The build of new short-term social action DB using RGB-D cameras

HoSub   Yoon

ETRI

22 November 2017

Contents

- **Introduction**

- **Previous social action DB**

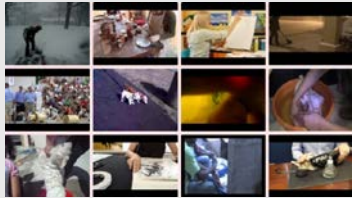- **ETRI social action DB**

- **Conclusion**

Introduction

- **A lot of action DBs**

- **But, rare databases of the focused on social behavior recognition**

- **we aim to construct a new action databased focused on "social behaviors"**

- **What is "social behaviors"**

  -> refers to human behavior that occurs in a social context or situation: Psychology glossary

  ->an Action is 'social' if the acting individual takes account of the behavior of others and is thereby oriented in its course: Max Weber

- **How to learn "social behaviors"**

  -> from observer's view: drama, movie, theater, sports,,, by broadcasting

  -> from interactor's view: simulations by DB making situation(head-mounted cameras)

Analysis of Previous Action DB 1
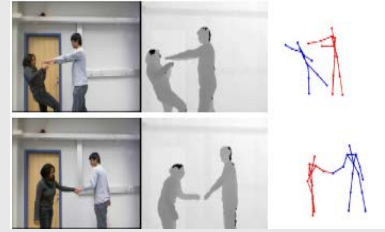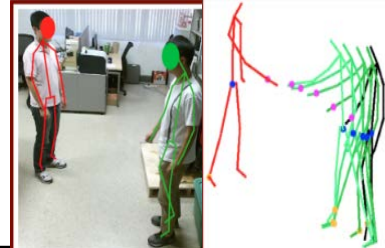
Action	DB Name	Depth	Affiliation	Contents	Example
Daily life action, Non-HHI	CAD-60/CAD-120	O	Cornell University	▪ Daily life action video on bedroom, ketch room, living room for robot learning ▪ 12 activities(60): rinsing mouth, brushing teeth, wearing contact lens, talking on the phone…… ▪ 10 high-level activities(120): making cereal, taking medicine, stacking objects, unstacking objects…… ▪ 60/120 video clips including skeleton images	
	IM-DailyDepthActivity	O	POSTECH	▪ Daily life action still images ▪ 15 person x 15 class actions: both hands waving, bending, boxing, clapping, cleaning, eating, exercise, kicking, phone conversation, reading an article, right hand waving, sit down, stand up, throwing and take an object. ▪ Silhouette image and skeleton joint position	
	MSRDailyActivity3D	O	Microsoft Research	▪ Daily life action video ▪ Drink, eat, read book, use laptop, paly game, call cellphone, write on a paper, use vacuum cleaner(8 class) ▪ RGB-D images by Kinect, 320 video clips	

Analysis of Previous Action DB 2

Action	DB Name	Depth	Affiliation	Contents	Example
Mixed	ActivityNet Dataset	X	Universidad del Norte, KAUST	▪ Human-Object interaction, Human-Human Interaction ▪ 200 classes: daily life, sports, social activity...... ▪ RGB video clips, 20,000 clips, 648hours	
	UCF-101	X	Univ. Central Florida	▪ Human-Object interaction, Body-motion Only, Human-Human Interaction, sports...... ▪ Cutting In kitchen, hammering, hula hoop, pizza tossing, shaving bread, skate boarding...... ▪ 101 human action class, 13320 clips, mean clip length 7.21 sec, 320x240 resolution, collected from Utube including audio(51 action)	
	THUMOS 2015	X	Univ. Central Florida	▪ Challenge data sets including UCF-101 DB ▪ Training Data: UCF-101, Background Data: New data sets ▪ Validation Data: including same time multiple action ▪ 24,000 clips , 45M frame video	
	Human Motion DataBase(H MDB-51)	X	Brown University	▪ Sports, every day behaviors, HHI ▪ Draw sword, dribble, drink, eat, kiss, laugh, pick, pour ▪ 51 classes, each class have minimum 101 clips, total 6766 clips	

