Multimodal Human-Robot Interaction

Vinay Krishna Sharma
PhD Student, I3D Lab
CPDM, IISc Bangalore

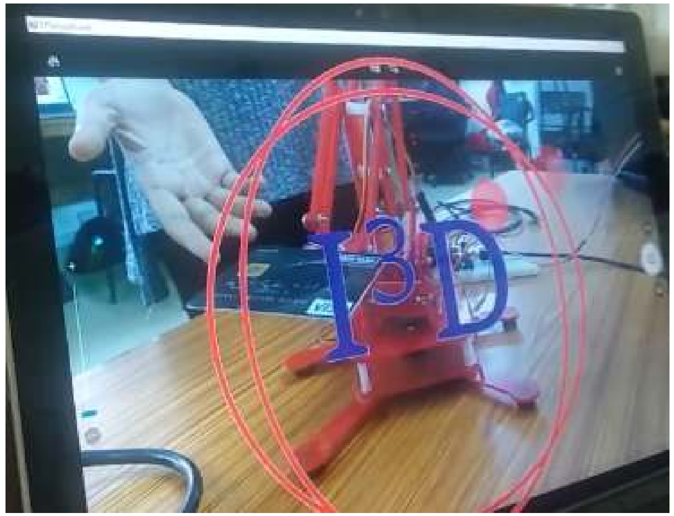
Supervised By
Dr Pradipta Biswas
Asst. Prof. IISc Bangalore

Contents

1. Introduction
2. Literature Survey of HRI
    - Human Factors
    - Robotics and Control Systems
    - Interaction Modalities
3. Research Gap and Research Problem
4. Proposed System
    - Output Module (User Interface)
    - Interaction Module
    - Cyber Physical Agent
5. Multimodal HRI Case Studies
    - Multimodal Joystick Controller
    - Smart Manufacturing Cobot in VR/MR
    - Inclusive Cobot
6. Conclusion & Future Work

Human Robot Interaction

- Understanding, designing, and evaluating robotic systems for use by or with humans
- Investigating interactions between humans and robots
- Combine contributions from human–computer interaction, artificial intelligence, robotics, natural language understanding, design, humanities and social sciences



Literature Survey on Human Robot Interaction

Human Robot Interaction

Human Factors
- Task analysis
- Situation Awareness
- Workload
- Task performance
- Model of human and robot(GOMS for HRI)
- Learning Requirements
- Operator Fatigue

Robotics
- Degree of Freedom
- Kinematics
- End Effector
- Control System
- Teleoperation
- Mobility
- Workspace
- Autonomy

Interaction
- Modalities

Eye Gaze, Gestures, Speech, Facial Expressions, Emotions, Thoughts

- Interfaces

AR,VR,MR, Audio, Visual, Haptic, BCIs

Human Factors

Task Analysis
- Operator understanding of system, task and controls
- Depends on the skill, familiarity, fear, user culture and experience
- Improves with practice

Situational Awareness
- Operator understanding of the environment
- Depends on the user interface and immersiveness
- Improves with feedback- vision, haptic, VR

Workload
- Proportional to task analysis and complexity
- Varies on experience, system, controller complexity
- Performance varies inversely

Model of Human & Robot
- Mental representation of Robot and environment
- Allows intuitive control and operation
- Improves with UI and vision perception
- GOMS for HRI: generic procedure to approach a task

Learning Requirements
- Skill, time and efforts for training
- Varies inversely with system and controller complexity
- Important for users and designers
- Simple and easier systems are better

Operator Fatigue
- Arises when operating master–slave devices
- Degrades performance due to controller complexity
- Proficiency vs Tiresomeness trade off

Robotics

Degrees of Freedom (dof)

- Task and operation specific
- Higher the dof higher the complexity
- Multimodal fusion suitable for higher dof

Kinematics

- Forward: end effector position for given joint angles, simple to implement, suitable to low dof
- Inverse: Joint angles from given end effector position, computationally heavier, suitable for higher dof

End Effector

- Depends on the task
- Responsible for manipulating the environment

Teleoperation

- Remote operation of agent
- Controller and agent are geographically separated
- Telepresence required for intuitive control

Feedback Variations

- Direct Control with no feedback : Speed vs Accuracy
- Bilateral Force Feedback : tactile and kinesthetic information, complement vision, increase awareness
- Virtual Fixtures: Guidance or Forbidden region overlays on display (AR,VR,MR) like trajectories, force clues, needs apriori knowledge

