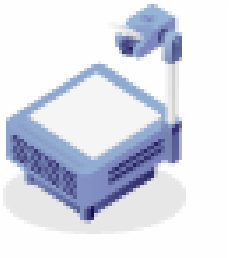
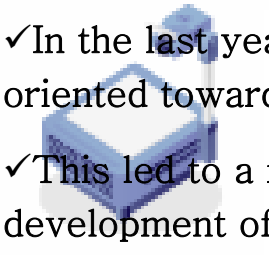
Propose the Interaction Model Architecture for HRI Components Standardization

2006. 10. 10

Contents



STATE-OF-THE-ART



✓In the last years, robotics research has focused more and more intensively on applications oriented towards service to humans, medical assistance, and human-friendliness in general.

✓This led to a number of impressive results in academic and industrial contexts, in the development of robots in diverse shapes, as humanoids, pets, medical tools, or appliances.

✓While the application range is wide, and so is the possible shaping of this kind of robots, a common element is that they are increasingly introduced into the Society of Humans.

✓Crucial aspects of robot design are therefore the modalities, mechanisms and tools of human-robot interaction, of communication with human beings, with the environment and possibly with other robots, and of interaction of the robotic systems with human environments, modelled on persons' needs and habits, and no more arranged according to the robot functions.

Robot as Team Member

✓Toward Human-Robot Collaboration, highlights the importance of creating robot capabilities and interfaces that address human concerns such as social appropriateness, safety, and quality of service.

✓Robots are, or soon will be, used in such critical domains as search and rescue, military battle, mine and bomb detection, scientific exploration, law enforcement, and hospital care.

✓Such robots must coordinate their behaviors with the requirements and expectations of human team members; they are more than mere tools but rather quasi-team members whose tasks have to be integrated with those of humans.

✓Robot as Team Member, highlights the importance of building core science and understanding the social and technical issues in human-robot interaction in the context of teams and groups

Special Issue on Human-Robot Interaction

- ✓ The recent trend toward developing a new generation of robots that can participate in our lives and exist in human environments has introduced the need for investigating the paradigms, techniques, and technologies for the interaction between people and robots.
- ✓ An important goal for the field of *Human-Robot Interaction* is to develop autonomous and semi-autonomous robots that operate within human spaces and play a beneficial role in the daily lives of ordinary people.
- ✓ Interaction between people and robots may potentially span physical, cognitive, task-based, social, or emotional dimensions.
- ✓ Human-robot interaction poses multi-faceted problems, requiring not only technical but also cultural, sociological, psychological, philosophical, and even ethical considerations.

Μιτσυβισηι, Ωακαμαρυ ροβοτ

Special Issue on Human-Robot Interaction

- ✓ How to model the interaction of a human being with a robot?
- ✓ How to manage the physical, intellectual, and emotional exchange between human beings and robots?
- ✓ How to realize an effective communication of robots with the humans and their environments?
- ✓ These questions and many others are the stimulus for this Special Issue of the human-Robot interaction, aimed at gathering the latest results by robotics researchers facing the diverse problems related to human-Robot interaction.

Μιτσυβισηι, Ωακαμαρυ ροβοτ

Getting to Know Socially Intelligent Robots

- ✓ Human-Robot Interaction and Communication is a quickly growing research area at the intersection of research fields such as robotics, engineering, psychology, ethology and cognitive science.
- ✓ Significant initiatives are currently underway funded by public, academic, governmental as well as industrial initiatives, exploring and aiming at advancing this research field and opening up novel and challenging applications.
- ✓ Robots moving out of laboratory and manufacturing environments face hard problems of perception, action and cognition.
- ✓ For robots to be accepted as assistants or companions in people's private homes and everyday environments technological solutions do not suffice: The `human in the loop', as the potential customer and user will decide on the ultimate success of a `home robot' as a product.
- ✓ Application areas that heavily involve human contact are a particularly challenging domain.

