인간-로봇 사회적 상호작용을 위한 핵심 기술

2018-01-24
JongSuk Choi
What is sHRI?
What is sHRI?

- Human-Robot Interaction in **the socially acceptable way** such as
  - Eye contact
  - Turn-taking by intention reading
  - Joint attention
  - Perspective-taking
  - Social navigation behaviors
  - Cooperative planning
  - Proactive behaviors to learn task semantics from demonstration
  - Emotional empathy or sympathy
  - ...

- **Social (Robot) Intelligence** is the key technology for sHRI
Social Robot Intelligence
Motivation

Strong Assumption!

- Human always approach the robot with their attention
- They are all newcomers to the place
- Human don’t stop paying attention to the robot until the robot stops its expression.
- ...

What If?

- If someone is just passing by w/o attention?
- If someone is already familiar with the place?
  - If someone feels bored and then wants to leave?
  - If the person is not a customer but a clerk?
  - If we change the robot?
  - ...

Social Intelligence!
Proposed System (from a sHRI-related project DeepTask)

“Development of Social Robot Intelligence for sHRI of Service Robots”

Social Robot Layer

Ontology for social interaction

Task Creator

Service Package

Context Model | Task Model | Domain Knowledge

Task Manager

Robot Intelligence Layer

Knowledge Manager

Robot behavior interface

Robot Behavior Layer

Recognition module

Social behavior of human

Action module

Social expression of robot

Approaches

1. **Social behavior recognition**
   (Electronics and Telecommunication Research Institute, Korea Institute of Science and Technology)
   - Deep learning based social behavior perception
   - Human-Human and Human-robot interaction data acquisition
   - Long term human behavior tracking technology
   - Social dialogue act model for Korean language

2. **Social expression and behavior of robot**
   (University of Auckland)
   - Robot personality model based on human behavior
   - Emotion and speech expression model
   - Robot speech and movement coordination

3. **Social task knowledge**
   (Hanyang University and University of Seoul)
   - BDI (Belief-Desire-Intention) based user intention prediction knowledge
   - Social interaction ontology
   - External web data based knowledge expansion

4. **User evaluation**
   (Ewha Womans University)
   - Social interaction model evaluation
   - Social cue based interaction model
[Perception]
WHO (사람의 얼굴 인식)

- Deep Learning 기반 신원 인식

- Deep CNN 기반 얼굴특징추출
- PCA 기반 차원축소
- High-dim. LBP
- SVM 학습

Collaborative research: KIST & POSTECH, 2017
WHERE (사람의 위치 추적)

• Modified faster R-CNN
  - 기존의 ROI를 Modified faster R-CNN (zf)에 활용 → 비휴먼(로봇, 문, 책상, 의자 등)검출제거
  - 30명 인원에 대한 사용자 출입, 검출실패, 가려짐 등에 강인한 다중 객체 추적
WHAT (사람의 행위 인식)

- 제스처 및 행동 인식 프레임크 개선
  - 딥러닝 (LSTM) 기반 행동 길이 변화 대응 가능 행동 인식 수행 → 연속적인 행동 인식 프레임워크 개발
    - 행동을 "start-middle-end" 단계로 구분지어서 인식
    - 3D CNN 구조를 이용하여 각 단계를 학습
    - 입력 정보가 불충분한 상태에서의 강화학습 행위 인식을 위해 현재의 판단을 기반으로 앞으로의 입력을 예측 수행
  - LSTM (Long Short Term Memory) 구조를 이용하여 "end" 단계 생성을 통해 분류 성능 개선

Collaborative research: KIST & KU, 2017
Results: 3W [WHERE + WHO + WHAT (A/V)]

Collaborative research:
KIST, POSTECH & KU, 2017
[Cognition]
Knowledge (Ontology)
Memory

Short-Term Memory

Long-Term Memory

Instant Memory
- Raw perception data
- 안식정보, 음성정보, Task 정보등 매순간마다 발생하는 모든 인식 데이터

Episodic Memory
- 과거 이벤트에 대한 기록 정보
- 이벤트의 시작/끝에 따라 모집된 Instant Memory들의 집합

Semantic Memory
- 일반적인 사실 (지식의 일부)
- Episodic Memory로부터 Encoding / Abstract된 지식 정보
Knowledge Base Management

Memory Generation Process

Knowledge Base Management

- Knowledge Manager
- Knowledge Generator:
  - Memory Generation Model
- DB Interface

Reasoning Answer

Ontology (Semantic Knowledge)

- Person
- Class

DataBase (Short/Long-Term Memory)

- LTM
  - Semantic Memory
- LTM
  - Episodic Memory
- STM
  - Instant Memory

Perception Data, Task Data, External Data.