Autonomy

- Remote Controlled: direct control, no autonomy
- Human-Assisted : Shared control (Man > Machine), Human as decision maker
- Human-Delegated: Shared control(Machine > Man), Human as guide
- Human-Supervised: Decision making and manipulation by machine, human for high level goal-oriented commands

Interaction

Interaction Modalities

- Eye Gaze: Point of gaze, hands free, can be non-intrusive, language independent, useful for especially abled
- Gesture: Hand, Head, Body. Non-verbal, versatile, can be non-intrusive, language independent, no mandate of UI/display
- Speech: Natural, easy, intuitive, non-intrusive,
- Facial Expressions: represent mental state(mood like pain, anger, excitement), non-verbal feedback, non-intrusive
- Emotions: motivate and modulate cognition and behavior, defines intentions and empathy
- Haptics: Being able to touch, feel, and manipulate objects in an environment, in addition to seeing (and hearing) them, provides a sense of immersion

Interface

- Master-Slave Controllers
- Semiautomatic Input Devices : Joysticks
- Computer Workstations : Laptop, Deskop, keyboard, Mouse
- Mobile Tablets : (like the one with cobot in smart factory)
- Invasive or non-invasive, wearable or handheld

Output / Display Types

- MR: AR / AV with overlaid controls, alerts, clues, depth information, Improves mental model
- VR: improve situational awareness, compensate delays, provide preview of actions, simulate environment
- Stereo Vision & Depth Imaging: Better end effector positioning, provide 3D data in 2D,

Research Gap and Research Problem

Research Gap

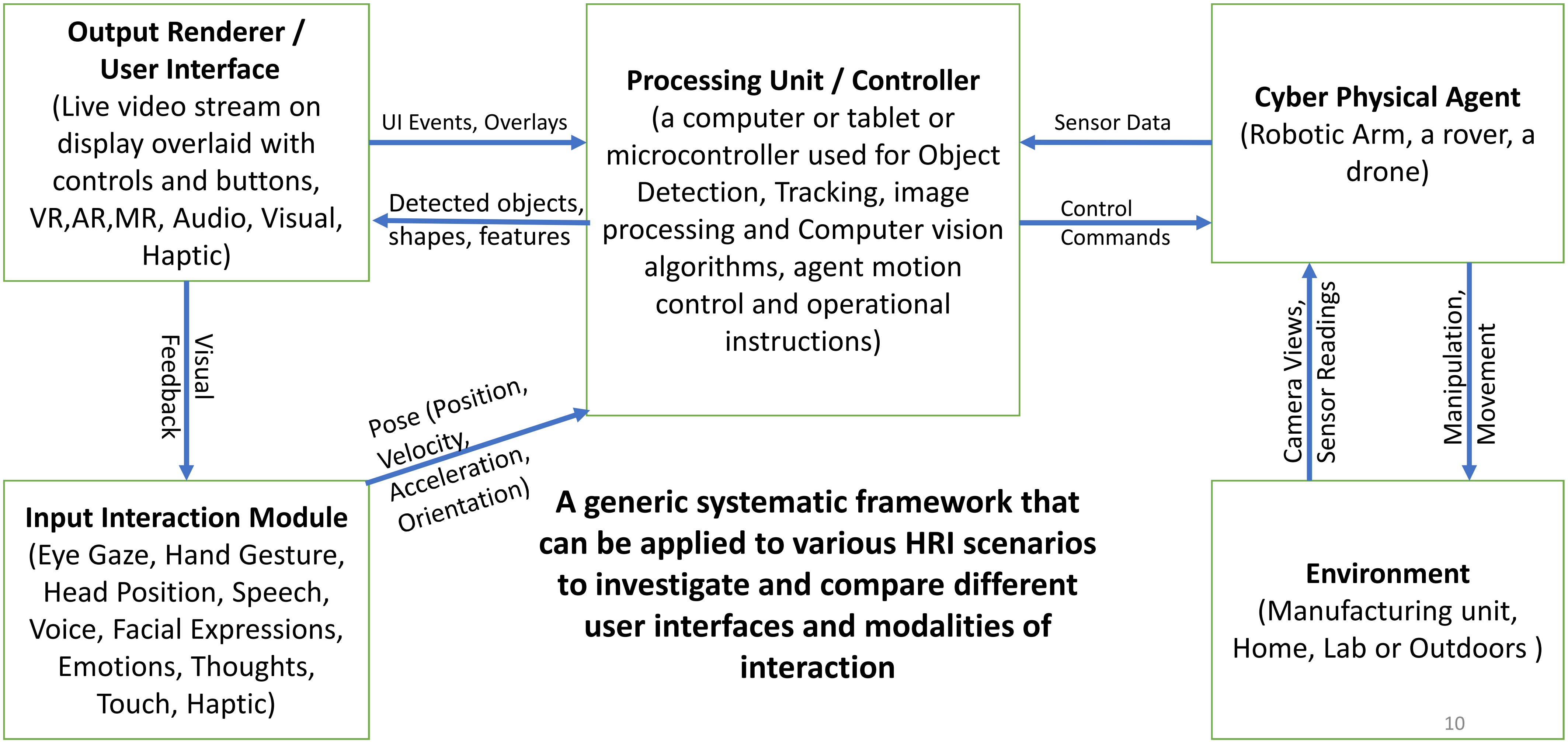
- Fusion of Multimodal Interactions system
- External Validity of HRI studies
- Sample Size
- Market Acceptance
- Non-Conventional Modalities