Analysis of Previous Action DB 3

Action	DB Name	Depth	Affiliation	Contents	Example
Human - Human Interaction	UT-Interaction	X	Univ. Texas at Austin	▪ Human-Human interaction ▪ 6 class: Hand shaking, Hugging, Kicking, Pointing, Pushing ▪ 20 video clips with 1 minute ▪ Including multi-people interaction(High level situation)	
	CMU Motion Capture	O	Carnegie Mellon University	▪ Human interaction, Interaction with environment, locomotion, physical activity & sports, situation and scenarios ▪ .tvd, .c3d, .amc, .mpg, animated format ▪ Frame rate 120, total 2605 clips	
	SBU Kinect Interaction	O	Stony Brook University	▪ one person is acting and the other person is reacting. ▪ 8 classes: approaching, departing, pushing, kicking, punching, exchanging objects, hugging, and shaking hands. ▪ 21sets, total 300clips, depth and motion capture	
	UCLA HHOI	O	UCLA	▪ Learning Social Affordance for Human-Robot Interaction ▪ 5 classes: shake hands, pull up, high-Five throw and catch, hand over a cup ▪ skeleton, rgb-D, Depth2RGB mapper, annotation, each class have 24 clips	

Analysis of Previous Action DB 4

Action	DB Name	Depth	Affiliation	Contents	Example
Human /Human Interaction	Physical sHRI	O	Osaka University	▪ robot view point Human robot interaction ▪ classes: jumping jack, jumping on both legs, jumping on right or left leg, running, walking, side jumps, skipping on left or right leg ▪ 14 days real environment acquisition 13,938 clips	
	Paired Egocentric Video	X	University of Tokyo	▪ dynamics of social interactions between two people using a head-mounted camera ▪ classes: pointing, attention, positive, negative, passing, receiving, gestures ▪ 1,226 clips pair	
	The MHI-Mimicry database	X	Imperial College London	▪ phenomena occurring in social interactions ▪ Annotation: dialogue acts, turn-taking, affect, head gestures, hand gestures, body movements and facial expressions ▪ In total there are 54 recordings, of which 34 are of the discussions and 20 of the role-playing game. 40 person(male: 28, female: 12), 11h 40min	

ETRI social action DB

- **Short-term social action DB**

  -> **short-term means temporal behavior within 10 seconds that can easily identify intent.**

  -> **long-term means that can be gained through long-term observations such as habits and attitudes.**

- **DB configuration**

  -> **RGB vs. RGB-D**

  -> **Observer's view vs. Interactor's view**

  -> **Daily life, sports vs. Social action only**

  -> **HOI(Human-Object interaction), HHI, HMI vs sHRI**

ETRI social action DB

- **6 categories by grouping**

  -> we proposed the 6 categories of social action

- **1: Greeting(4)**

  -> Bow head, Shake hands, Give out the handshake, Snuggle
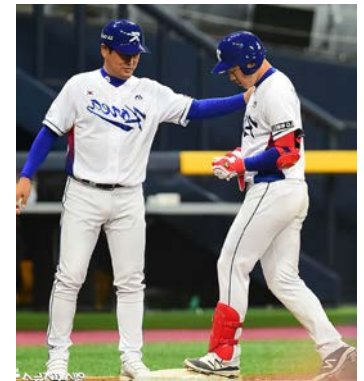
ETRI social action DB

- **2: Agreement(3)**

  -> Nod, Rolling head, Tilting head

- **3: Encouragement (5)**

  -> Clench fist, Stick out thumb, Make O.K finger sign, Make heart using finger, Knock one's shoulders

ETRI social action DB

- **4: Considering (4)**

  -> Fold arms, stretch head, chin in hand
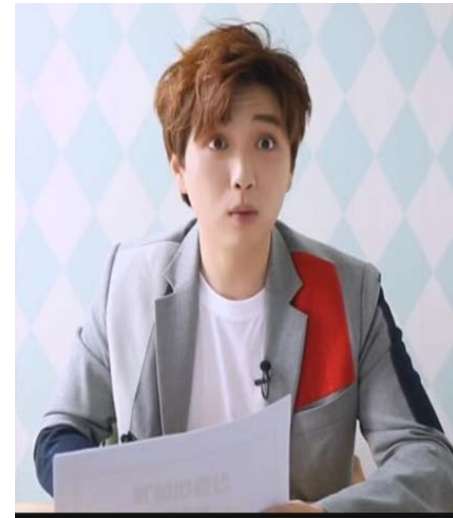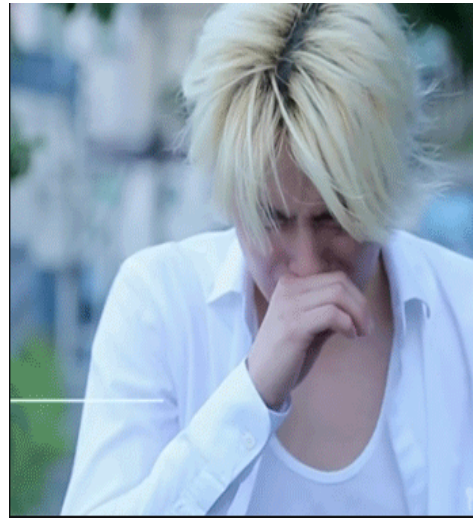

