Part V

Relevant Sample Applications

20

An Internet of Things Approach to "Read" the Emotion of Children with Autism Spectrum Disorder

Tiffany Y. Tang and Pinata Winoto

Wenzhou-Kean Autism Research Network, Assistive Technology Research and Development Center, Department of Computer Science, Wenzhou-Kean University, Zhejiang Province, China

20.1 Introduction

Following the discussion on the contributions of Internet of Things (IoT) in health science and management in previous chapters, this chapter describes a prototype of IoT system designed specifically to help children with autism spectrum disorder (ASD) who are notoriously known to lack some social skills. Specifically, the application is designed to help both users with or without ASD to recognize their emotion in order to facilitate social interaction. Since the effectiveness of the prototype has not been clinically tested, readers are advised to consider it as another future IoT application in health and learning management. In this chapter, we describe the design and development of such an IoT prototype.

During a social interaction, it is important for all participants to interpret correctly the emotional nature of the message being communicated in order to avoid misunderstanding. Prior researches in the past decades have pointed out the lack of emotion recognition skills and social interaction skills among children and adults with ASD and the possible causal link between the former and the latter (Harms et al., 2010, Baron-Cohen et al., 1993). Since all participants in a social interaction should be held accountable for correctly interpreting other's emotional state, the ability of neurotypical (NT) individuals to recognize the emotion of children with ASD should also be evaluated. A small number of recent works began to study these issues by empirically evaluating the emotion expressiveness of individuals with ASD (Brewer et al., 2016; Macdonald et al., 1989; Faso et al., 2015; Stagg et al., 2014). However, the research on the ability of NT individuals in "reading" the emotions of those with ASD is still rare. The study involving advanced technology is even rarer, appealing, and yet seems

Internet of Things A to Z: Technologies and Applications, First Edition. Edited by Qusay F. Hassan.

to be more feasible with the overwhelming focus on and success of sensing and wearable technologies. In particular, would it be possible for a collection of sensors and wearable devices acting as "eyes" and "ears" to collectively "label" the emotion of individuals situated in a social environment engaging in natural interactions? Drawn from earlier studies on psychology that have successfully linked emotions with expressive body movements (Boone and Cunningham, 1998; De Meijer, 1989), researchers have advanced our understandings of discern emotion of NT individuals from such multimodal data, including facial expression, hand gestures, body movements, and so on (Camurri et al., 2003; Kapur et al., 2005; Gunes and Piccardi, 2007; Robinson, 2014). However, there is a lack of work to assist NT individuals to understand the emotional states of those with ASD under natural social interactions (Tang, 2016), which motivates this study.

In this chapter, we report our early attempt to construct an IoT-based natural play environment designed specifically to "read" the emotion of children with ASD. The rationale behind this play environment is that the behavioral data including children's body movements and hand gestures provide rich data for an emotion-learning algorithm (Mitchell, 2009), which is coined as an area called computational sensing (Rehg et al., 2014). In order to capture valid behavioral data, a naturalistic IoT environment was created that contains embedded sensors, toys (e.g., LEGO® toys), and other objects, combined with the facial and behavioral data captured by smartwatch, IP camera, and Kinect™ sensor to "generate" emotion labels in the provided play environment where ASD children and other NT individuals interact (Figure 20.1).

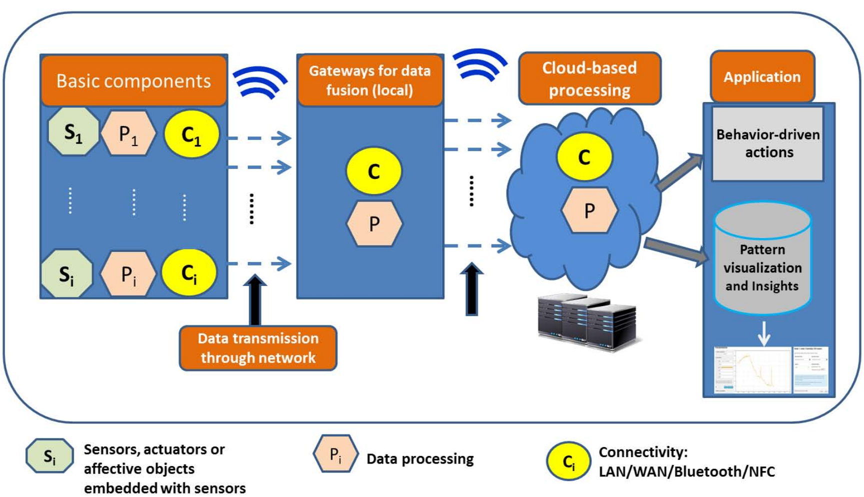

Figure 20.1 The box-level view of a typical IoT environment. (Adapted from Karimi (2013).)

20.2 Background

In this section, some background information on current research and development, as well as the challenges of technology-based intervention on autism spectrum disorder in China will be presented.

20.2.1 Current Approaches of Technology-Based Intervention on Autism Spectrum Disorder in China[1]

Thanks to the numerous impressive works on technology-based intervention (TI) (among many, Carter and Hyde, 2015; Beykikhoshk et al., 2014; Hong et al., 2015, Seo et al., 2015, Tartaro et al., 2015; Tang et al., 2015), significant positive outcomes have been achieved medically and clinically.

However, compared to the plenty of research efforts on the design and evaluation of TI on ASD in the United States and Europe, there are a few recently published works (Wang, 2015; Tang, 2016; Tang et al., 2015, 2017a, 2017b; Tang and Flatla, 2016; Winoto et al., 2016a, 2016b; Winoto, 2016) in Eastern Asia, but compared with its Western counterpart, the progress still lag far behind (Sun et al., 2015). Until now, the number of ASD population (both adults and children) in China is still undetermined (Sun et al., 2015). Indeed, very few researches had been published regarding the assessment, diagnosis, and early intervention strategies for children with ASD in China until recently (Sun et al., 2015).

20.2.2 The Challenges of Technology-Based Intervention on Autism Spectrum Disorder in China

As discussed earlier, the special education system in China is largely different from that in developed countries such as the United States in that the majority of children with autism would end up going to private educational centers to receive education. There are also a number of government-funded special education schools in each province or city, but due to the extremely limited resources, the majority of these special education schools are largely inaccessible to these children. In addition, there is a huge gap between the available number of teachers who are trained to deal with children with autism and the needed number.

1 The focus of this chapter is on the research and practices in mainland China (excluding Taiwan and Hong Kong) where the diagnosis, assessment and technology-based intervention have been far lag behind its neighboring countries and regions. For a more complete discussion of the current practices and research progress in mainland China, readers can refer to (Tang and Flatla, 2016).

Developing an affordable, portable, and personalized application that can be delivered at home has become time-pressing. Familial, social, and cultural factors could affect the participation, acceptability, and even outcomes of therapeutic approaches (Zwaigenbaum et al., 2015). Therefore, to include socially and culturally diverse populations in intervention, research is essential so as to enhance and deepen our understanding of how technology can be accepted, adapted, and deemed helpful across populations of individuals with ASD (Tang and Flatla, 2016).

20.3 Related Work

Social interaction is inherently bidirectional (Halberstadt et al., 2001), requiring individuals to recognize each other's emotion and intent that is vital for their action and behavior (Tang et al., 2017a, 2017b). In this section, previous works related to emotion recognition abilities and emotion expressivity of individuals with ASD will be discussed.

20.3.1 Emotion Recognition in Autism Spectrum Disorder

Individuals with ASD generally have a typical or delayed emotion processing (World Health Organization, 1993). Most works in this area focus on examining their abilities to perceive others' emotions via their static facial expressions (Harms et al., 2010; Baron-Cohen et al., 1993), which is typically attributed to anomalies in facial expressions (Simmons et al., 2009). However, these empirical studies have yielded mixed and even some contradictory outcomes: For example, it is argued that previous research might have underestimated the emotion recognition abilities via face among children with ASD (Peterson et al., 2015); additionally, it is revealed that ASD children are more skilled in recognizing emotion via body movements (Peterson et al., 2015). Others have failed to find any significant differences between NT and ASD groups involving basic emotion recognition tasks (Baron-Cohen et al., 1993, Baron-Cohen et al., 1997; Harms et al., 2010).