Research Problem

- Investigation and comparative analysis of available modalities of interaction
- Investigation and evaluation of possible fusion of multiple modalities
- Exploration of possible multimodal human robot interaction use cases and applications
- Finding out the suitable modalities for respective use-cases and validating effectiveness
- Exploration of suitable modalities for AR/VR/MR based applications

Proposed Approach

Multimodal Human Robot Interaction System



Implementation of Proposed HRI System

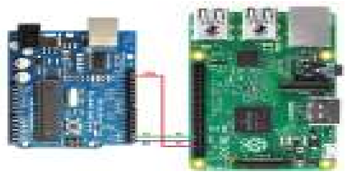
Output Module / User Interface

Input / Interaction Module

Processing Unit (Controller)

Computer and Tablets

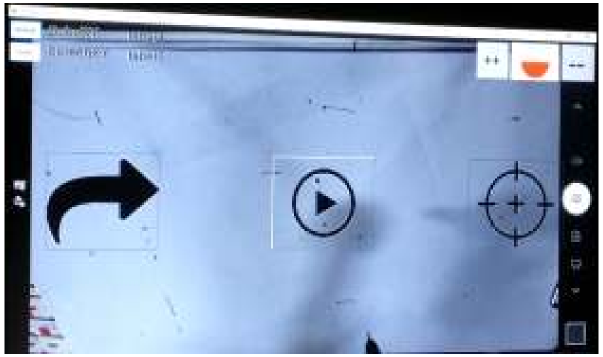
Microcontroller Arduino, R Pi



Cyber Physical Agent

Output Module (Multimodal User Interface)

- Live video **see-through display** with transparent controls
- UI preferably is
  - **Language independent**
  - **Noninvasive**
  - **Hands free**
- UI Controls are
  - Navigational (four-way controls, Eight Way)
  - Operational (Switch Modality, motion control, type & function of end effector)
- VR,AR,MR Interfaces
  - Immersiveness
  - Better situational awareness
  - Natural Feedback and Control
- Auditory and Haptic feedback interfaces
  - Complements vision
  - Sense of presence



Input Interaction Module

- Interaction module **houses the various sensors**
  - Eye-gaze (Tobii Eye Trackers)
  - Hand gesture (LeapMotion)
  - Speech (microphone)
  - Expressions (cameras)
  - Full body motion trackers (Microsoft Kinect)
  - Thoughts (Emotiv EEG headset)
  - Haptic Glove
- The user can
  - Switch between different modalities
  - use a combination
- Modalities are independent of each other making the system **modality agnostic**
- Currently **only eye-gaze tracking and hand gesture** tracking has been evaluated

*** We have all these in the I3D Lab**

Cyber Physical Agent

- Physical embodiment of the functional unit
- Equipped with sensors, power source and controls
- Capable of navigation and manipulation through the environment

Present Status

Affordable Do It Yourself 3dPrinted Robotic Arm: 4 degree of freedom , hobby servo motors ,limited motion and payload capacities

Dobot Magician Robot Arm: Industrial grade commercially available robust high precision table-top 4 degree of freedom robotic arm with multiple end-effectors