Reasoning

Encoding

Abstracting

Instant Memory
Episodic Memory
Semantic Memory
Knowledge Base Management

Knowledge(Personality) Generation in Memory System

**Thin Slice of Behaviors**

*(Non-Verbal Features Set)*

- **Thin Slice based Personality**

- **Long-Term Personality**

---

**로봇 연구단**

Robotics Research Group
[Expression]
Robot’s characteristics

- 5 Factor model + GMM

![Diagram showing the relationship between Robot's characteristics, Internal State, Generative Probability Model, Bayes' Rule, and Emotion.]

By courtesy of Prof. Ho Seok Ahn, UoA
Robot’s behavior generation

- Platform independent, multi-modal expression

By courtesy of Prof. Ho Seok Ahn, UoA
Speech-Gesture Behavior

대화 상황에 적절한 로봇 제스처 생성 기술 개발

- 로봇 제스처 생성 모델 설계 및 구현
  - Recurrent Neural Network 기반 제스처 분류 모델
  - 언어 해석을 위해 Twitter로부터 학습된 Word Embedding 사용
  - 수집된 TED 영상 및 자막으로 Network 학습
  - NAO 로봇에서 발화 문장에 대한 제스처 생성 시스템 개발
  - 31명을 대상으로 한 설문 평가 결과, 학습된 모델을 가장 선호 함 (학습된 모델과 랜덤 선택을 포함한 3개 모델 간 비교)

---

< 제스처 선택 모델 >

<table>
<thead>
<tr>
<th>LSTM</th>
<th>LSTM</th>
<th>LSTM</th>
<th>LSTM</th>
<th>LSTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Embedding</td>
<td>Word Embedding</td>
<td>Word Embedding</td>
<td>Word Embedding</td>
<td>Word Embedding</td>
</tr>
</tbody>
</table>

This is what I meant

< 사용자 대상 평가 결과 (선호도) >

- 제안방법 (RNN기반) 53.3%
- 랜덤방법 #1 30%
- 랜덤방법 #2 16.7%

By courtesy of Dr. Minsoo Chang, ETRI
Speech-Gesture Behavior

So obviously when we think about nonverbal behavior.
[Simple Integration]
Social HRI: *Encounter*
Project plan and applications (from *DeepTask*)

**Project Plan**

<table>
<thead>
<tr>
<th>1st Year</th>
<th>2nd Year</th>
<th>3rd Year</th>
<th>4th &amp; 5th Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greet and start to form close rapport</td>
<td>Attract person’s attention</td>
<td>Acquire information and deliver necessary information</td>
<td>Empathize and encourage human</td>
</tr>
</tbody>
</table>

**Social interaction applications**

**Healthcare service**

- Deliver basic medical information to elders
- Understand the social relation with users especially with elder people

**Reception service**

- Reception service for department store, university library, etc
Concl usion

- **sHRI**: HRI in the way of social acceptance
  - Eye contact
  - Turn-taking by intention reading
  - Joint attention
  - Perspective-taking
  - Social navigation behaviors
  - Cooperative planning
  - Proactive behaviors to learn task semantics from demonstration
  - Emotional empathy or sympathy
  - …

- **sRI (Social Robot Intelligence)** is the key technology for sHRI
  - Perception
  - Cognition
  - Expression

- **sDS (Social Domain Services)** need to be set-up to show the necessity of sRI
  - Healthcare service
  - Reception service
  - …

- **Multi-disciplinary Research**
  - AI, Robot, Psychology, Human-Factor, UI/UX, …
Thank You