20.3.2 Emotion Expressiveness of Individuals with Autism Spectrum Disorder

Unlike the many research to study the emotion recognition ability of individuals with ASD, very few empirical studies have been reported on their emotion expressiveness (Brewer et al., 2016; Macdonald et al., 1989; Faso et al., 2015; Stagg et al., 2014). Emotion expressiveness includes the abilities to understand, mimic, and pose various emotions; of these, the first requires mental understanding of emotion and its communicative value, and the last two require facial muscle movement skills and proper feedback. Prior researches showed that children with ASD have reduced facial muscle movement during playing (Czapinski and Bryson, 2003), “atypical” looks during emotional story-telling (Grossman et al., 2013), and decreased proprioception awareness levels toward their own facial muscle movements (Weimer et al., 2001), all of which might decrease their emotion expressive abilities (Brewer et al., 2016). Taken together, it might reduce the quality of their social interaction, and negatively affect the participating NT individuals.

20.3.3 Emotion Recognition by Neuro-Typical Individuals

To date, very few literatures have reported empirical studies or technology-based design for helping NT or ASD individuals to recognize the emotions of individuals with ASD. For example, Park et al. (2012) proposed a system framework for teaching children with ASD to identify their own emotion through body language, without implementing it. A recent study discussed an early pilot study of utilizing a portable motion sensor to continuously capture the facial landmark data of children with ASD when they are watching a video; the temporal facial data will then be used to automatically generate emotion labels to inform NT individuals so as to facilitate the social interaction between NT individuals and those with ASD (Tang et al., 2017a, Tang et al., 2017b). Compared with an overwhelming number of previous studies probing the impairments of emotion recognition among individuals with ASD and implementing various computerized emotion-recognition training application for individuals with ASD, applications such as those proposed in Tang et al. (2017a, 2017b) that can assist them in informing others (especially those NT individuals) of their emotion are highly desired (Virnes et al., 2015; Tang et al., 2017a, 2017b; Tang, 2016).

20.3.4 Affective Computing, Multisensory Data Collection in Naturalistic Settings, and Ubiquitous Affective Objects

20.3.4.1 Naturalistic Settings and Ubiquitous Affective Objects

Thanks to more affordable sensory and wearable technologies, impressive improvements have been made in affective computing (Williams et al., 2015). Physiological data can be obtained easily from various wearable affect monitors (e.g., Fitbit Surge) and affordable sensors, including pressure sensors, capacitive sensors, and GSR; user’s environmental information can also be collected using other sensors, such as ambient light sensors, temperature meter, PIR infrared sensor, and so on, which could provide relevant contextual information to predict users’ emotion and eventually to help in mediating their effects (Williams et al., 2015). Ertin et al. (2011) designed a multisensor suit (a total of six sensors) to continuously measure a user’s stress level where both body and ambient temperature data were obtained to monitor the body thermoregulatory and nervous system activation. These raw sensory data can then be transmitted via Bluetooth to a central system consisting of an algorithm to compute user emotions. Mitchell (2009) argued that various speakers' information such as their intonation, physical distance, body language, and other upper body movements can be used to characterize their interpersonal interactions.

Affective objects refer to ". . . any physical object which has the ability to sense emotional data from a person, map that information to an abstract form of expression and communicate that information expressively, either back to the subject herself or to another person" (Scheirer and Picard, 2000). This is particularly important especially for people with physical or learning disabilities (Williams et al., 2015). For the visually impaired who cannot identify emotions of others, a glove (named as *VibroGlove*) with vibrations can convey that information to the wearer (Krishna et al., 2010). Williams et al. (2015) go even further by designing a fashionable actuator-based scarf that can be used to sense the wearers' emotional state using attached sensors and then helping them to mitigate and convey their emotions to others via some actuators (e.g., LEDs). However, in the initial user testing (including one with high functioning autism), regarding whether to publicly share the user's emotion, all participants, including the one with ASD, strongly disapprove the design. As for the acceptability of group emotion broadcast, all participants are more willing despite overall reluctance except for the person with ASD and a visually impaired female participant. The results showed a strong privacy concern on broadcasting user emotion either at individual or group levels. In addition, the representation and evaluation of emotions are simplified in the study, for instance, low body temperature is associated with both "stressed" and "excited." However, ambient temperature might also have a significant impact on human being's physiological responses to daily physical activity (Tyka et al., 2009).

20.3.4.2 Sensing the Emotion from Behavioral Data Analysis

Some earlier attempts have focused on associating a specific emotion type with a set of behavioral data. Pollick et al. (2001) showed that different qualities of body movements can effectively be associated with distinct emotions. Castellano et al. (2007) proposed an emotion recognition technique based on body movement and gesture expressivity analysis by which emotional states such as anger, joy, pleasure, and sadness can be learned. Rehg et al. (2014) pointed out the drawbacks of current methods for acquiring social and communication behavioral data, and thus proposed the adoption of a so-called behavioral imaging to efficiently and effectively collect multimodal behavioral data through audio, video, and wearable sensing. Since human behaviors are multimodal in nature, these behavioral cues might provide rich information to construct user's behavior modeling. Such physiological data such as electrodermal, respiratory, and cardiovascular ones can shed light on the quality of social interactive behaviors that are not typically visible or observable to a therapist directly.

Similar to Rehg et al., 2014, Robinson (2014) suggested that emotional cues can be automatically learned from facial expressions, tone of voice, body postures, and gestures, which, if appropriately broadcasted, can assist individuals with ASD overcome their notoriously known difficulty of being understood by others.

20.3.5 The Internet of Things in Monitoring and Tracking Individuals for ASD Intervention

Although IoT has become increasingly popular in the health care industry, few published research has been reported for its adoption in ASD till recently. Sula et al. (2013) constructed a SmartBox featuring vibrator, light, smell, and sound control with an aim to provide a calming and motivating environment for autistic children when they are engaging in various learning activities. For example, the room light will be changed to adapt to the child's visual reading preferences (reading mode); a smell control to maintain pleasant room smell; vibrating the chair or bed to keep the child relax and calm. Children were asked to wear wireless body sensors to detect their body and hand movement as well.

Karimi (2013) discussed an IoT framework for remote emotive computing where such physiological variables and states were discussed for emotion data retrieval; these variables include muscle relaxation and contraction (via a pressure sensor); heart rate variability (via an on-chip two-electrode ECG), sweat (via a capacitive sensor), attitude (via an accelerometer to monitor wearer's body), and hand movements. These data are suggested to be measured in a natural environment where contextual awareness can be ensured and made meaningful for software designers to conduct behavioral analysis in the application. When an object is embedded with such sensors, it becomes affective object (Scheirer and Picard, 2000).

In summary, although most of these previous studies have, to a certain extent, effectively demonstrated the great potentials of probing the behavioral pattern of both individuals with ASD and NT individuals with more unobtrusive and efficient means, little research has been done on a wider adoption of various kinds of embedded sensors on objects or toys and on-body sensors, and how these sensing data can be combined with facial expression analysis, which motivates the research here. Figure 20.1 illustrates the functional box-level view of an IoT environment.

In the following sections, the system design and some initial analysis on the potentials of such an environment will be presented.

20.4 The Internet of Things Environment for Emotion Recognition

Driven by previous studies, an integrated IoT platform was developed and discussed in this section.

20.4.1 System Background and Architecture

To solve the privacy issues, this model is primarily designed for activities at home, school, or medical centers; and it focuses on embedding some sensors into "objects" children would play with (e.g., toys). The goal is to sense their emotion indirectly or noninvasively because individuals with ASD tend not to like wearable objects (Williams et al., 2015). Selective sensors are embedded into children's natural play environments to obtain their routine behavioral data combined with other information retrieved from monitoring devices such as Kinect and web cams to recognize the children's emotional state so as to inform their teachers, parents, or therapists.