Autonomous Ground Vehicle (AGV): a remote-controlled toy car as a representation of a ground vehicle with camera mounted on roof

Unmanned Air Vehicle (UAV): a remote-controlled toy drone for the proof of concept

Case Studies

Multimodal HRI User Studies

1. **Multimodal Joystick Controller**
   - Semi-autonomous UGV – EG and Gesture controlled RC car with Waypoint Navigation and Landmark Detection

2. **Smart Manufacturing**
   - **Cobot:** Eye Gaze and Hand Gesture Controlled Robotic Arm for Pick & Place and Stamping Task with Object Detection and Feature Tracking
   - VR Based Co operation of Cobot and sAGV

3. **Inclusive Cobot:** Gaze Controlled Robotic Arm for Speech and Motor Impaired persons (SSMI)
   - Pick & Drop Task
   - Reachability Task

1. Multimodal Joystick Controller

- Designed and developed inhouse
- Can be used to **control any joystick, thumb-stick or a gaming console** of COTS product (like a robot arm, RC Car, Drone, machine tool)
- Allows **natural and intuitive modes of interaction** like eye-gaze, gesture, speech etc.. **Easy to learn** and does not require special training or skills
- Customizable & 3D Printed design

Eye-gaze & Gesture Controlled RC Car: Semi-Autonomous Ground Vehicle(sAGV)

What

Analyze Eye-gaze and gesture to control a ground vehicle

Why

- Understand eye-gaze and gesture-based control
- Color, shape and QR detection to improve operation
- Explore efficient user interface designs

How

- Drive the car from source to destination using the manual and autonomous modes
- Define Waypoints and landmarks
- Switch operation modes as per requirement

Uses

Transport materials, Logistics, monitoring, security and safety, remote operation

2. Smart Manufacturing Cobot

Collaborative Robot (Cobot)

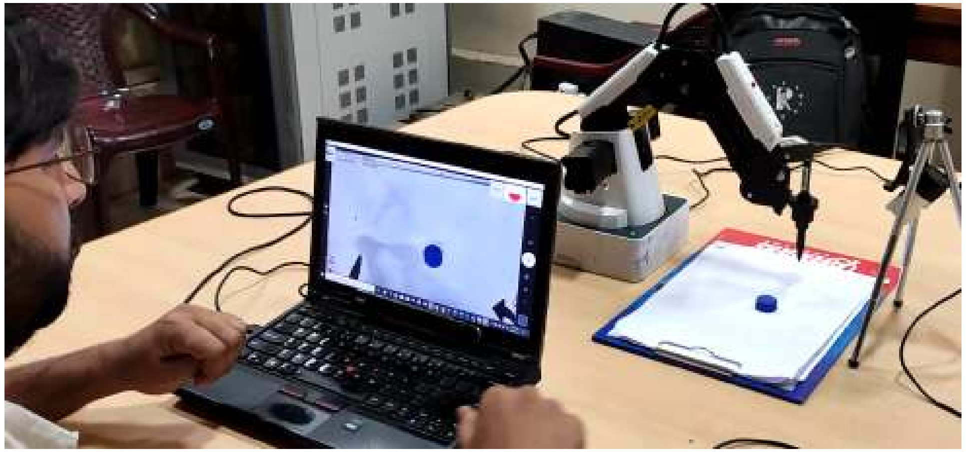
- Operator aware and intelligent
- Safe and smart
- Potentially replace low-level manual tasks

Aim

- Explore possible Human Robot Interaction in a manufacturing set up and analyze the effectiveness of such a system.
- Comparative analysis of different modalities of interaction

Design

- Eye gaze and hand gesture controlled robotic arm system to perform simple manufacturing tasks.



Multimodal Cobot for Manipulation

What

Analyze eye-gaze and gesture modalities to control a robotic arm for a stamping task

Why

- Use eye-gaze and gesture to operate a robotic arm
- Explore possible tasks
- Improve safety. Allow remote operation

How

- Stamp on the paper at the desired location using eye-tracking, gestures with the multimodal UI
- Point-to-Point motion control for bigger strides
- Four-way control for precision Control
- Stamping icon for pressing down the stamp

Applications

- Feature detection and tracking to track a blue object (can be operator's finger, teach pendant, marker etc.)
- Laser cutting, welding, drawing, sawing etc

VR Based Co-operation of Cobot and sAGV

Explore the Smart Factory in VR

Fusion of Multimodal User Interface with VR for a task involving two users operating sAGV and Cobot simultaneously using non- traditional modalities of interaction