Μιτσυβισηι, Ωακαμαρυ ροβοτ

Getting to Know Socially Intelligent Robots

- ✓ Human societies have easily assimilated new technologies, such as mobile phones, but it is less clear in which application areas robots will be accepted.
- ✓ Robots as embodied beings, physical, possibly humanoid or android entities that share our living environments and accompany our lives will have a certain degree of autonomy, initiative, cognitive skills and will communicate and interact with people in ways inspired by human-human contact.
- ✓ Interaction and communication of embodied physical robots with humans is multi-modal, and involves deep issues of social intelligence, communication and interaction that have traditionally been studied primarily in psychology and other areas.
- ✓ The design of a robot's behaviour, appearance, and cognitive and social skills is highly challenging, and requires interdisciplinary collaborations across the traditional boundaries of established disciplines.

Μιτσυβισηι, Ωακαμαρυ ροβοτ

I. Review of HRI Technology

1. Definition

2. Characteristics

3. Classification

I-1. Definition

- **Human-Robot Interaction (HRI) is a core technology that can naturally interact between human and robots through robot camera, microphone, and various sensors for intelligent service robots.**

Μιτσυβισηι, Ωακαμαρυ ροβοτ

I-2. Characteristics

- **HRI technology is different from Human-Computer Interaction (HCI) in that robots have an autonomous movement, bidirectional feature of interaction, and diversity of control level.**

- **To development an effective HRI, the system with module architecture to implement convenience between human and robots, cooperation, and friendship should be needed.**

I-3. Classification

- **Vision-based HRI : face recognition, gesture recognition, behavior recognition, facial expression recognition**
- **Audio-based HRI : speech recognition, speaker recognition, sound localization, emotional recognition**
- **Others : PDA-based HRI interface, emotion generation and expression**

II. Introduction of HRI Technology (HRI Team Research Activities in ETRI)

1. Vision-based HRI technology

2. Audio-based HRI technology

3. others

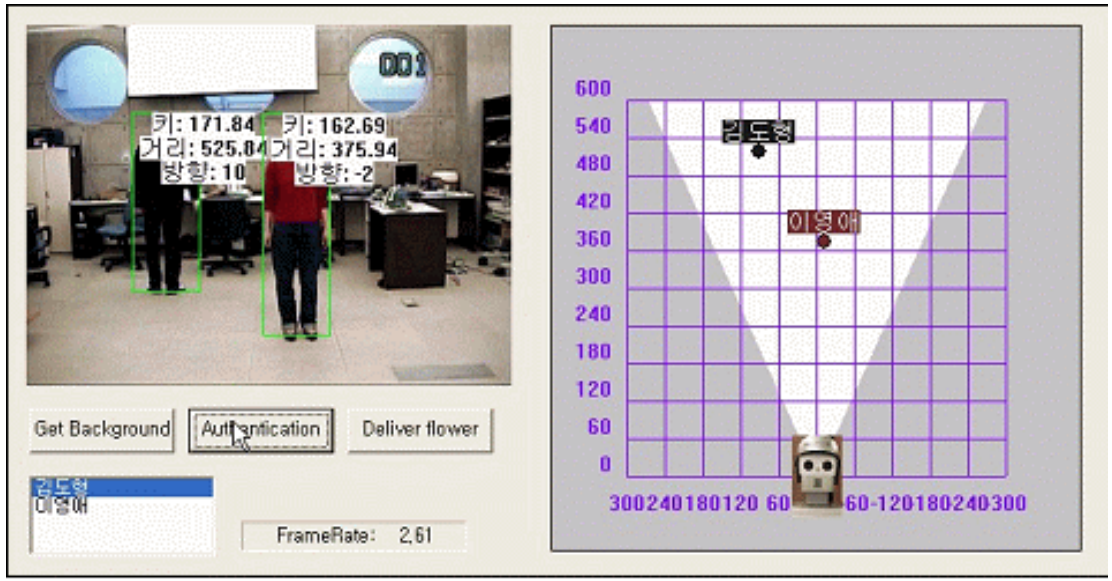
II-1. Vision-based HRI

- **Face detection/tracking/recognition:** the face detection/tracking allows robot to detect a human face from images obtained through robot camera and to track a movement of the human face during natural conversation with the robot. The face recognition/verification allows robot to recognize a member of family and to verify the identity of the human face.