Unlike the work by Sula et al. (2013) that requires the autistic children to wear body sensors, the current IoT system did not consider it due to observations from the numerous field testing in autistic educational centers and the feedback from parents and special education teachers. These wearable sensors are considered more intrusive to autistic children.

Figure 20.2 shows the overall system architecture, and Table 20.1 lists the emotion-related physiological data and emotion-sensitive environmental data recorded and stored in the system that follows the recommendation for remote-emotive computing in Karimi (2013). Two types of sensors have been designed to obtain information either from the human being or from the ambient environment (Figure 20.2).

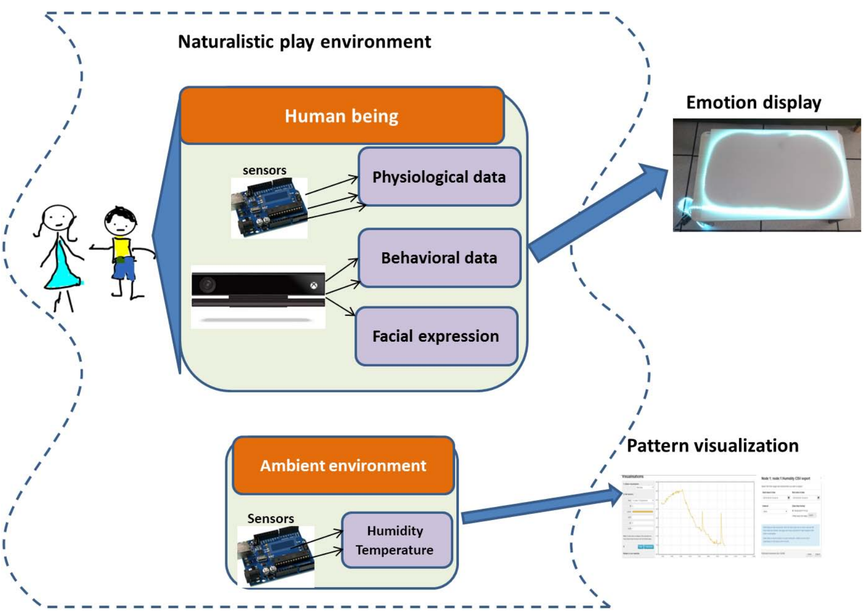

Figure 20.2 Overall system architecture.

Table 20.1 Emotion-related physiological data and emotion-sensitive environmental data in an IoT environment for emotion recognition.

Emotion-related physiological data	Sensors	Current IoT environment
Muscle relaxation	Pressure sensor	√
Muscle contraction	Pressure sensor	√
Sweat	Capacitive sensor	√
Body movement	Accelerometer	×
Heart rate variability	Wearable affect monitors	×
Emotion-related behavioral data	**Sensors**	
Attitude via body, head, and hand movements	Motion capture sensors such as RealSense™, Kinect, Leap Motion™, and Myo™ Gesture Control Armband	√
Facial expression	Motion capture sensors such as RealSense™, Kinect	√
Ambient factors contributing to emotion	**Sensors**	
Temperature	Thermometer	√
Humidity	Barometer	√
Light	Light sensor	×

In the next section, the details of the system setup, running environment, and report on some initial findings will be presented here.

20.4.2 The Naturalistic Play Environment

In our current design, there are three types of measurements used concurrently to sense users' behavioral patterns and individual and group emotions (see Table 20.1):

a) *Physiological (individual) measurement.* heart rate and perspiration (obtained via a set of Microsoft Band 2 worn by the target user)
b) *Behavioral (individual) measurement.* upper-body movements (including head and hands), gestures and motions (obtained via pressure and touch sensors), and facial expression (will be included in the future system design)
c) *Sociometers.* embedded sensors and affordable depth and RGB-B sensors (e.g., two sets of Kinect V2 sensors).

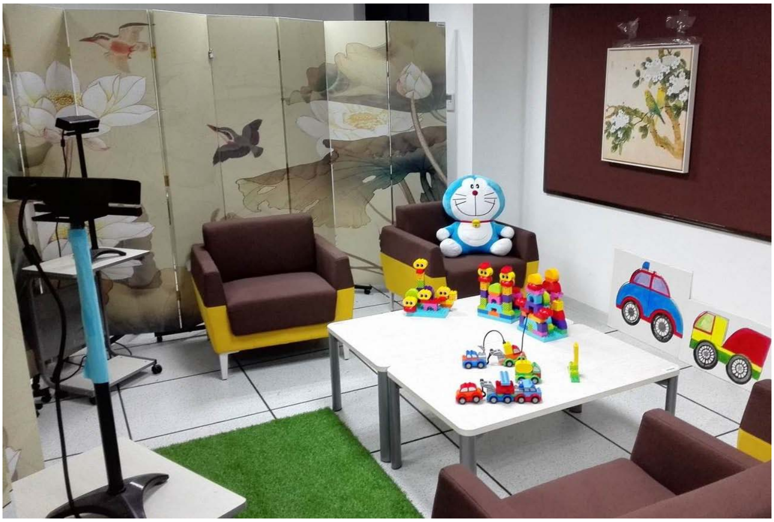

Figure 20.3 An example of a testing room consisting of two sets of Kinect sensors, a humidity + temperature sensor, two foam-like drawings equipped with capacitive sensors, a 1 m^2 synthetic grass equipped with pressure sensor (max. 40 kg), and some LEGO bricks.

Figure 20.3 illustrates the initial setup of playing environment. An IP camera was installed (above the play tables) to capture the play interactions (Figure 20.4). Figure 20.5 demonstrates a play moment captured by the IP camera; as can be seen, a player can touch the picture of a car where two capacitive sensors were installed behind its wheels (Figures 20.6 and 20.7). The collected touch-force data will be stored in the Cloud for later analysis. Currently, our design goal is to use the sensors to obtain the correlation between the level of touch force and the player's mood.

20.4.3 Sensors and Sensor Fusion

20.4.3.1 Hardware Design on Emotion and Actuation

Motivated by prior work on the link between various emotional state and biological and environmental triggers (Williams et al., 2015; Robinson, 2014, Rehg et al., 2014), we embedded a number of sensors for the purpose of collecting users' behavioral, physiological, and environmental data. And similar to the work described in Williams et al. (2015) and Tang et al. (2015), the sensors were employed in the Arduino platform due to a wide variety of compatible and affordable sensors available on the market. Table 20.2 lists the sensors used in this work. In addition, there is also a plan to make use of wearable devices to obtain users' physiological data (Microsoft Band 2).

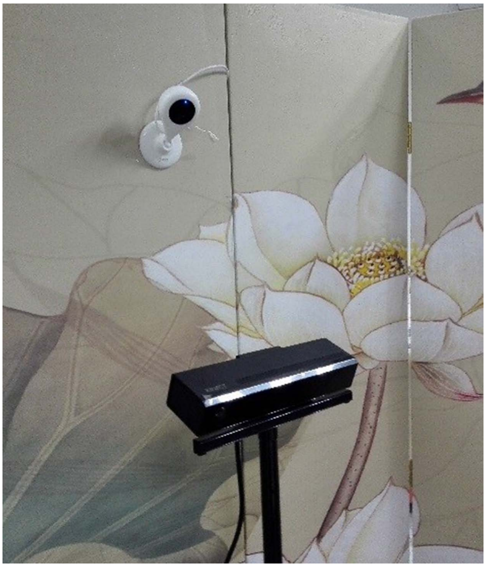

Figure 20.4 An IP camera has been installed to capture the interactions between players and the "affective" objects.

Figure 20.5 An image captured by the IP camera shown in Figure 20.4.