Drive the sAGV in VR Interface

Autonomous Pick & Drop

Park the sAGV in Cobot Workspace

3. Inclusive Cobot

Speech and Motor Impaired persons(SSMI)

- SSMIs are not capable using natural hand control, speech and gestures
- **Reason:** Damage to brain or spinal cord leading to involuntary movements
- **Cause**: By Birth, Accident, Disease or other complications

Aim

- Explore possible multimodal Human Robot **interaction in assistive** and rehabilitation robotics
- Design and analyze a multimodal HRI system that can be employed to support and **empower differently abled persons**

Design

Two tasks involving eye tracking based manipulation of a robotic arm to move objects and draw in the robot workspace

1. Pick and Drop Task
2. Reachability Task

User	Age(yrs)	Gender	NIOS Std	Finger	Head	Speech
A	12	F	10th	No	Nod	No
B	20	F	12th	No	Nod	Little
C	15	M	10th	No	No	No(Blinks)
D	6	F	3rd	Point	No	Little
E	6	M	2nd	No	Little	No(Smiles)
F	12	F	PreVoc.	Yes	Yes	Little
G	7	M	2nd	Yes	Little	No(Blinks)
H	20	F	12th	Yes	Yes	Yes
I	23	M	12th	No	Yes	Little

Pick & Drop Task

What

Analyze the effectiveness of eye gaze controlled robotic arm by SSMIs for pick and drop task

Why

- Enable SSMI to manipulate
- Utilize unexplored modalities of SSMI
- Education and Entertainment

How

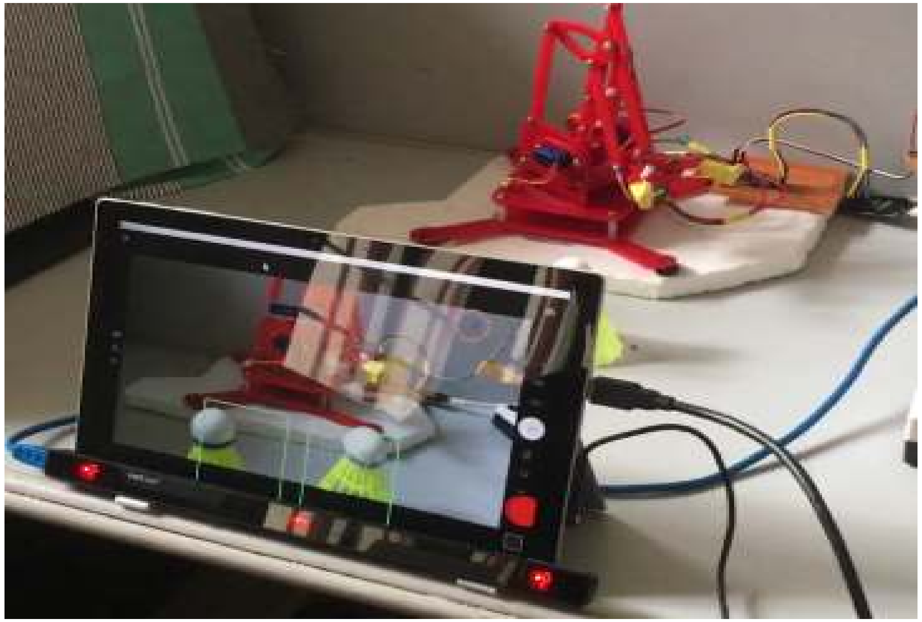
- Pick object from designated source and place on a destination using eye gaze controlled robotic arm
- Users must only use eye-gaze
- Dwell for 500 ms for pick and drop
- Scaled eye gaze 2D control of robotic arm

