ON QUANTIFYING FACIAL EXPRESSION-RELATED ATYPICALITY OF CHILDREN WITH AUTISM SPECTRUM DISORDER

Tanaya Guha¹, Zhaojun Yang¹, Anil Ramakrishna¹, Ruth B. Grossman², Darren Hedley², Sungbok Lee¹, Shrikanth S. Narayanan¹

¹ Signal Analysis and Interpretation Lab (SAIL), University of Southern California, Los Angeles

² Emerson College, University of Massachusetts Medical School, MA
• Introduction
  • Problem statement, motivation and goals

• Database
  • Data collection
  • Preprocessing

• Data Analysis
  • Complexity analysis
  • Region-based analysis

• Conclusion
• 1 in 68 children has autism spectrum disorder (ASD) in US
• 1% prevalence in Asia, Europe, North America, 2.6% in South Korea
• 1 in 68 children has autism spectrum disorder (ASD) in US
• 1% prevalence in Asia, Europe, North America, 2.6% in South Korea

• Children with ASD
  • often preserve IQ, language and cognitive skills
  • have significantly impaired social communication skills
• 1 in 68 children has autism spectrum disorder (ASD) in US
• 1% prevalence in Asia, Europe, North America, 2.6% in South Korea

• Children with ASD
  • often preserve IQ, language and cognitive skills
  • have significantly impaired social communication skills

• Facial expressions of children with ASD
  • perceived *awkward* or *unusual* by observers
  • awkwardness is perceived just after 1 sec
  • clinically accepted measure of autism
• Psychologists rely on visual inspection
• Psychologists rely on visual inspection
• Little computational effort to understand atypicality in facial gesture
• Psychologists rely on visual inspection
• Little computational effort to understand atypicality in facial gesture
• Computational work quantifying atypicality in prosody [Bone et al. 2012], speech and body gesture [Yang & Narayanan 2014]
Psychologists rely on visual inspection
Little computational effort to understand atypicality in facial gesture
Computational work quantifying atypicality in prosody [Bone et al. 2012], speech and body gesture [Yang & Narayanan 2014]

**Goal:**
To analyze and quantify perceived atypicality in facial-expressions of children with autism.
• Introduction
  • Problem statement, motivation and goals

• Database
  • Data collection
  • Preprocessing

• Data Analysis
  • Complexity analysis
  • Region-based analysis

• Conclusion
**MOTION CAPTURE (MOCAP) DATA**

- Mocap technique to record facial motion
- 24 ASD, 21 typically developing (TD) participants
- 32 facial markers (4 stability markers) record (x,y,z) locations
- Mimic facial expressions shown in stimuli video
- Each participant mimics 18 expressions (e.g. smile, tearful, excited)

*Figure: (left) Facial marker positions, and (right) sample frame from stimuli video for expression *smile*
PREPROCESSING

- Artifacts removal, missing data interpolation
- Head motion removal
- Face normalization to remove subject-specific structural variability
- For each task performed by each participant, time-series data from 28 channels is used for further analysis

[Metallinou et al, Proc. ICME 2013]
• Introduction
  - Problem statement, motivation and goals

• Database
  - Data collection
  - Preprocessing

• Data Analysis
  - Global analysis
  - Region-based analysis

• Conclusion
• Knowledge-driven grouping of the tasks
  emotion groups: anger, disgust, fear, joy, sadness and surprise
• Knowledge-driven grouping of the tasks
  emotion groups: anger, disgust, fear, joy, sadness and surprise

• Differences between ASD and TD for each emotion group
  • Global analysis: Dynamical complexity
    multivariate multiscale entropy (MMSE)
  • Local analysis: region-based dynamics
    psychology-inspired measures
    time series modeling
- Reduced complexity in physiological systems under disease or disorder
- Facial expression as a multivariate dynamical system
- Complexity is measured in terms of Multivariate Sample Entropy
  

- Detects dynamic structures within and across channels at multiple temporal scales.
Consider a 2-channel system: \([x(1), x(2), \ldots, x(T)], [y(1), y(2), \ldots, y(T)]\)