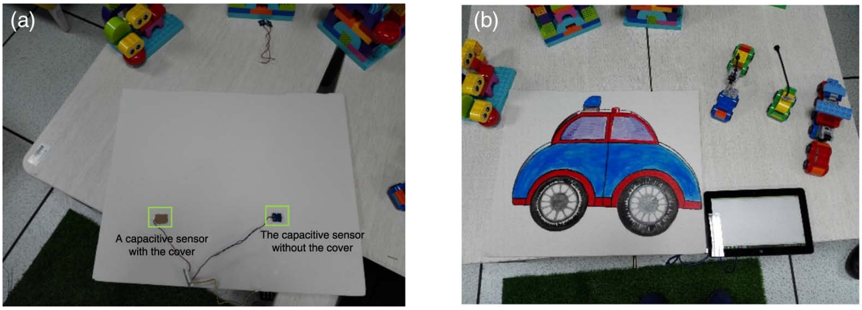

Figure 20.6 Two capacitive sensors embedded in the car wheel (a). A Windows tablet is put to visually present the touch force at real-time (b).

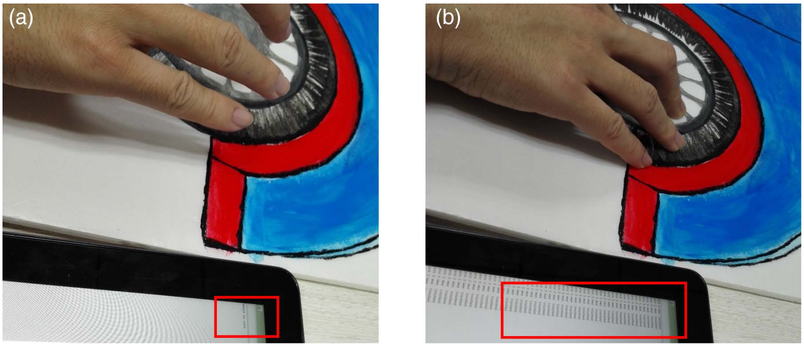

Figure 20.7 The testing of capacitive sensors, before (a) and after (b) touching the car wheel; the data bar is seen increasing on the screen.

Table 20.2 List of sensors used in current environment.

Function	Hardware components
Indoor temperature and humidity detection	(a) SEEED Grove – Temperature and Humidity sensor DHT22[a)] (b) Arduino UNO R3 (a microcontroller board)[b)] (c) Dragino Yun Shield V1.1[c)] (*a strong* shield for Arduino Board for Internet connectivity and storage issue) (d) SEEED Base Shield V2 Grove for Arduino[d)] (can be plugged into an Arduino as an expansion board)
Painting interaction	(a) SEEED Grove – I2C Touch Sensor[e)] (b) Arduino Conductive Wire with Capacitive Sensing Library Supported (for sensing the electrical capacitance of the human body) (c) Arduino UNO R3 (d) Dragino Yun Shield V1.1 (e) SEEED Base Shield V2 Grove for Arduino
Step-on detection	(a) Pressure Sensor – 40 kg Pressure Sensor with HX711 AD Module[f)] (b) Arduino UNO R3 (c) Dragino Yun Shield V1.1 (d) SEEED Base Shield V2 Grove for Arduino

a) http://wiki.seeed.cc/Grove-Temperature_and_Humidity_Sensor_Pro/
b) https://www.arduino.cc/en/main/arduinoBoardUno
c) http://www.dragino.com/products/yunshield/item/86-yun-shield.html
d) http://wiki.seeed.cc/Base_Shield_V2/
e) https://www.seeedstudio.com/Grove-I2C-Touch-Sensor-p-840.html
f) http://www.sunrom.com/p/loadcell-sensor-24-bit-adc-hx711

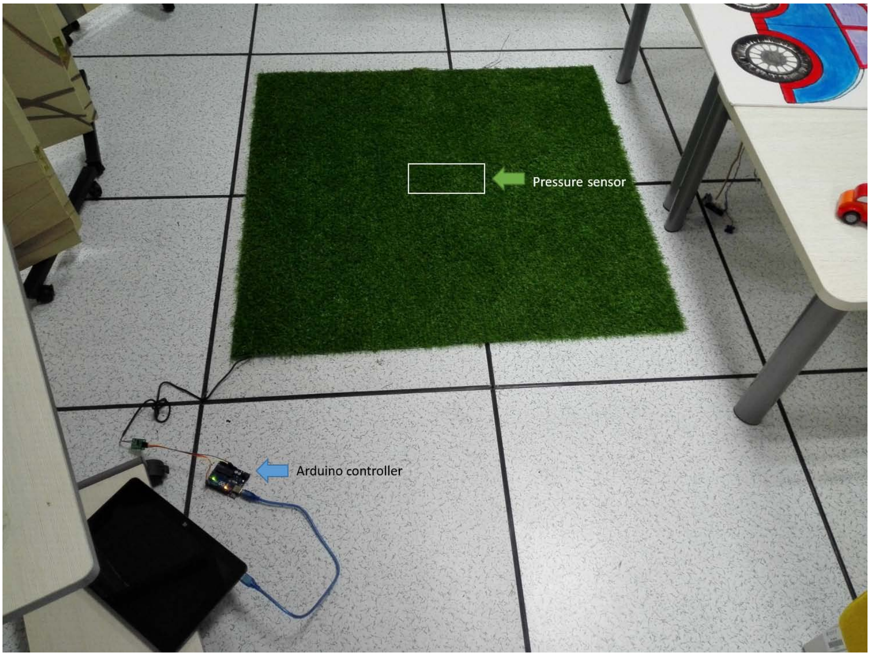

Figure 20.8 The artificial grass with a 40 KG sensor at the center and controlled by an Arduino board.

20.4.3.2 Pressure Sensors: Two Exemplary Play Scenarios

Two types of pressure sensors had been tested in the pilot studies: 40 KG pressure sensor and a capacitive one. The pressure sensor is controlled by an Arduino board (shown in Figure 20.8, at the center of the grass).

Each time the pressure sensor is activated, the self-calibration is conducted to correct the initial self-weight value. Intuitively, the initial pressure value is 0, and after the user steps on the grass (Figure 20.9), the value returned by the pressure sensor will be adjusted based on the level of pressure.

A second type of the pressure sensor has been embedded in the back of the two car wheels of a paper car toy (see Figure 20.6) that provides additional touch points for children during their play.

Figure 20.10 shows one of the typical play scenarios where a tablet is used to visualize the real-time temporal pressure values during the play. Figure 20.7 shows two testing outputs of the before and after the continuous pressures derived from users' interaction with the wheel.

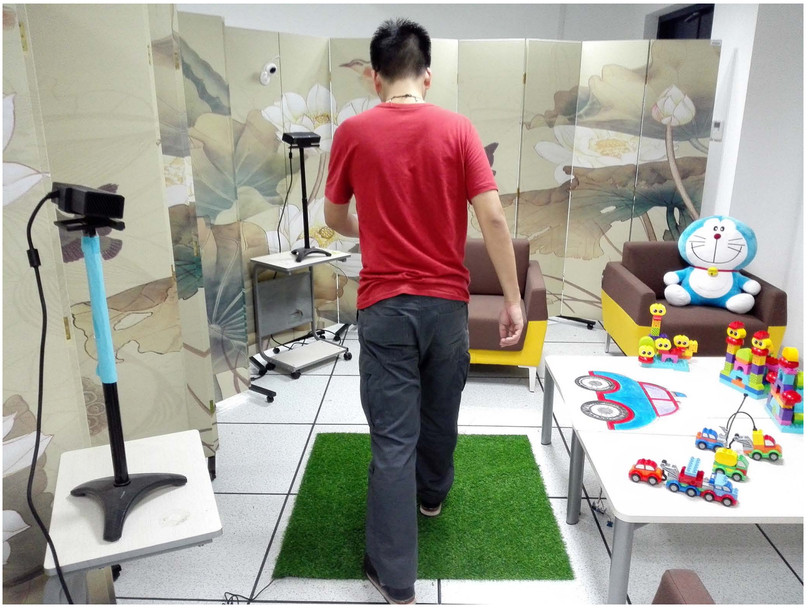

Figure 20.9 The tester stepping on the grass when heading to the front.