NEX, Παπερο ροβοτ

II-1. Vision-based HRI

- **User identification using semi-biometrics**
  - The face recognition/verification is furthermore allows robot to know the direction, distance, and identity of the user using semi-biometrics information such as user's cloth color and height.



II-1. Vision-based HRI

- **Gesture recognition:** Meaningful gestures could be represented by both temporal hand movements and static hand postures. It can be efficiently used in noisy environment and can also give active services by recognizing user's intension for intelligent robots

Ηονδα, ΑΣΙΜΟ ροβοτ

II-1. Vision-based HRI

- **Facial expression recognition:** Robotic systems for natural user interaction performs facial expression recognition, since facial expressiveness is regarded as a key component to developing personal attachment. The six basic facial expressions are the expressions of the six basic emotions of humans: anger, disgust, fear, happiness, sadness, and surprise.

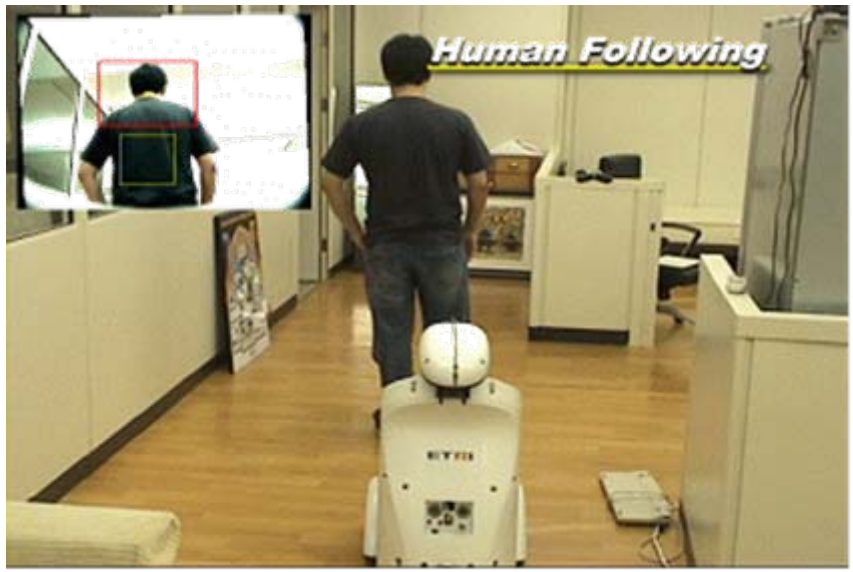
MIT, KISMET ροβοτ

II-1. Vision-based HRI

- **Caller identification**
  - The robot recognizes the face image of caller with hand gesture. The hand gesture detection is obtained by hand shape around face image of caller. After recognizing hand gesture, robot moves forward caller.

II-1. Vision-based HRI

- **Human following**
  - The robot follows specific person by using color histogram of cloth and information of depth map obtained from stereo camera of moving robot. Human following obviously requires real-time ability to respond to the changing position of the person to be followed.

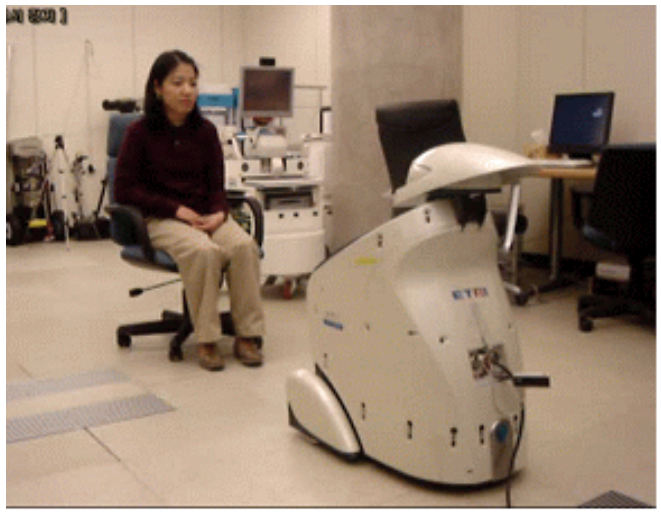


II-1. Vision-based HRI

- **Human detection and tracking**
- **Object recognition**
- **Human following**
- **Behavior recognition**
- **Posture recognition**

II-2. Audio-based HRI

- **Speaker recognition:** the text-independent speaker recognition allows robot to recognize the identity of a speaker during natural conversation with robot. Furthermore, this technique allows a speaker to communicate with robot through spontaneous speech.