- Scale \(\epsilon = 1\), time lag vector \(= [1,1]\), embedding vector \([1, 1]\),
- Compute all possible composite delay vectors: 
  \(C_1 = [x(1), y(1)]\), 
  \(C_2 = [x(2), y(1)]\), ...
- \(B^m(r) = \text{fraction of the composite vector pairs for which } d(C_i, C_j) \leq r\), where \(i \neq j\)
- \(B^{m+1}(r) = \text{same measure for } (m+1)\text{-dim space}\)
- Includes vector pairs within and across the embedded subspaces

\[
\text{Multivariate sample entropy} = - \ln \frac{B^{m+1}(r)}{B^m(r)}
\]
DYNAMICAL COMPLEXITY ANALYSIS

Temporal scale ($\epsilon$) vs. multivariate entropy for different emotions and groups:

- Angry
- Disgust
- Fear
- Happy
- Surprise
- Sadness
**Figure:** significant group difference for disgust and sadness $p \leq 0.05$
• Introduction
  • Problem statement, motivation and goals

• Database
  • Data collection
  • Preprocessing

• Data Analysis
  • Global analysis
  • Region-based analysis

• Conclusion
• Divide markers in 8 facial regions:
  Left and Right Eye brow, Eye, Cheek and Mouth

Figure: Partition of markers into eight facial regions
• Divide markers in 8 facial regions:
  Left and Right Eye brow, Eye, Cheek and Mouth
• Compute psychology-inspired measures
  • mimicry quality
  • left-right (motion) activation symmetry
  • upper-lower (motion) activation divergence

Figure: Partition of markers into eight facial regions
• Stimuli:
  • Experts score each region (0 - 5) depending on how activated is a region during an expression
  • Human annotated motion activation vector: \( \mathbf{v} = [v_1, v_2, ..., v_8] \)
• Stimuli:
  • Experts score each region (0 - 5) depending on how activated is a region during an expression
  • Human annotated motion activation vector: \( \mathbf{v} = [v_1, v_2, ..., v_8] \)

• Subjects:
  • Motion activation measured from participant’s mocap data
  • Motion activation_{reg} = total distance traveled by all markers from their respective rest positions
  • Let \( \mathbf{a} = [a_1, a_2, ..., a_8] \) be the activation vector
Table: Results of statistical tests, \( N = 45 \)

<table>
<thead>
<tr>
<th>Mimicry quality</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Corr between computed activation and manual ratings</td>
<td>Lower correlations for ASD, ( p = 0.024 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Left-Right activation symmetry</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Corr between left and right regions</td>
<td>Lower correlations for ASD, ( p = 0.0554 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Upper-Lower activation divergence</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference in activation between upper and lower face</td>
<td>Lower divergence for ASD, ( p = 5.23e-4 )</td>
</tr>
</tbody>
</table>
• Autoregressive (AR) model of order $p$ to model the dynamics of each region:

$$d_t = \sum_{i=1}^{p} \alpha_i d_{t-i} + \sigma_t,$$

(1)

where $\sigma_t$ is white noise, $\{\alpha_i\}_{i=1}^{p}$ are the model parameters.

• model order = 4

Figure: AR model order determination
Dynamics of each facial region of a subject is represented by a 4-dimensional vector.

Compute distances between all ASD and TD subject pairs for each emotion group.

**Observations:** Eye regions are more different than mouth region, disgust showing highest difference.
CONCLUSION

• Contribution
  • Quantification of facial-expression related atypicality in autism
  • Identifying expressions which induce more awkwardness
  • Identifying the facial regions inducing perception of atypicality
• Contribution
  • Quantification of facial-expression related atypicality in autism
  • Identifying expressions which induce more awkwardness
  • Identifying the facial regions inducing perception of atypicality

• Observations
  • ASD group has reduced dynamical complexity supporting the well-known complexity loss theory
  • Differences in dynamics is higher in eye region
  • ASD subjects underperform in mimicry
  • Differences are pronounced for expressions of disgust and sadness (negative valence)
Thank you