Figure 20.10 The play scenario where the tester is observed triggering the capacitive sensors.

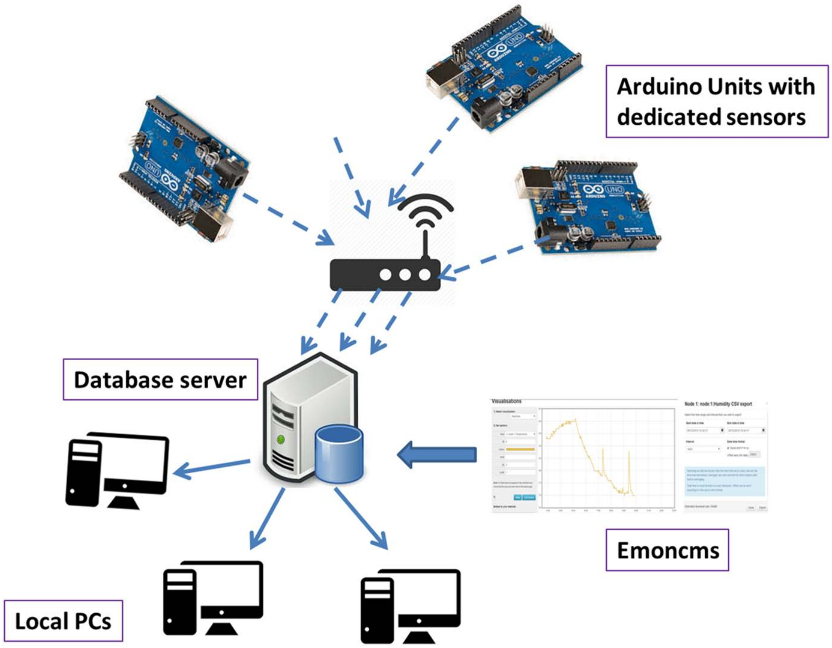

Figure 20.11 An illustration of the data transmission and network management architecture.

20.4.3.3 Data Management and Visualization for Indoor Temperature and Humidity Detection

For the purposes of managing real-time data from multiple sensors, a simple yet efficient data transmission and management system had been set up whose structure is shown in Figure 20.11.

Data capturing can be achieved by expanding a single Arduino board so that it consists of four modules: Arduino Uno, Base Shield, Yun Shield, and the sensors to serve their corresponding purposes (Figure 20.12 and Table 20.2). Data collected by the sensor will be stored at a local server via a Wi-Fi router; therefore, it will allow local computers to access and manipulate it accordingly.

Emoncms,[2] an open-source web-app for data processing and visualization of environmental data, had been deployed on the local database server. The application provides rich functions to store, visualize, and export the sensor data (Figure 20.13). It also supports JSON format making data transmission between the Arduino Yun unit and the database server more feasible and

2 https://emoncms.org/

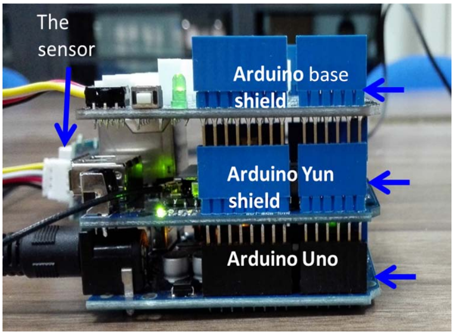

Figure 20.12 The Arduino unit at a closer look.

efficient. At present, the humidity sensor data had been tested with satisfying preliminary results (Figure 20.13).

20.5 The Study and Discussions

20.5.1 Emotion Recognition through Microsoft Kinect

So far, the capacitive sensors placed on paintings, and the pressure sensors on grass, as well as the HD Face SDK in Kinect 2.0 has been tested.

20.5.1.1 The Emotional Facial Action Coding System (EMFACS) and Kinect HD Face API

Facial Action Coding System (FACS) is a widely applied approach for objectively describing the facial animation (Ekman and Friesen, 1978). Six universal basic emotions of happiness, sadness, anger, surprise, disgust, and fear were proposed. Friesen and Ekman further proposed a methodology—Emotional Facial Action Coding System (EMFACS)—to improve the aforementioned universal emotions extraction strategy by directly analyzing the facial activities (Friesen and Ekman, 1983; Kanade et al., 2000; Mao et al., 2015) that was adopted in our current study.

In our preliminary experiment, only happy and sad emotions were studied.

Compared with the old Face Tracking API, HD Face API in Kinect V2 not only supports face *tracking* but also provides face *capture* capabilities, the latter providing high-definition raw facial data nodes for further manipulation (Kinect, 2016). Kinect V2 API natively offers 17 Action Units (AUs); all AUs are assigned with the numeric weight varying from 0 to 1, except for Jaw Slide Right, AU R4, and AU L4 whose values vary from −1 to 1.

It is known that several essential AUs in the FACS algorithm cannot be directly retrieved with the HD Face API; therefore, in our experiment, only AU L12, AU R12 (as Lip Corner Puller Left, Lip Corner Puller Right) are used to infer *happiness*, while AU L4, AU R4, AU L15, and AU R15 (as Left Eye Brow Lowerer, Right Eye Brow Lowerer, Lip Corner Depressor Left, and Lip Corner Depressor Right) are

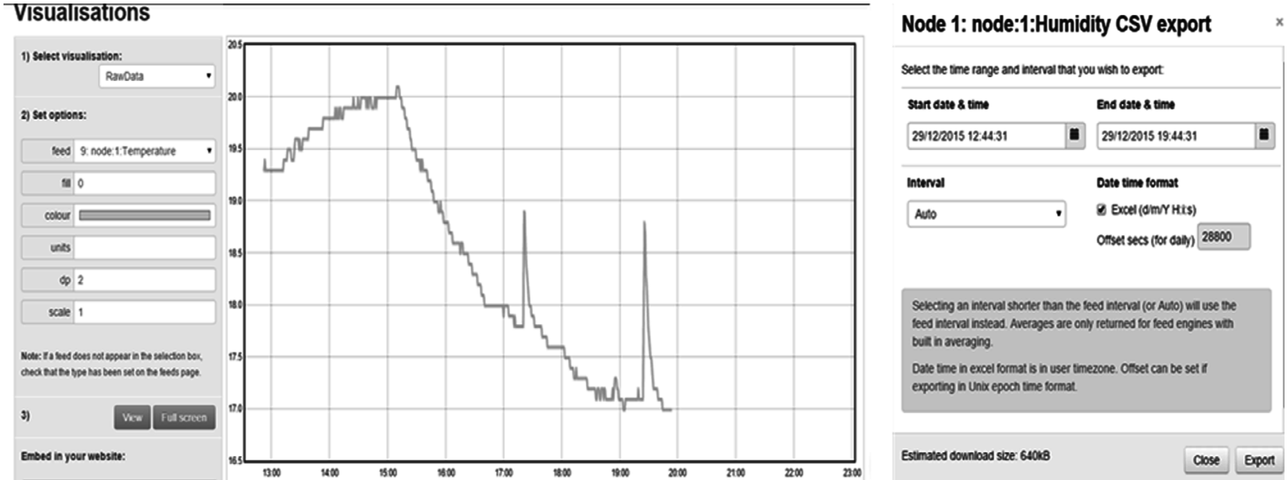

Figure 20.13 The Emoncms user interface where humidity data are visualized and shown.