ETPI, Ωεωερ ροβοτ

II-2. Audio-based HRI

- **Sound localization:** the main goal of this technique is to find the direction of the call voice by uttering its name or hand clapping sound. After that, robot moves toward the position where the sound is generated,

Τοσηιβα, Απρι σηαρπ εαρ ροβοτ

II-2. Audio-based HRI

- **Speaker and speech recognition**
  - This technique allows a speaker to communicate with robot through spontaneous speech. With text-independent speaker recognition, it is able to provide useful information such as daily life schedule and TV program suitable to the speaker.

II-2. Audio-based HRI

- **Emotion recognition**
- **Sound source separation**
- **Speech understanding**
- **Speech synthesis**

II-3. Others

- **PDA-based HRI interface**
  - PDAs can be used to interact with robots. Furthermore, they provide a suitable interaction device for teleoperation and a touch-screen interaction capability

II-3. others

- **Emotion generation and expression**
  - This is a technique that the robots generate and express their emotion corresponding to human's behavior

KAIST, Albert HYBO

KITECH, EveR-1

III. Propose the standardization for HRI Components

What do users want?

Requirements from users	*We may need .. for standards*
Simple and natural interfaces	*Voice or Vision(Gesture, face, etc.) interfaces*
Simple and natural commands (don't want remember voice or gesture commands)	*Common or natural words interfaces Natural gesture interfaces*
If they have to remember keyword, They want to remember same word of their language for the same service of different robot makers. (If you want to know the weather, you just use "weather")	*Keyword standards*
If they have to use keyword and If they are Korean, They want to use Korean keywords. (If they want to know the weather, they want to use "날씨" means weather)	*Keyword Internationalization Standard*

What do robot makers want?

Requirements from robot makers	*We may need .. for standards*
They want to know interface guide line For HRI functions (speech, vision, etc.)	→ *Standard interfaces*
They want to use existing functions about HRI technologies	→ *Standard components*
They cannot have enough machine power to use HRI functions and they want to use just HRI services from remote server.	→ *Remote HRI Service Component Standards*
They want to upgrade functions or algorithms easily with low maintenance costs.	→ *Algorithm Independent architecture*

What do App. Developers want?

Requirements from App.
Or function developers

We may need .. for standards

They want same interface independent on robot types and makers to develop applications.

