

Towards Socially Intelligent HRI Systems: Quantifying Emotional, Social, and Relational Context in Real-World Human Interactions

Jesse Parent

Undergraduate Student

Yelin Kim

PI

INSPIRE (Interaction Sensing and Perception in Real Environment) Lab

College of Engineering & Applied Sciences

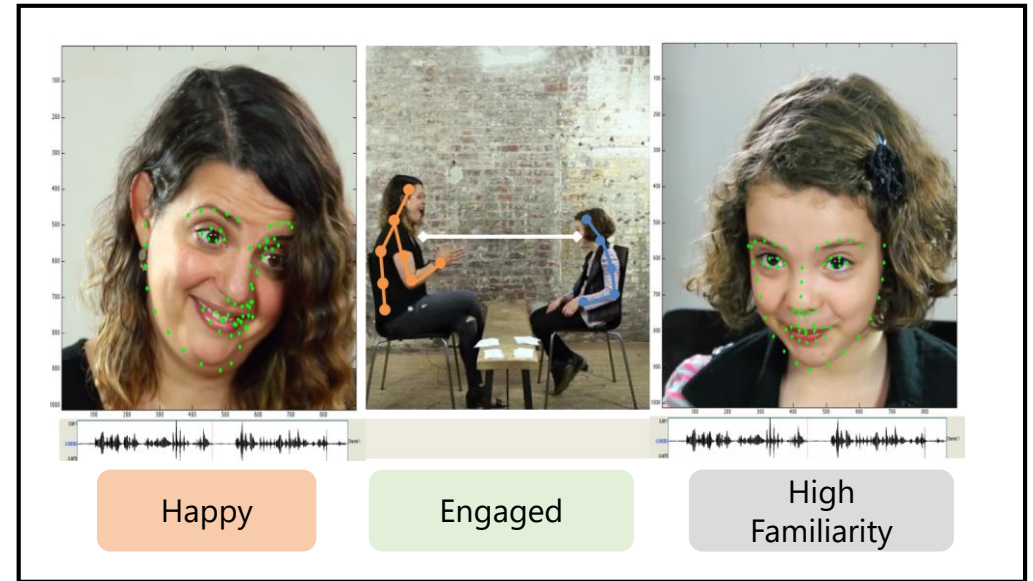
University at Albany, SUNY

Scope of Research

- Vision:
Socially Intelligent AI
- Mission:
Create new Data-set with real-world relational context
- Specific Target:
Quantify & Label Social Context

Vision

- Vision:
Socially Intelligent AI



- Mission:
Create new Data-set with real-world relational context
- Specific Target:
Quantify & Label Social Context

Mission

- Vision:
Socially Intelligent AI
- Mission:
Create new Data-set with real-world relational context
- Specific Target:
Quantify & Label Social Context

Target

- Vision:
Socially Intelligent AI
- Mission:
Create new Data-set with real-world relational context
- Specific Target:
Quantify & Label Social Context

Development of HRI Systems with Automatic Social and Relational Context Discovery

Advancing Socially-aware
Emotion Recognition Systems

New Dataset Creation with
Real-world Relational Context

Research Focus of This Paper:

Identifying Emotional, Social, and Relational Labels

Novelty: Social Context Label

- **Proposed Label: Engagement**
- *Regions of High Engagement*
 - *Isolating “emotionally high stakes” regions of interaction*

Novelty: High-Quality Film Footage

- Unscripted emotion display
- Real-life relationships
- Organic interaction



{THE AND}

Documentary Films

Skin Deep Media Corp

Q&A prompts discussion, **real dialogue** containing **organic interaction** and **affective display**

- Real-life interpersonal relationships
- High quality documentary style footage
(4080x1080p total resolution)
- Authentic: no actors or staged emotion



Transcript: Mother & Daughter		Activation	Valence
M1: What scares you the most?		4	3
D1: Oh, lots of things scare me, the first is you being murdered!		5	1
M2: Oh, that's the first?		5	1
D2: Yes!and the Boogie Man...		5	3

Findings

High Engagement Regions	Duration
Average length of a HE Region	44.3 seconds
Total number of HE Regions	10 events
Total time of all HE Regions	00: 07:23 minutes (18.6%)
Total time of conversation	00: 39:40
Utterance-Specific Ratings of Mother and Daughter	Mean Activation & Valence, (Standard Deviation)
M: All Utterances	A 3.24 (0.74), V 2.73 (0.80)
M: HE Regions	A 4.23 (0.59), V 2.54 (1.05)
D: All Utterances	A 3.49 (0.89), V 3.08 (0.84)
D: HE Regions	A 4.31 (0.48), V 3.00 (0.91)

Findings

High Engagement Regions	Duration
Average length of a HE Region	44.3 seconds
Total number of HE Regions	10 events
Total time of all HE Regions	00:07:23 minutes (18.6%)
Total time of conversation	00:39:40
Utterance-Specific Ratings of Mother and Daughter	Mean Activation & Valence, (Standard Deviation)
M: All Utterances	A 3.24 (0.74), V 2.73 (0.80)
M: HE Regions	A 4.23 (0.59), V 2.54 (1.05)
D: All Utterances	A 3.49 (0.89), V 3.08 (0.84)
D: HE Regions	A 4.31 (0.48), V 3.00 (0.91)

Findings

High Engagement Regions	Duration
Average length of a HE Region	44.3 seconds
Total number of HE Regions	10 events
Total time of all HE Regions	00: 07:23 minutes (18.6%)
Total time of conversation	00: 39:40
Utterance-Specific Ratings of Mother and Daughter	Mean Activation & Valence, (Standard Deviation)
M: All Utterances	A 3.24 (0.74), V 2.73 (0.80)
M: HE Regions	A 4.23 (0.59), V 2.54 (1.05)
D: All Utterances	A 3.49 (0.89), V 3.08 (0.84)
D: HE Regions	A 4.31 (0.48), V 3.00 (0.91)

Future Work & Let's Talk!