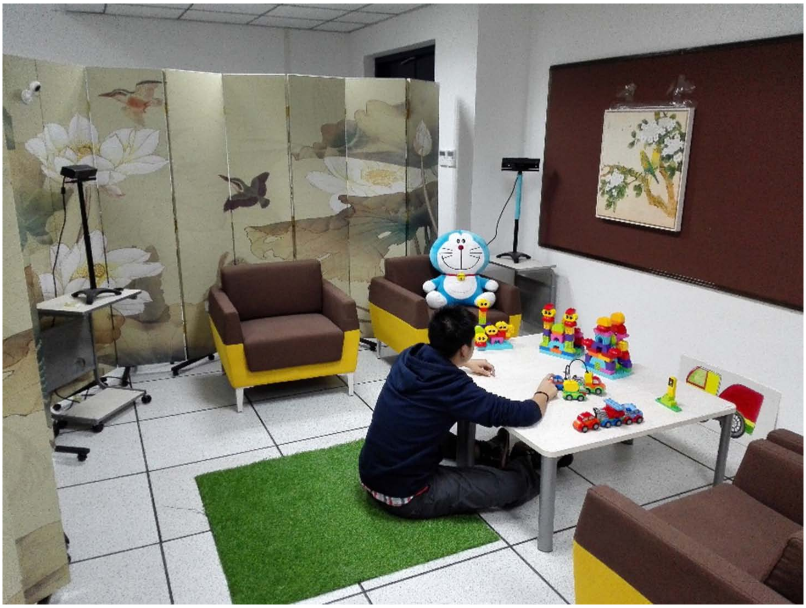

Figure 20.14 The face tracking testing environment where a set of Kinect sensors is positioned to capture player's facial expressions.

applied to infer *sadness*. Additionally, some initial experiments were conducted to retrieve the proper thresholds for the different dedicated AUs and these updated thresholds were fine-tuned before the actual preliminary experiment.

20.5.1.2 Emotion Recognition: Preliminary Testing Results

Due to space limitations, an initial testing on happy and sad emotions will be reported here. During the preliminary experiment, both sitting and standing status were examined. In each type of experiment, an NT tester was instructed to play the toys laid out on the table freely where images of testing moment in the sitting status are captured by Kinect V2 sensor (Figures 20.14 and 20.15).

At present, the objective of our preliminary testing is merely to find the minimum weights used in the calculation of different AUs so that we could establish a relationship between the combinations of these AUs and the corresponding emotion labels. We tested both sitting and standing positions in our preliminary experiments. In each type of experiment, an NT tester was instructed to play the toys laid out on the table freely (Figure 20.11). During each play, the tester was randomly instructed (for testing purposes) to express a happy or sad emotion without inducing his/her moods. Figure 20.16 shows the user interface of the module.[3]

3 The testing videos can be watched at: https://www.youtube.com/watch?v=8Y17gj7zrx4&feature=youtu.be and https://www.youtube.com/watch?v =ZSyR1AEJg4s&feature=youtu.be

Figure 20.15 The face tracking testing environment where the tester can be seen sitting while playing the LEGO blocks.

20.5.2 Emotion Visualization and Broadcasting through Affective Object

After obtaining the emotion label, instead of simply returning the emotion label to the user, the emotion label was visualized and broadcasted through colored stripes (referred to as *affective objects* (Scheirer and Picard, 2000)) mainly due to the following reasons (Figure 20.2):

- Individuals with ASD tend to appear very different outwardly than what they might really be (through their behavioral and/or interactions in a social/individual environment) (Picard, 2009); thus, it is vital to inform others of their emotional state.

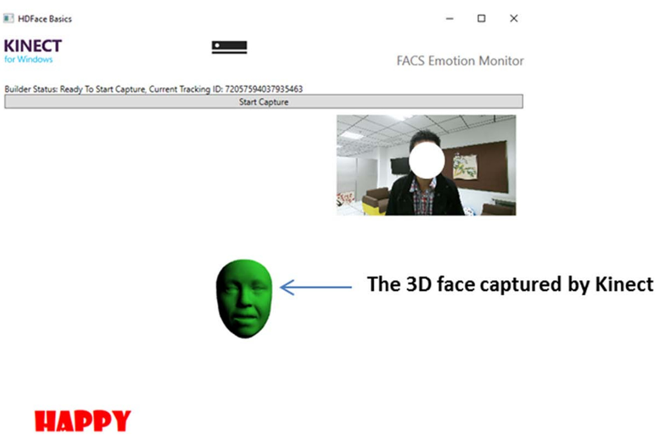

Figure 20.16 The Kinect HD facial capture moment.

- Previous studies have shown that many individuals prefer not to "broadcast" their emotion/mood to others due to the privacy concern except for individuals with ASD who have difficulties expressing their own emotions (Williams et al., 2015).
- Identifying and broadcasting the emotion states (of individuals either with ASD or TD) to others (again including both types) are vital for successful and effective communication (Williams et al., 2015; Picard, 2009)

In summary, instead of "announcing" the mood of the child to all the people in the environment, such emotion display through "light-up table" is considered to be more private—although it offers enough clues to other individuals to act during social interactions. The affective object designed in our experiment is a neo-pixel stripe embedded under a play table (Figure 20.17) that contains three layers: The neo-pixel stripe was positioned on top of an ordinary play table and covered with an acrylic board on which children can play with toys.

Our emotion visualization and broadcast module consists of two basic functional units: emotion tracking and responsive units, as shown in Figure 20.18.

The table is controlled through the Arduino Uno board that is linked to a local PC. Through the digital pin, the Arduino controller sends commands to the stripe to display a corresponding emotion color (see Table 20.3 for the displayed emotion color scheme used in the system (Nijdam, 2009)).

Figure 20.19 captures one of the testing moments when the player was happy in playing.

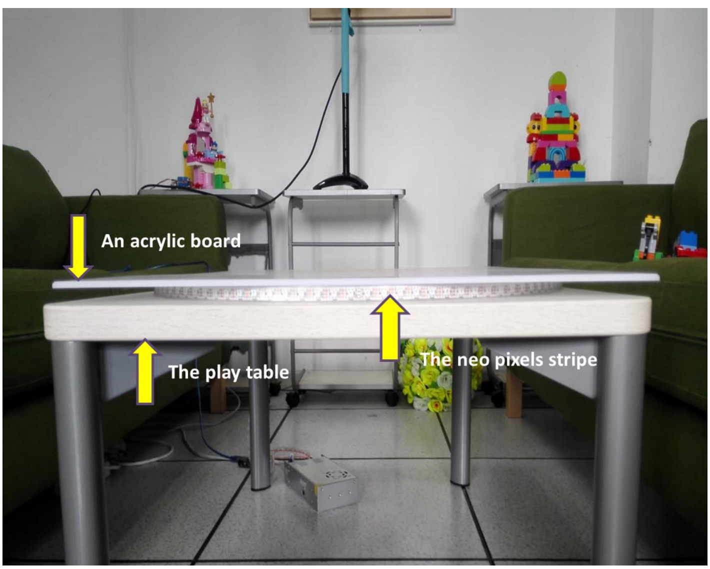

Figure 20.17 The play-table embedded with the LED stripe (three layers).

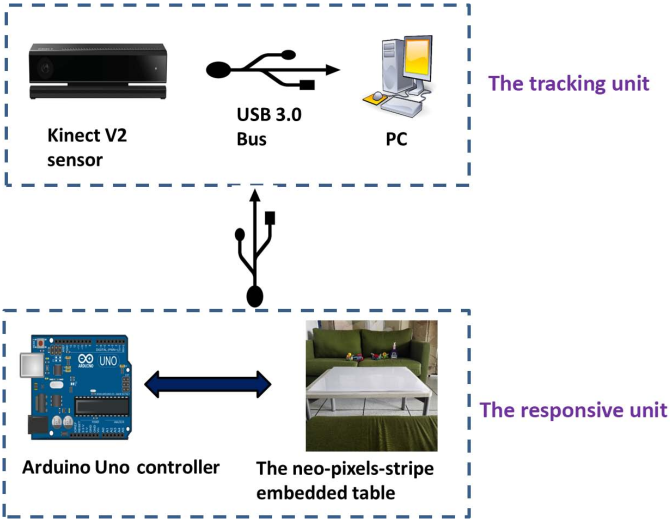

Figure 20.18 Two functional units in the emotion visualization and broadcast module.

