Early Diagnosis of Autism through Analysis of Pre-speech Vocalizations

Keshi Dai

College of Computer and Information Science, Northeastern University, Boston, USA

daikeshi@ccs.neu.edu

Introduction

Autism is a complex neurodevelopmental disorder that typically appears before age of three years old. The narrower term of autism always refers to classic autistic disorder in the spectrum of pervasive developmental disorder (PDD). In this paper, "autism" and "autistic" are all used to refer to the entire range of autism spectrum disorders (ASD).

Although autism was first described in 1943 by Kanner, there is still no quantitative method to accurately diagnose it. Currently, autism can be reliably diagnosed by the age of 3. However, studies suggest that many children with autism may be accurately identified at 12 months or even younger [1]. The most widely used diagnostic methods nowadays rely on evaluation by professionals of behavioral characteristics. This kind of diagnosis usually requires a two-stage process that includes developmental screening and comprehensive evaluation by a multidisciplinary team [1]. In the developmental screening, parents need to recall when their children reached each milestone in their development through reviewing video recordings, photos, and so on. In the next step, a screening instrument is used to assess whether children are typically developing based on parent reports and observations. Among screening instruments available today are Checklist of Autism in Toddlers (CHAT) [2], the Screening Tool for Autism in Two-Year-Olds (STAT) [3], and the Social Communication Questionnaire (SCQ) that is for 4 years old children or older [4]. These four are mainly to identify children with severe autism. They may not be able to detect some mild ASD. Other screening instruments like Autism Spectrum Screen Questionnaire (ASSQ) [5] and Childhood Asperger Syndrome Test (CAST) [6] are designed to identify mild ASD like high-functioning or Asperger syndrome. If some behaviors of a child in the screening process are identified as indicators of ASD, this child will need a comprehensive diagnostic evaluation. This evaluation is performed by a multidisciplinary team that includes a psychologist, a neurologist, a psychiatrist, a speech therapist, or other professionals in ASD. Apart from traditional diagnostic methods, new technologies such as movement analysis [7] and eye gaze analysis [8] are being developed as potential clinical tools in the diagnosis of autism.

Early diagnosis of autism via pre-speech vocalization

Our purpose is to develop an objective clinical tool for the early diagnosis of autism. We will use technologies developed in Early Vocalization Analyzer (EVA) and visiBabble system to extract features from vocalizations of infants and toddlers from 6 to 18 months old. By focusing on a population with moderate risk for the development of autism, we hope to be able to derive classification rules, based on these detected speech features, which distinguish those children who are eventually diagnosed with autism from other children in the group. We hope that our classifier can be used, in the future, for early diagnosis of autism and therefore for early intervention.

EVA and VisiBabble

EVA is a computer program that automatically analyzes infant vocalizations and derives a "vocalization age" based on detected features. It is able to clinically distinguish infants, six to fifteen months old, who may be at risk for later communication or other developmental problems from typically developing infants [9, 10]. The visiBabble system relies on EVA to process vocalizations in real-time. It responds to the infant's syllable-like productions with visual feedbacks and records the landmark analysis. This system reinforces the production of syllabic utterances that are associated with later language and cognitive development [11].

Technical Methods

We first find landmarks in the acoustic signal and then use them to extract other features. Landmark detector used in our work is built on the Liu-Stevens landmark detection theory [12]. Essential to this theory are landmarks, pinpointing the abrupt spectral changes in an utterance, which mark perceptual foci and articulatory targets. Listeners often focus on landmarks to obtain acoustic cues necessary for understanding the distinctive features in the speech.

In this work, we focuses on three types of landmarks: glottis (marks the time when the vocal folds transition from freely vibrating to not freely vibrating), sonorant (marks sonorant consonantal closures and releases), and burst (marks stop/affricate bursts and points where aspiration/frication ends due to stop closure) [9, 10]. We used three measurements related to landmarks: landmarks per word or utterance, voice onset time that is the time between when a consonant is released and when the vibration of the vocal folds begins, and landmark rate that is the rate of each landmark type in an utterance. Figure 1 shows landmarks produced by our landmark extractor.