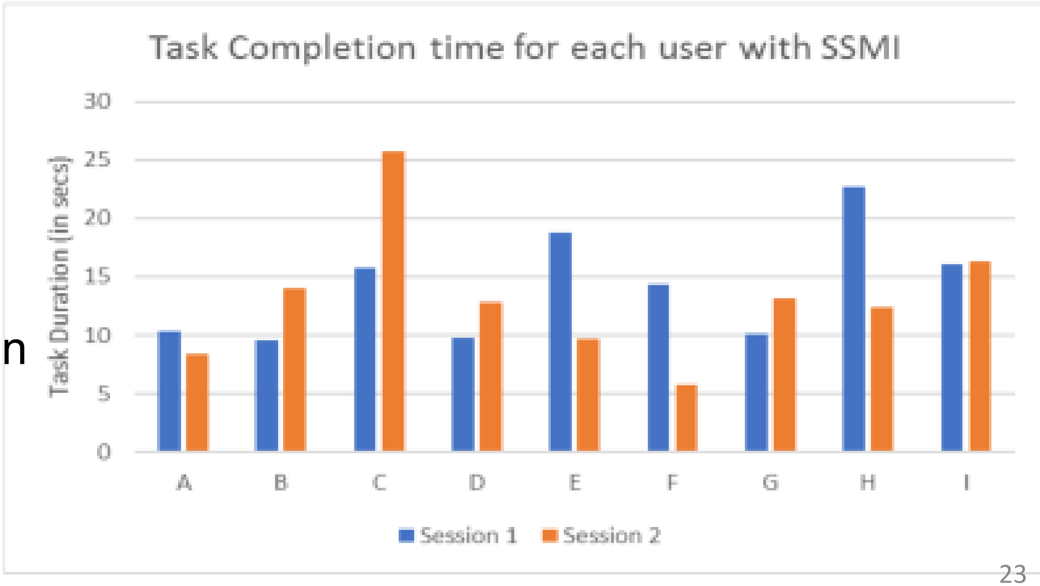
Results

All 18 participants performed successfully, and we found no significant difference in task completion time between participants groups H(1,16)=2.94, p>0.05

Application

- SSMIs Could hold and place objects
- Training for tasks like fabric printing, assembling, stamping





Reachability Task

What

Analyze eye-gaze control of a pen attached to Robotic arm

Why

- Explore Eye-gaze to write and draw
- Explore effective UI designs
- Understand SSMI behavior

How

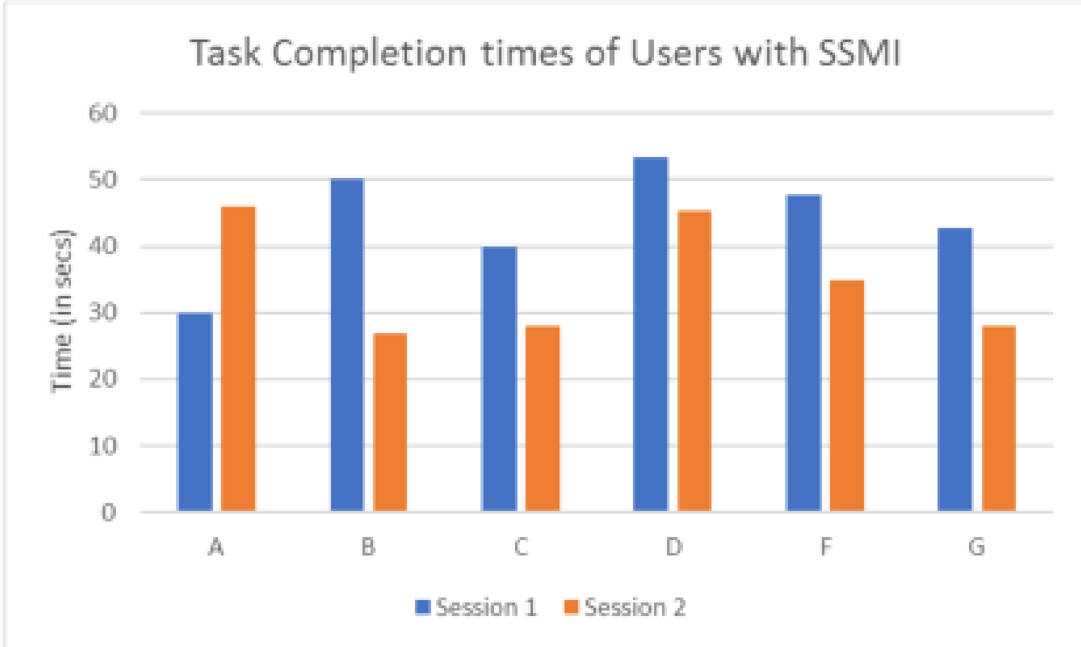
- Use eye gaze to reach the designated target from a randomly assigned starting point on a paper
- Video see through four-way control on the UI
- Dwelling for selection