Table 20.3 Correlation between each emotion and the displayed color.

Emotion	Corresponding color	RGB value
Happy	Orange red	(255,69,0)
Sad	Light blue	(34,219,182)
Fear	Purple	(128,0,128)
Surprise	Hot pink	(255,105,180)
Neutral	Light green	(144,238,144)

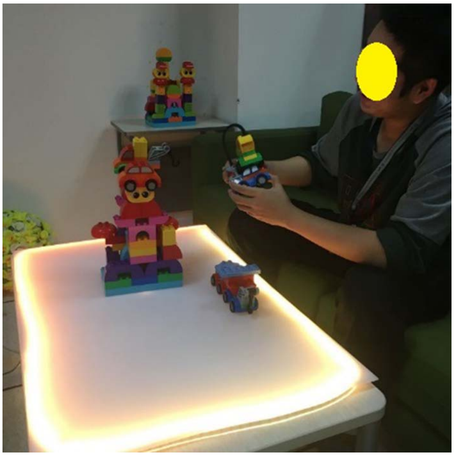

Figure 20.19 The happy moment broadcasted through the neo-pixel stripe-embedded play table.

20.6 Conclusions

In this chapter, the integration of a naturalistic multisensory IoT play environment has been presented; such an integration aims at capturing emotion-related behavioral, physiological, and ambient environment data for emotion recognition. The long-term goal of such a platform is to help NT individuals “read” the emotions of children with ASD. By adopting IoT, it may reduce direct intervention and interruption during children’s playing moment; hence, children could enjoy their activities more naturally. The chapter presented the system design and offered some early insights derived from the conducted testing sessions on the potential of such environment.

At present, some sensors (including the facial data captured by the Kinect sensor) are being tested separately. Achieving data fusion (i.e., integrating data captured from the different sensors) for the purposes of generating meaningful emotional label is one of our future challenges. Meanwhile, we have also used and tested the Emotion API (part of the Microsoft Cognitive Services[4]) to label user’s emotions: Although the application has been successfully experimented on NT individuals, how or whether it can be applied to children with ASD is unclear. As we have already mentioned, inferring the emotion of those with ASD is very challenging (Tang, 2016, Tang et al., 2017a, 2017b); therefore, it could be very useful if we could have a training data set obtained from children with ASD that includes their indicative emotional labels and behavioral patterns. One avenue to pursue is to collect emotional behaviors of children with ASD by inducing their emotion; another is to invite caregivers and parents living with them to rate the generated labels from natural events (e.g., playing moments). Either way, since the interest is to help caregivers and parents to better recognize their children’s emotional behaviors in daily activities, the natural setting is very important here. Inferences based on other emotional behaviors (e.g., heart rate and perspiration) should be generalizable from NT individuals, since those characteristics are sympathetic nervous responses.

Moreover, there are two immediate questions that require our attention here: (1). How will the children’s emotion be conveyed to others (e.g., synthetic speech, LEDs, etc.)? (2). Does the severity of disorder influence the children’s emotional affect? If yes, how can this be incorporated into the existing models? Regardless of these two questions, current work has demonstrated an early effort to build an IoT-based natural play environment that would allow NT individuals to better recognize ASD children’ emotion with the goal to open up another possibility to establish better social interactions between them.

4 https://www.microsoft.com/cognitive-services

Acknowledgments

The authors would like to acknowledge the financial support from the Academic Affairs Office for this project. Thanks also go to Relic Yongfu Wang for implementing the system when he was an undergraduate student at the University, and Odd Aonan Guan for his time and efforts in the experiments. We greatly acknowledge our editor, Dr. Qusay F. Hassan, for his insightful comments and efforts to help improve the readability of this chapter.

References

Baron-Cohen, S., Spitz, A., and Cross, P. (1993) Can children with autism recognize surprise? *Cognition and Emotion*, **7**, 507–516.

Baron-Cohen, S., Wheelwright, S., and Jolliffe, T. (1997) Is there a “language of the eyes”? Evidence from normal adults and adults with autism or Asperger syndrome. *Visual Cognition*, **4**, 311–332.

Beykikhoshk, A., Arandjelovic, O., Phung, D., Venkatesh, S., and Caelli, T. (2014) Data-mining twitter and the autism spectrum disorder: a pilot study. Proceedings of the ASONAM’2014, Calgary, Canada, IEEE Press, pp. 349–356.

Boone, R.T. and Cunningham, J.G. (1998) Children’s decoding of emotion in expressive body movement: the development of cue attunement. *Developmental Psychology*, **34**, 1007–1016.

Brewer, R., Biotti, F., Catmur, C., Press, C., Happé, F., Cook, R., and Bird, G. (2016) Can neuro-typical individuals read autistic facial expressions? Atypical production of emotional facial expressions in autism spectrum disorders. *Autism Research*, **9**(2), 262–271.

Camurri, A., Lagerlof, I., and Volpe, G. (2003) Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, **59**, 213–225.

Carter, E.J. and Hyde, J. (2015) Designing autism research for maximum impact. Proceedings of the ACM CHI’2015, Seoul, Korea, ACM Press, pp. 2801–2804.

Castellano, G., Villalba, S.D., and Camurri, A. (2007) Recognising human emotions from body movement and gesture dynamics. Proceedings of the ACII 2007, Lisbon, Portugal, LNCS 4738, pp. 71–82.

Czapinski, P. and Bryson, S.E. (2003) Reduced facial muscle movements in autism: evidence for dysfunction in the neuromuscular pathway? *Brain and Cognition*, **51**(2), 177–179.

De Meijer, M. (1989) The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, **13**, 247–268.

Ekman, P. and Friesen, W.V. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologist Press, Palo Alto, CA.

Ertin, E., Raij, A., Stohs, N., al’Absi, M., Kumar, S., and Shah, S. (2011) AutoSense: unobtrusively wearable sensor suite for interring the onset, causality and consequences of stress in the field. Proceedings of the ACM SenSys’11, Seattle, WA, ACM Press, pp. 274–287.

Faso, D.J., Sasson, N.J., and Pinkham, A.E. (2015) Evaluating posed and evoked facial expressions of emotion from adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, **45**(1), 1–15.

Friesen, W.V., and Ekman, P. (1983) EMFACS-7: Emotional Facial Action Coding System. Unpublished manual, University of California, California.

Grossman, R.B., Edelson, L.R., and Tager-Flusberg, H. (2013) Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *Journal of Speech, Language and Hearing Research*, **56**(3), 1035–1044.

Gunes, H. and Piccardi, M. (2007) Bi-modal emotion recognition from expressive face and body gestures. *Journal of Network and Computer Applications*, **30**(4), 1334–1345.

Halberstadt, A.G., Denham, S.A., and Dunsmore, J.C. (2001) Affective social competence. *Social Development*, **10**(1), 9–119.

Harms, M.B., Martin, A., and Wallace, G.L. (2010) Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychology Review*, **20**(3), 290–322.

Hong, H., Gilbert, E., Abowd, G.D., and Arriaga, R.I. (2015) In-group questions and out-group answers: crowdsourcing daily living advice for individuals with autism. Proceedings of the ACM CHI’2015, Seoul, Korea, ACM Press, pp. 627–636.

Kanade, T., Cohn, J.F., and Tian, Y. (2000) Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, IEEE Press, pp. 46–53.

Kapur, A., Kapur, A., Babul, N.V., Tzanetakis, G., and Driessen, P.F. (2005) Gesture-based affective computing on motion capture data. Proceedings of the ACII 2005, Beijing, China, pp. 1–7.