→ *Standard Interfaces*

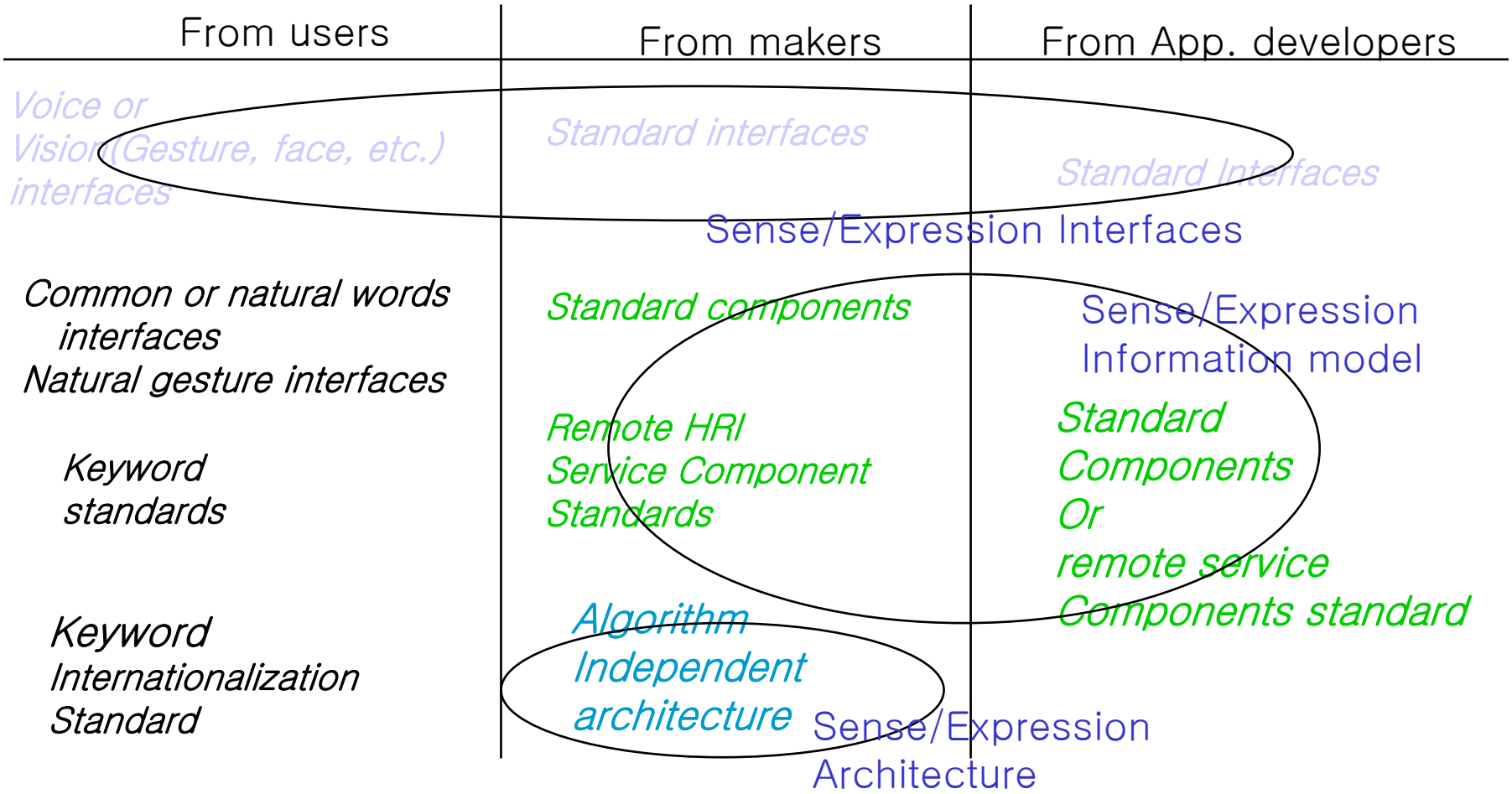
They do not want to develop new applications for the same services or functions again for each different robot