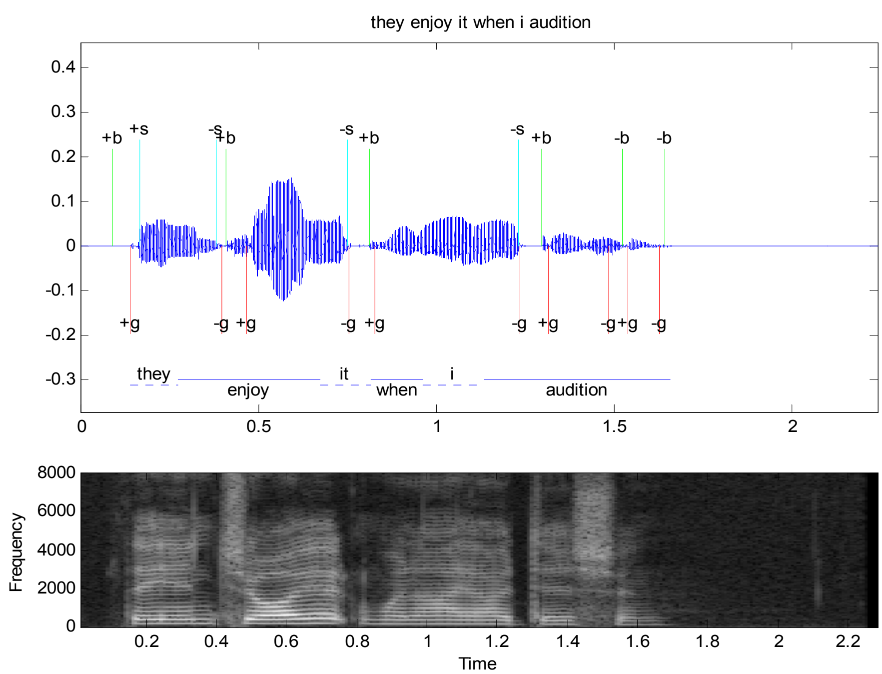

Figure 1: Landmark plot produced by our landmark detector

In Figure 1, we can see that the region between +g and –g is the voiced region. +s/-s landmarks only happen in voiced region, and +b/-b landmarks only appear in unvoiced region. In the spectrogram, the energy of the fundamental frequency in voiced region is the strongest. The +s landmark happens when there is an increase in energy from the Bands 2 (0.8-1.5 kHz) to Bands 5 (3.5-5.0 kHz) and the –s landmark signifies energy decrease in these frequency bands. A +b landmark is detected when a silence interval is followed by a sharp energy increase in high frequency from Bands 3 (1.2-2.0 kHz) to Bands 6 (5.0-8.0 kHz). On the contrary, a –b landmarks signifies a sharp energy decrease in high frequency followed by a silence interval.

Sequences of landmarks are grouped into standard syllable patterns. A syllable is a unit of sound, and is typically made up of a vowel with optional initial and final margins. In our syllable detector, syllable is based on the order and spacing of detected landmarks. It must contain a voiced segment of sufficient length. 38 possible syllables were recognized. 11 syllables begin with +g landmark, 22 begin with +b, and 5 begin with +s. 4 types of syllable features are extracted: syllable rate, syllable number, landmarks per syllable, and syllable duration.

With above-mentioned landmark and syllable features, we can construct new features for early diagnosing autism. Specific patterns of landmarks, particular types of syllable, and anomalous pitch patterns in the utterance may also be important indicators. In addition, features used to estimate vocalization age in [10] may be also useful. Vocalization age delay may be a possible indicator of autism, because autism is also characterized by impaired social interaction and communication.

Planned Study

Subjects

We plan to work with a group at Children's Hospital, Boston, which is studying younger siblings of children with autism. We plan to record these children for 10 minutes every 3 months from one to two years. Features will be extracted from these recordings and we will look for patterns that can distinguish those children who eventually develop autism from those who don't.

Data Processing and Analysis

First we will examine the recordings we collected. We will remove utterances not produced by our subjects, non-speech utterances, and other extraneous noise or voices, and only keep clear babbling. Then, landmark detector will be used to extract landmarks. Syllables and utterances analysis will also be applied to landmarks. We will record landmarks, syllables, and utterances for further analysis. After children are clinically diagnosed as autism (at 2.5 to 3 years), a comprehensive analysis will be conducted based on landmarks, syllables, and utterances collected before. In this step, we hope to find features that distinguish those children who eventually develop autism from those who don't.