Karimi, K. (2013) The Role of Sensor Fusion and Remote Emotive Computing (REC) in the Internet of Things. Freescale Semiconductor, SENFEIOTLFWP REV 1, White Paper.

Kinect (2016) High definition face tracking. Available at https://msdn.microsoft.com/en-us/library/dn785525.aspx (retrieved October 9, 2016).

Krishna, S., Bala, S., McDaniel, T., McGuire, S., and Panchanathan, S. (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proceedings of the ACM CHI’2010, Atlanta, GA, pp. 3637–3642.

Macdonald, H., Rutter, M., Howlin, P., Rios, P., Conteur, A.L., Evered, C., and Folstein, S. (1989) Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology, Psychiatry, Allied Disciplines*, **30**(6), 865–877.

Mao, Q.R., Pan, X.Y., Zhan, Y.Z., and Shen, X.J. (2015) Using Kinect for real-time emotion recognition via facial expressions. *Frontiers of Information Technology & Electronic Engineering*, **16**, 272–282.

Mitchell, T. (2009) Mining our reality. *Science*, **326**, 1644–1645.

Nijdam, N.A. (2009) *Mapping Emotion to Color*, University of Twente, The Netherlands.

Park, J.H., Abirached, B., and Zhang, Y. (2012) A framework for designing assistive technologies for teaching children with ASDs emotions. Proceedings of the ACM CHI'2012, Austin, TX, pp. 2423–2428.

Peterson, C.C., Slaughter, V., and Brownell, C. (2015) Children with autism spectrum disorder are skilled at reading emotion body language. *Journal of Experimental Child Psychology*, **139**, 35–50.

Picard, R. (2009) Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B*, **364**, 3575–3584.

Pollick, F., Paterson, H., Bruderlin, A., and Sanford, A. (2001) Perceiving affect from arm movement. *Cognition*, **82**, 51–61.

Rehg, J.M., Rozga, A., Abowd, G.D., and Goodwin, M.S. (2014) Behavioral imaging and autism. *IEEE Pervasive Computing*, **13**(2), 84–87.

Robinson, P. (2014) Computation of Emotions. Proceedings of the ACM ICMI 2014, Istanbul, Turkey, pp. 409–410.

Scheirer, J. and Picard, R. (2000) Affective objects. MIT Media Laboratory Perceptual Computing Section, Technical Report No. 524.

Seo, J.H., Sungkajun, A., and Suh, J. (2015) Touchology: towards interactive plant design for children with autism and older adults in senior housing. Proceedings of the CHI 2015 Seoul, Korea, pp. 893–898.

Simmons, D.R., Robertson, A.E., McKay, L.S., Toal, E., McAleer, P., and Pollick, F.E. (2009) Vision in autism spectrum disorders. *Vision Research*, **49**(22), 2705–2739.

Stagg, S., Slavny, R., Hand, C., Cardoso, A., and Smith, P. (2014) Does facial expressivity count? How typically developing children respond initially to children with autism. *Autism*, **18**(6), 704–711.

Sula A., Spaho E., Matsuo K., Barolli L., Xhafa F., and Miho R. (2013) An IoT-based framework for supporting children with autism spectrum disorder, in Park, J., Barolli, L., Xhafa, F., Jeong, H.Y. (eds.), *Information Technology Convergence*. Lecture Notes in Electrical Engineering, vol. 253, Springer, Dordrecht, The Netherlands.

Sun, X., Allison, C., Auyeung, B., Zhang, Z., Matthews, F., Baron-Cohen, S., and Brayne, C. (2015) Validation of existing diagnosis of autism in mainland China using standardised diagnostic instruments. *Autism Research*, **9**(8), 1010–1017.

Tang, T. (2016) Helping neuro-typical individuals to “Read” the emotion of children with autism spectrum disorder: an Internet-of-Things Approach. Proceedings of the ACM IDC’2016, Manchester, UK, pp. 666–671.

Tang, T. and Flatla, D. (2016) Autism awareness and technology-based intervention research in china: the good, the bad, and the challenging. Proceedings of the Workshop on Autism and Technology: Beyond Assistance & Intervention, C in conjunction with the *ACM CHI’2016*, San Jose, CA.

Tang, T., Wang, Y. You, Y., Huang, Z., and Chen, P. (2015) Supporting collaborative play via an affordable touching + singing plant for children with autism in China. Proceedings of the ACM UbiComp/ISWC’15, Osaka, Japan, ACM Press, pp. 373–376..

Tang, T., Falzarano, M., and Morreale, P.A. (2017a) Assessment of the utility of gesture-based application for the engagement of Chinese children with autism. Universal Access in the Information Society, Special Issue on Information Society Skills. Is knowledge Accessible for All? Springer.

Tang, T., Winoto, P., and Chen, G.X. (2017b) Emotion recognition via face tracking with RealSense™ 3D camera for children with autism. Proceedings of the ACM IDC’2017, Stanford, CA, pp. 533–539.

Tartaro, A., Cassell, J., Ratz, C., Lira, J., and Nanclares-Nogués, V. (2015) Accessing peer social interaction: using authorable virtual peer technology as a component of a Group Social Skills Intervention Program. *ACM Transactions on Accessible Computing (TACCESS)*, **6**(1), 1–29.

Tyka, A., Palka, T., Tyka, A., Cison, T., and Szygula, Z. (2009) The influence of ambient temperature on power at anaerobic threshold determined based on blood lactate concentration and myoelectric signals. *International Journal of Occupational Medicine and Environmental Health*, **22**, 1–6.

Virnes, M., Kama, E., and Vellonen, V. (2015) Review of research on children with autism spectrum disorder and the use of technology. *Journal of Special Education Technology*, **30**(1), 13–27.

Wang, S. (2015) In China, the making of an app for autism. The Wall Street Journal. Available at http://blogs.wsj.com/chinarealtime/2015/05/19/in-china-the-making-of-an-app-for-autism/ (accessed October 17, 2016).

Weimer, A., Schatz, A., Lincoln, A., Ballantyne, A., and Trauner, D. (2001) “Motor” impairment in asperger syndrome: evidence for a deficit in proprioception. *Journal of Developmental & Behavioral Pediatrics*, **22**(2), 92–101.

Williams, M.A., Roseway, A., O’Dowd, C., Czerwinski, M., and Morris, M.R. (2015) SWARM: an actuated wearable for mediating affect. Proceedings of the TEI’2015, Stanford, CA, ACM Press, pp. 293–300.

Winoto, W. (2016) Reflections on the adoption of virtual reality-based application on word recognition for Chinese children with autism. Proceedings of the ACM IDC’2016, Manchester, UK, ACM Press, pp. 589–594.

Winoto, P., Chen, G.X., and Tang, T. (2016a) The development of a Kinect-based online socio-meter for users with social and communication skill impairments: a computational sensing approach. Proceedings of the IEEE International Conference on Knowledge Engineering and Applications (ICKEA'2016), Singapore, IEEE Press, pp. 139–143.

Winoto, P., Xu, N., and Zhu, A. (2016b) “Look to Remove”: a virtual reality application on word learning for Chinese children with autism. Proceedings of the 2016 Human-Computer Interaction International Conference (HCII'2016), Toronto, Canada, Springer, pp. 257–264.

World Health Organization (1993) *International Statistical Classification of Diseases and Related Health Problems*, 10th edn (ICD-10), World Health Organization, Geneva.

Zwaigenbaum, L., Bauman, M.L., Choueiri, R., Fein, D., Kasari, C., Pierce, K., Stone, W.L., Yirmiya, N., Estes. A., Hansen, R.L., McPartland, J.C., Natowicz, M.R., Buie, T., Carter, A., Davis, P.A., Granpeesheh, D., Mailloux, Z., Newschaffer, C., Robins, D., Roley S.S., Wagner, S., and Wetherby, A. (2015) Early identification and interventions for autism spectrum disorder: executive summary. *Pediatrics*, **136** (Suppl. 1), S1–S9.