→ *Standard Components Or remote service Components standard*

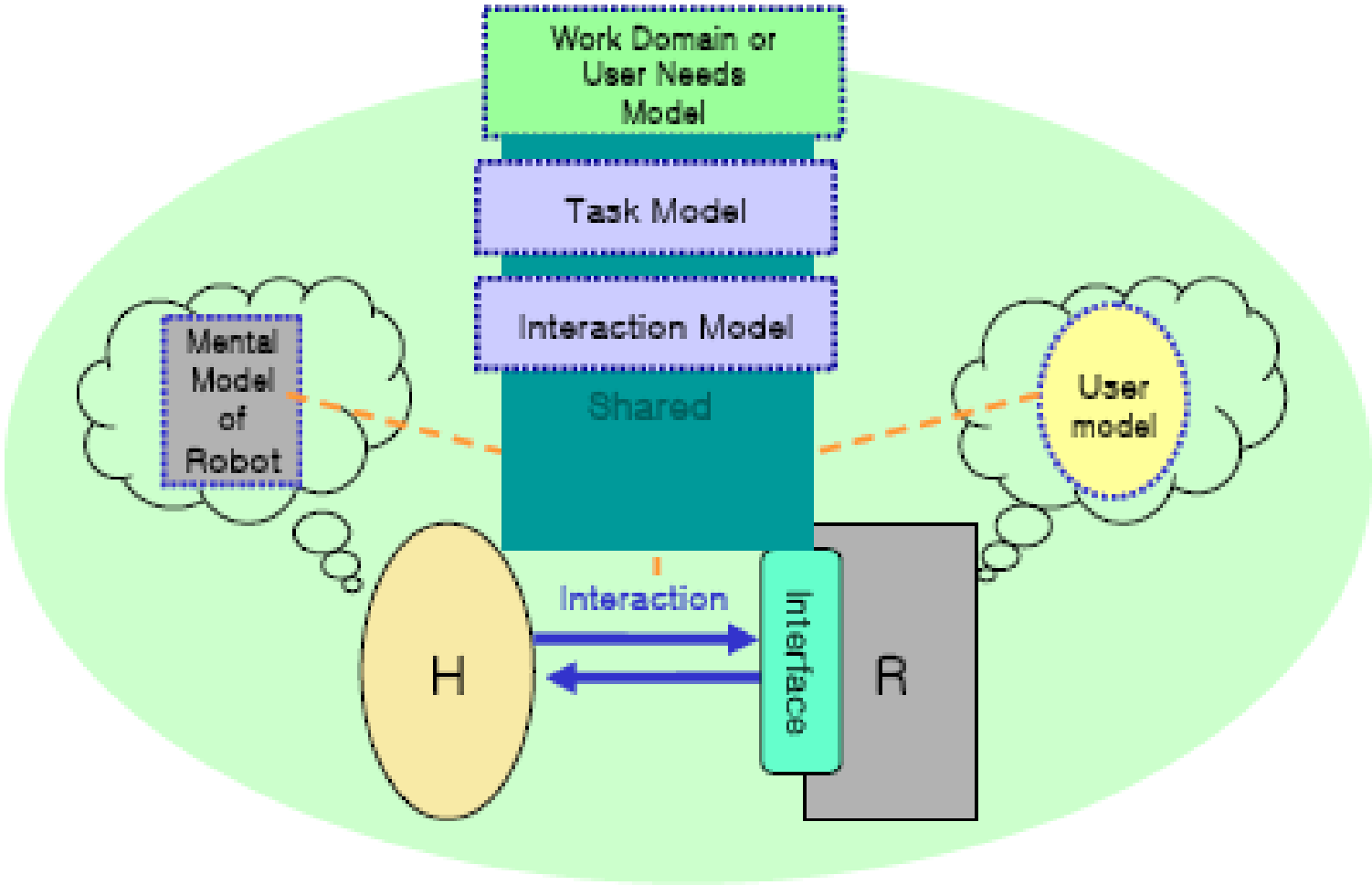
Standard Items

From users	From makers	From App. developers
Voice or Vision(Gesture, face, etc.) interfaces	*Standard interfaces*	*Standard Interfaces*
Common or natural words interfaces *Natural gesture interfaces*	*Standard components*	
Keyword standards	*Remote HRI Service Component Standards*	*Standard Components Or remote service Components standard*
Keyword Internationalization Standard	*Algorithm Independent architecture*	