Results

- All 12 participants were able to complete the task successfully
- Task completion time reduced in 2nd session implying that practice can improve task performance

Applications

- SSMI can draw & write
- Rehabilitation and motor control exercises



Future Work

Research | **Applications**

Future Work: Research

- Explore and investigate various **non-traditional modalities** of interaction like hand-body gestures, speech, expressions, thoughts, emotions etc.
- Explore and investigate **possible use-cases** for Human Robot Interaction
- Investigate and **classify** best suited **interaction modality** for respective use cases and applications
- Investigate the **role of human factors** like emotion, cognition, experience, perception, situation awareness etc. and other unknown(now) factors in a successful human robot interaction
- Analyze factors and dependencies of **adoption** of human robot interaction systems in existing infrastructure

Future Work: Applications

- Telepresence: Rendering realistic feedback in real time
- Telecare and Telemedicine:
  - ✓ Multimodal contact less interaction between contagious patients, doctors, relatives
  - ✓ Mobile manipulators to take samples, deliver medicines, cleaning and disinfecting the area
  - ✓ Monitor vitals and record data (like temperature, heartbeat, blood pressure)
- Military, Police, Defense and Airforce :
  - ✓ Multimodal intelligent autonomous search and rescue bots for disaster management
  - ✓ Ensuring safety of personnel and staff
  - ✓ multimodal AGVs and Drones capable of remote operation and manipulation using non-traditional modalities of interaction
- Agriculture:
  - ✓ Remote monitoring and crop health maintenance through drones and AGVs
  - ✓ Telegardening, watering and cultivation
- Nuclear Power Plant
  - ✓ Radio-active material and waste handling
  - ✓ Remote multimodal operation of robotic manipulators and logistic robots
- Assistive & Rehabilitation Robotics:
  - ✓ Social inclusion, education , training and entertainment for especially abled
  - ✓ Rehabilitations and motor control therapy and exercises

Conclusion

- Non-traditional modalities of interaction like eye gaze and gesture have the potential to be used as natural mode of interaction as these are intuitive, **easy to learn** and requires minimum skill
- Using non-traditional modalities of interaction may **improve operator safety**, product quality and repeatability.
- Coupling non-traditional interaction modalities with some level of autonomy **reduces human effort** and allow human to focus their skills on high level and more productive tasks
- Eye tracking **enabled the limited mobility persons** to pick & place and write and draw which was otherwise not possible for them physically