Envisioned Contributions

Current research estimates that 2-6 per 1000 children have an ASD in the United States [1]. It is a surprising increase over rates reported in 1980s and 1990s. Our method makes early diagnosis of autism at pre-speech age possible. Early diagnosis allows for early intervention, which is critical, because children with autism who receive early intervention at younger ages make greater improvements than those who receive it at older ages [13, 14].

Early intervention can prevent declines in intellectual development, and improve the communication skills, motor, and cognitive development [15, 16]. Early intervention should commence as soon as possible after diagnosis. It will have a great impact on the lives of children with autism and their families.

References

1. Autism Spectrum Disorder (Pervasive Developmental Disorders). National Inst. of Mental Health, 2004.
2. Baird, G., Charman T., Baron-Cohen, S., Cox, A., Swettenham J., Wheelwright, S., and Drew, A. A screening instrument for autism at 18 months of age: A 6-year follow-up study. Journal of the American Academy of Child and Adolescent Psychiatry, 39 (2000), 694-702.
3. Stone, W. L., Coonrod, E. E., and Ousley, O. Y. Brief report: Screening tool for autism in two-year-olds (STAT): development and preliminary data. Journal of Autism and Developmental Disorders, 30, 6 (2000), 607-612.
4. Berument, S. K., Rutter, M., Lord C., Pickles, A., and Bailey, A. Autism Screening Questionnaire: diagnostic validity. British Journal of Psychiatry, 175 (1999), 444-451.
5. Ehlers, S., Gillberg, C., and Wing, L. A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. Journal of Autism and Developmental Disorders, 29, 2 (1999), 129-141.
6. Scott, F. J., Baron-Cohen, S., Bolton, P., and Brayne, C. The Cast (Childhood Asperger Syndrome Test): preliminary development of a UK screen for mainstream primary-school-age children. Autism, 2, 1 (2002), 9-31.
7. Teitelbaum, P., Teitelbaum, O., Nye, J., Fryman, J., and Maurer, R. G. Movement analysis in infancy may be useful for early diagnosis of autism. Psychology, 95, 23 (Nov. 1998), 13982-13987.
8. Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. General Psychiatry, 59, 9 (Sep. 2002), 809-816.
9. Fell, H. J., MacAuslan, J., Ferrier, L. J., and Chenausky, K., Automatic Babble Recognition for Early Detection of Speech Related Disorders. J. Behaviour & Inf. Tech., 18, 1 (1999), 56-63.
10. Fell, H. J., MacAuslan, J., Ferrier, L. J., Worst, S. G., and Chenausky, K. H., Vocalization Age as a Clinical Tool. In Proceedings of ICSLP ‘02, (Denver, CO, Sep. 2002), 2345-2348.
11. Fell, H. J., Cress, C., MacAuslan, J., and Ferrier, L. J., visiBabble for reinforcement of early vocalization. In Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility, (Atlanta, GA, Oct. 18-20, 2004), 161-168.
12. Liu, S. Landmark Detection for Distinctive Feature-based Speech Recognition. Ph.D. Thesis. M.I.T. Cambridge, MA, 1995.
13. Harris, S.L., and Weiss, M. J. Right from the start: Behavioral intervention for young children with autism. Woodbine House, Bethesda, MD, 1998
14. Sheinkopf, S. J., and Siegel, B. Home based behavioral treatment of young children with autism. Journal of Autism and Developmental Disorders, 28, 1 (1998), 15-23.
15. Guralnick, M.J. Effectiveness of early intervention for vulnerable children: A developmental perspective. American Journal on Mental Retardation, 102, 4 (1998), 319–345.
16. McGee, G. G., Morier, M. J., and Daly, T. The Walden Early Childhood Programs. In Handleman, J. S., Harris, S. L. Preschool Education Programs for Children with Autism 2nd ed. Pro-Ed, Austin, TX, 2001.

About the author

Keshi Dai is a second-year Ph.D. student in the College of Computer and Information Science of Northeastern University. His research interests include speech processing, early vocalization analysis, and emotion recognition. Currently, he is planning to work with a group at Children's Hospital, Boston to collect utterances of younger siblings of children with autism. He hopes to find features that can distinguish those children who eventually develop autism from those who don't.