in movement changes and more hurriedness in general is a sign of greater tension. Findings such as these support the strength and validity of Laban Movement Analysis as a quantitative tool for measuring body language features in order to relate to emotional states.

V. Exact Equations for LMA Parameters

Numerous studies have established variations of equations to quantify the LMA parameters used. In our research, we initially, as previously described, utilize the time and weight LMA parameters. We also utilize the concept of a "motion unit" to quantify a certain movement.

Va. Weight Term

Part of LMA effort classifiers, the weight term gives an indication of the strength or force of movement exerted by body for a particular movement sequence. A weight term can be light or strong. One approach to calculate the weight term is to do a weighted sum of all the joints torques'. The two variables that are intrinsically dependent on the subject are radius and mass of the joint.

Equation A [5]:
Note: $_i$ is an arbitrary joint
In general, the equation for weight term can be described by:
Weight per frame/second=(∑**torque$_i$)/#joints=(∑(r$_i$F$_i$sin(theta$_i$)))/#joints**

Vb. Time term

Another effort classifier, the time term helps distinguish between movements which are sustained versus those that are sudden. Therefore, the time term can represent the speeds of a particular movement sequence and the speeds' changes over time and can be useful, as an example, for identifying meditation and concentration states (sustained states) versus frustration and excitement states (sudden states).

Equation A [10]:
One approach is based on assessing changes in acceleration and is represented as follows:
Time term per frame/second=**abs((∑derivative of acceleration)$_i$)/#joints**

Equation B [5]:
Another approach is:
Time term per frame/second=**((∑(angular velocity)$_i$)/#joints**

Vc. Motion Unit

While a "motion unit" is not particularly a LMA feature, it is also useful in characterizing motion sequences like the other features. A motion unit can be said to be a pause in a whole-body movement. We can say that a motion unit is analogous to a minimum in the movement's acceleration feature, or an instance when the absolute velocity of the movement is approximately zero. The absolute velocity will not be exactly zero though due to noise, so an approximate threshold will be determined based on trial. One approach is therefore to subjectively assess a few motions for the number of motion units that can be observed with the eye and use Kinect at the same to provide data for plotting the acceleration for that movement sequence and track the minimums and correlate with the number of observable motion units. Then, an approximate velocity threshold for when a new motion unit occurs can be garnered.

VI. Discussion and Future Work

In our next steps, we also create sessions where the robots react to music and perform music-based movements and emotional gestures. Rather than just analyzing the Kinect information after the session, the robots will be able to react to the children's reactions in real time based on the Kinect motion sequence information. This will make it more efficient in forming rapport. The robots will mimic the child's actions, using input from the Kinect. Different types of songs, and different reactions from the children, will generate new robotic actions.

VIII. References & Acknowledgments

1. Bernhardt, D. Emotion Inference from Human Body Motion. Diss. U of Cambridge, n.d. N.p.: n.p., n.d. Print.

2."Color Talk - Colorfully Inspired." Weblog post. *: Color & Autism*. N.p., n.d. Web. 05 Mar. 2016. <http://colorturners.blogspot.com/2011/03/color-autism.html>.

3. "Do2Learn: Educational Resources for Special Needs." *Do2Learn: Educational Resources for Special Needs*. N.p., n.d. Web. 05 Mar. 2016. <http://do2learn.com/organizationtools/EmotionsColorWheel/index.htm>.

4. "Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis." *Principles of Practice in Multi-Agent Systems Lecture Notes in Computer Science* (2009): 322-34. Web.

5. Masuda, Megumi, Shohei Kato, and Hidenori Itoh. "Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis." *Principles of Practice in Multi-Agent Systems Lecture Notes in Computer Science* (2009): 322-34. Web.

6. Morita, Junya, Yukari Nagai, and Tomoyuki Moritsu. "Relations between Body Motion and Emotion: Analysis based on Laban Movement Analysis."

7. Rett, Jorg, Jorge Dias, and Juan Manuel Ahuactzin. "Bayesian Reasoning for Laban Movement Analysis Used in Human-machine Interaction."*International Journal of Reasoning-based Intelligent Systems IJRIS* 2.1 (2010): 13. Web.

8. Russel's Circumplex of Affect image - http://www.frontiersin.org/files/Articles/13744/fneng-05-00003-HTML/image_m/fneng-05-00003-g001.jpg

9. Ricks, Daniel J., and Mark B. Colton. "Trends and Considerations in Robot-assisted Autism Therapy." *2010 IEEE International Conference on Robotics and Automation* (2010): n. pag. Web.

10. Wakayama, Yuki, Seiji Okajima, Shigeru Takano, and Yoshihiro Okada. "IEC-Based Motion Retrieval System Using Laban Movement Analysis." *Knowledge-Based and Intelligent Information and Engineering Systems Lecture Notes in Computer Science* (2010): 251-60. Web.

This research is supported by the National Institutes of Health (NIH) under grant # 5 R01 HD082914-04 through the National Robotics Initiative.