Standard Items

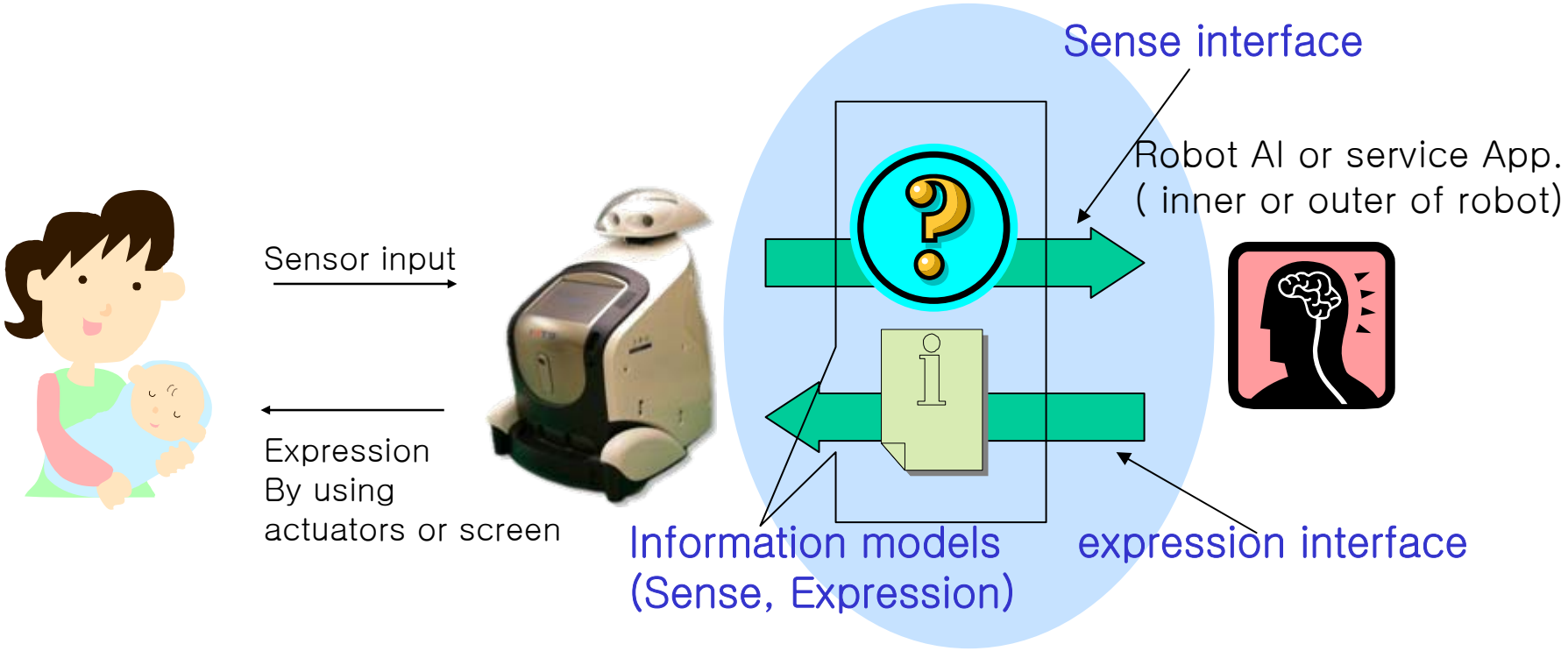


Model of HRI Component Function



Interaction Model Architecture for HRI

What do we suggest for the standard



Information Model

- Our proposed standard for HRI consists of one model and two interfaces
  - information model and recognition and expression interfaces
- Information model designs the types and structures of information objects as well as their management methods.
- The objects are two types
  - ✓ recognition and expression.
- First of all, information model consists of the types structures and protocols of information from sensors and applications of robot.

Sense Information Model

- Sense information model says When, Who say, What, Where from Human to Robot information
- Sense Information model consists of 4W Model
  - ✓ Who information : face recognition, speaker recognition, etc.
  - ✓ Where information : vision based location information, landmark based location info., etc.
  - ✓ What information : speech recognition, gesture recognition, etc.
  - ✓ When information: scheduler, clock, etc.

Sense Information Model Detailed

Continue…

- Who information model
  - ✓ Generally it can be obtained by face recognition , speaker recognition, etc.
  - ✓ We define whoinfo as path + userId information
    - Ex) “//korea/south/taejon/etri/mrsong” : “//korea/south/taejon/etri/” Is path, “mrsong” is id
  - ✓ If the robot do not know the correct user id, they send WhoHint(face image or voice sound) to server
    - Ex) WhoHint = { Face Image or Voice sound }
    - if server do not know location
      - WhoHint = { (Image or Sound), Location path}

Sense Information Model Detailed

Continue…

- Where information model
  - ✓ Generally it can be obtained by location finding algorithm
  - – We define whereinfo as address(or path) + position information
    - Ex) Can be direct type or near type
      - direct type: “//korea/south/taejon/etri/7thbuilding/L864/x=10;y=13”
      - near type : →have to translate direct position.
      - “//korea/south/taejon/etri/robotdivion/hriteam/song”