- Towards Socially Intelligent AI Systems
 - Quantify & label interaction
 - Measuring trajectory of relationship / closeness?
 - Affective cues across relationship types
- HRI applications
 - Interpersonal skills for Robots
 - Socially aware systems

Introduction

Socially Intelligent HRI Systems

AI systems that are able to comprehend social dynamics between multiple human speakers

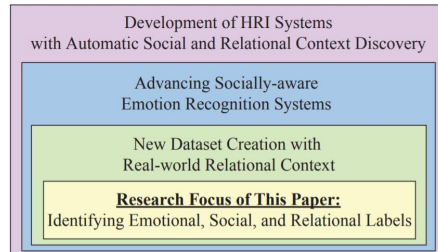
This poster presents:

Analyses of interaction via a **social context label**, **engagement**, with source data of real-life relationships and affective emotion display. Foundation for future work in quantifying complex human social interaction.

Building a Foundation

Goal: intelligent systems capable of understanding **social and relational context**

- Socially intelligent systems will need training datasets authentic human interactions
- Here, we work on quantification and labeling approaches that can capture emotional, social, and relational context



Novelty of Our Method

- {The AND} Dataset, given relationship status
- Building on established literature to quantify affective expression
- Combining given **relationship status** and **affective cues** to interpret **social contexts**
- New social context label: engagement


Towards Socially Intelligent HRI Systems

Quantifying Emotional, Social, and Relational Context in Real-World Human Interactions

Jesse Parent and Yelin Kim

University at Albany, College of Engineering and Applied Science, Albany NY, USA
jtparent@albany.edu, yelinkim@albany.edu

Authentic Emotion Within Real-Life Relationships



{The AND} Documentary Films

Q&A prompts discussion, **real dialogue** containing **organic interaction and affective display**

- Real-life interpersonal relationships
- High quality documentary style footage
- Authentic: no actors or staged emotion

Transcript: Mother & Daughter

	Act.	Val.
M1: What scares you the most?	4	3
D1: Oh, lots of things scare me, the first is you being murdered!	5	1
M2: Oh, that's the first?	5	1
D2: Yes!and the Boogie Man...	5	3

Quantifying and Labeling Social Context

Social context labels: quantifiable labels identifying or describing components of human interaction

Proposed Label: Engagement

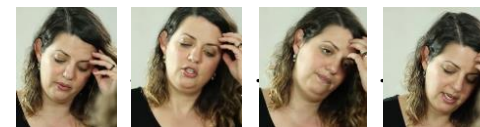
Regions of High Engagement

High Engagement (HE): demarcating 'emotionally high-stake' regions, may lend insight into underlying relational qualities

High Engagement Defined

- At least **4 utterances**, two turn changes between speakers, lasting at least 20 seconds
- Activation** ≥ 4 from both speakers: consistent expression of interest or significance
- Valence deviates ± 2** within the segment: fluctuation as indicator of non-standard conversation flow

Context Clues: Repetitive Postures



15:08 - 22:01 (M) Time marker - 2:05 - 27:59 (Q2 5)

What effective or nonverbal cues could indicate **social context**? When asking questions, M may show:

- Head shielding, pausing, delayed eye contact
- Distance or uneasiness, negative anticipation

M shows this **7 of 14** questions asked, such as:

- (Q17), "What do you love and hate about having a younger sister?"
- (Q21) "Do we spend enough time together?"
- (Q25) "How can I be a better mom?"

Results

Statistics of high engagement (HE) regions

High Engagement Regions	Duration
Average length of a HE Region	44.3 seconds
Total number of HE Regions	10
Total time of all HE Regions	00:07:23 minutes (18.6%)
Total time of conversation	00:39:40

Utterance-Specific Ratings of Mother and Daughter	Mean Activation & Valence, (Standard Deviation)
M: All Utterances	A 3.24 (0.74), V 2.73 (0.80)
M: HE Regions	A 4.23 (0.59), V 2.54 (1.05)
D: All Utterances	A 3.49 (0.89), V 3.08 (0.84)
D: HE Regions	A 4.31 (0.48), V 3.00 (0.91)

Discussion

Asymmetrical High Engagement Behavior

- Lower valence during HE Regions**
M displayed stronger emotional negativity within HE, D had less variation from normal to HE
- Repeated postures as indicators**
M displayed shields head during questions, lower valence, more concern

Longest HE Regions: Significant Relationships

- Top 2 HE events (70s, 80s) contained discussion of intimate relationships (dad, sister, family)

Different Approaches to Interaction

- M appears more wary, potentially more responsibility towards responses
- Implications for relationship status; imbalance

Future Work

- Significance of **interaction over time** for social context labels. Does time spent in certain social contexts imply trajectory of relational development?
- Variety of factors** contribute to social states – how can these be **quantified**, supported by literature?
- What topical or lexical conversations constitute HE periods in **across different relationship types**?