- **5: Eye contract (3)**

  -> view the philtrum, view the pupil, view the air
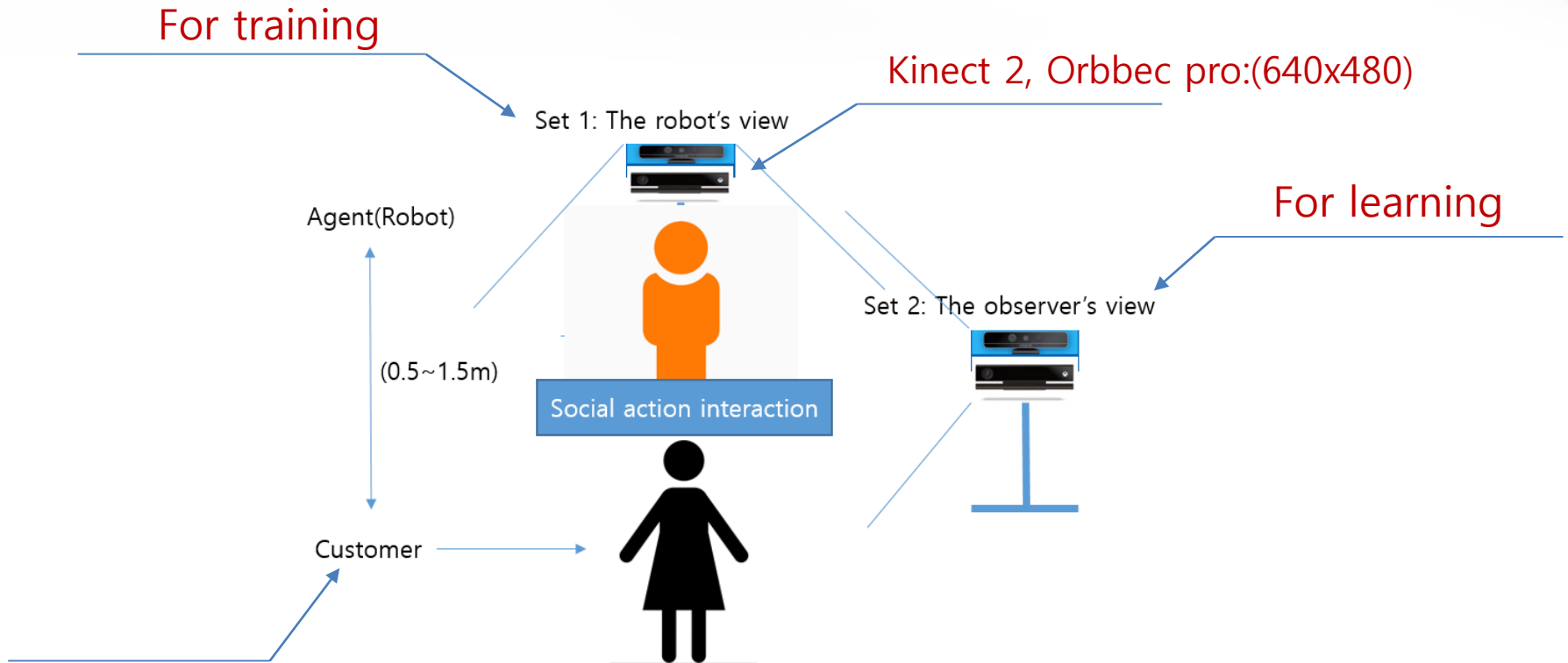
ETRI social action DB

- **6: Facial emotion expression(4)**

  -> Happy, Angry, Sad, Surprise (disgust, fear)

ETRI social action DB

- **Making environment of social action DB**

For training

Kinect 2, Orbbec pro:(640x480)

Set 1: The robot's view

For learning

Agent(Robot)

Set 2: The observer's view

(0.5~1.5m)

Social action interaction

Customer

- elementary school students to 60s or older
- total of 140 subjects, 10 men and 10 women at 10-year intervals
- repeats action 10 times for 24 social actions of the above 6 groups
- Including profile information as sex, age, hair style, lipstick color, clothes style, clothes color, accessories

ETRI social action DB

- **Examples**

Conclusion

- aim to construct a new action databased that is specified to social behaviors

- define various short-term social behaviors in to 6 groups according to their meaning

- describe the surrounding environment for acquiring them

- play an important role in recognizing and understanding the social interaction between human and robot

- need to make long-term social action DB in the future.