Sense Information Model Detailed

Continue…

- What information model
  - Generally it can be obtained by speech recognition, gesture recognition, any other command recognition methods.
  - We define whatinfo as natural string or path command
    - Ex)
      - Natural string
        - “change the TV channel”, or “show today schedule”
      - Command format with path information
        - TV (“//korea/south/taejon/etri/7thbuilding/L864/38TV”). channel.changeUp();
  - If the robot do not understand a correct command, the robot send WhatHint(gesture image, voice or any other command data) to server
    - Ex) WhatHint = { Image, voice or other command data(IR remote controller info, etc.) }

Sense Information Model Detailed

Continue…

- When information model
  - Generally it can be obtained by clock and scheduler, etc.
  - WhenInfo show current time
  - We define wheninfo as time information
    - Ex)
      - Date Time : from clock
        - “2006-09-13-14:33:23”
      - If command use when info(in case users say “response until tommorow”), it can be processed robot AI with Whatinfo and Wheninfo.
  - If the robot do not send wheninfo, servers use their internal clock.

Sense Interfaces

- Sense Interfaces use sense information model as their arguments and results.
- include interfaces for 4W info model and Hint model
- include sensor type interfaces and processor type interfaces
  - Sensor type: active, event style
  - Processor type: passive, it can be used in a application.

Sense Interfaces Detailed

- Sensor(Event) type Interface:
  - On… interfaces
    - OnWho(WhoInfo), OnWho(WhoHint)
    - OnWhere(WhereInfo), OnWhere(WhereHint)
    - OnWhat(WhatInfo), OnWhat(WhatHint)
    - OnWhen(WhenInfo)
    - OnSee(ViewImage), OnSound(SoundData)
      - If robot cannot sense anything, In some case, robot send all images or sounds to be processed to server.

Sense Interfaces Detailed

- Processor(Query) type Interface :
  - Get… interfaces
    - GetWho(), GetWhoHint()
    - GetWhere(), GetWhereHint()
    - GetWhat(), GetWhatHint()
    - GetWhen()
    - GetSee(), GetSound()
  - Get… interfaces have block mode and non-block mode

Expression Information Model

- Expression Information model consists of
  - How information
    - Just say abstract expression commands: first step
    - It can be translated each methods for each robots: second step
    - Therefore, Expression information model require some special architecture

Expression model (abstract)
ExpressionHint model (detailed actions)
ExpressionRender architecture

Expression Information Model Detailed

- Expression Model
  - It can be produced from Robot AI or service application in server
  - It can be processed on robot side or server side.
    - In case of robot side, robots process expression words
    - In case of server side, servers process words by calling robot control interfaces.

Expression Information Model Detailed

- Expression Model
  - It have to be defined with standard word→ we have to define some robot expression word.
    - Ex) “say(<text>)”, “approach(<user>)”, etc.
    - It may match to same interface
      - “say(hello)” can match to robot.say(“hello”)
  - Some case, an expression model may not matched to an expression interface
    - Ex) expression can be “greeting” but a robot may not do “greeting” because they do not know the meaning of “greeting”
      - In this case, a robot use expression hint.

Expression Information Model Detailed

- Expression Hint Model
  - If a robot cannot process expression, they want expression hint.
  - ExpressionHint shows action flow instead of abstract words.

Expression Information Model Detailed

- Expression Hint Model
  - It can be differently implemented by each robot makers or robot service application developer .
    - ExpressionModel: Greeting
      - A maker:
        - Expression(Greeting) = ExpressionHint(TTS(“Hello”))
      - B maker
        - Expression(Greeting) = ExpressionHint(robot.arm.shakeHands());

Expression Interfaces

- Expression Interfaces use expression information model or expression hint model as their arguments and results.
  - Expression Interface…
    - ExpressionRenderer.Expression(“some expression”)
    - ExpressionRenderer.Expression(“expression”,ExpressionHint hint)

Expression Renderer

- Expression Renderer process Expression, ExpressionHint
- If Renderer cannot process Expression, it try to get ExpressionHint from robot makers.
- Renderer is a process engine or task engine.
- It can be placed on application server or robot side.

Conclusion

We need to many opinions

and join from others

Thank You