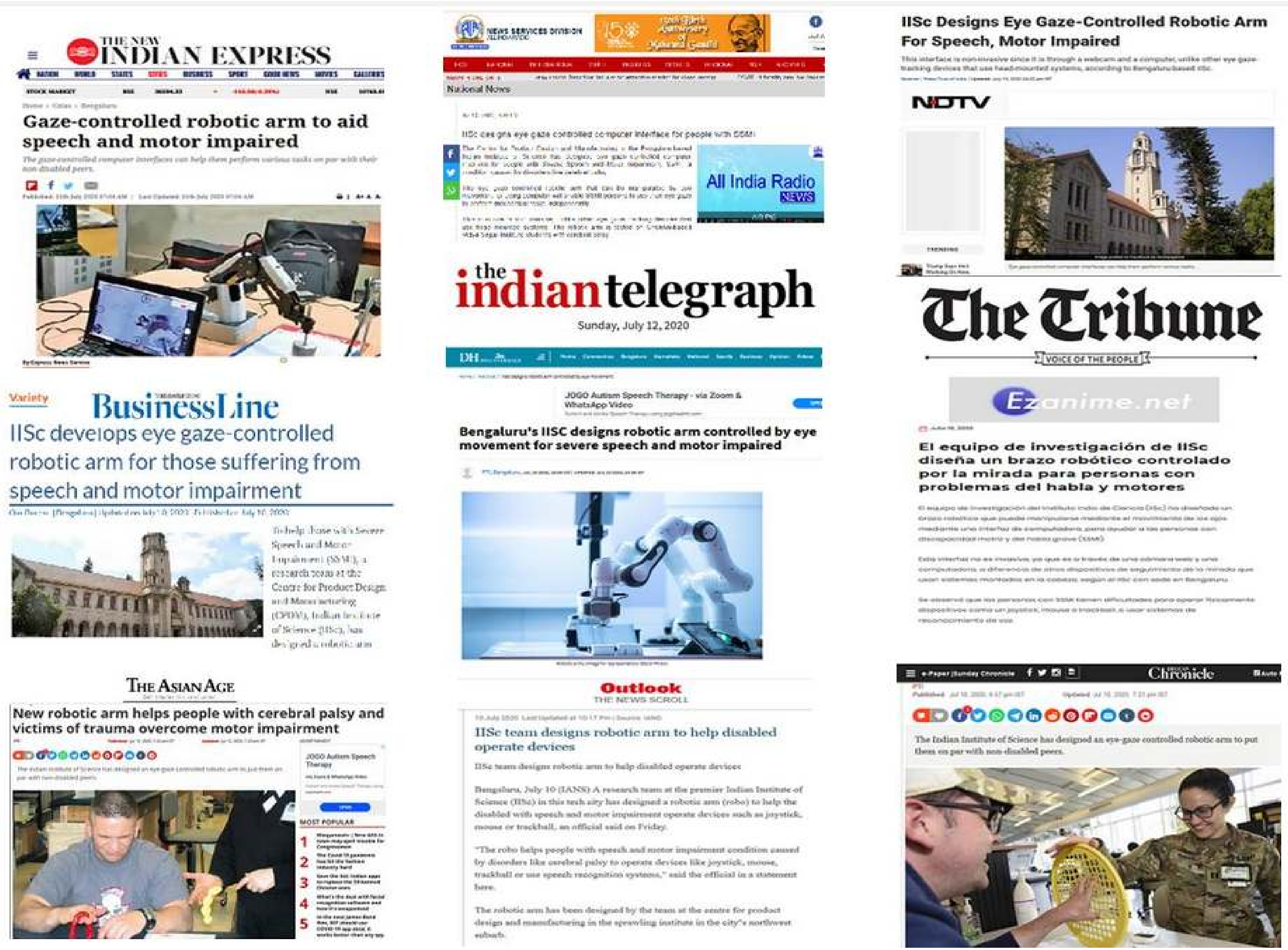
Publications

- Sharma, V.K., Saluja, K., Mollyn, V., & Biswas, P. (2020, June). Eye Gaze Controlled Robotic Arm for Persons with Severe Speech and Motor Impairment. In ACM Symposium on Eye Tracking Research and Applications (pp. 1-9).
- Sharma, V. K., Murthy, L. R. D., Singh Saluja, K., Mollyn, V., Sharma, G., & Biswas, P. (2020). Webcam controlled robotic arm for persons with SSMI. Technology and Disability, (Preprint), 1-19.
- G Prabhakar, A Ramakrishnan, LRD Murthy, VK Sharma, M Madan, S Deshmukh and P Biswas, Interactive Gaze & Finger controlled HUD for Cars, Journal of Multimodal User Interface, Springer, 2019

Patents

- Sharma V. K. and P. Biswas, System for Operating Joystick, Indian Patent Application no. 201941044740
- P. Biswas, S. Deshmukh, G. Prabhakar, M. Modiksha, V. K. Sharma and A. Ramakrishnan, A System for Man- Machine Interaction in Vehicles, Indian Patent Application No.: 201941009219, PCT International Application No. PCT/IB2020/050253

Press Coverage



DEMO

3. Virtual Reality Digital Twin

Aim: Design and test a system where a user can drive the sAGV and operate the Cobot from a virtual environment (virtual replica of smart factory).

Application: Simulated immersive environment, used for training of multiple persons, allows teleoperation, saves energy and time, ensures safe operation

Methodology

- Integrated the multimodal UI in the VR environment to operate the sAGV in a virtual reality environment wearing a VR Headset (Oculus or HTC Vive)
- The task here was a bit complex and collaborative in nature as it required two users simultaneously operating the system (one for driving sAGV other for controlling the Cobot)
- The user must drive the semi-autonomous ground vehicle carrying a package in VR from a starting point to the workspace of the Cobot
- The driving controls were transferred from the UI to the VR joystick controllers
- As the user drive the sAGV near to the Cobot, the other user tries to pick up the package using natural eye-gaze and gestures.

Result: The users were able to operate and complete the task with little bit to intervention and help.

Literature Survey on Human robot Interaction

Human Factors

- Task analysis
- Situation Awareness
- Workload
- Task performance
- Model of human and robot
- GOMS for HRI(explain GOMS in brief)
- Major human factors research challenges include- task analysis that includes dynamics, economics, and other factors; considering how both human and robot have mutual models of each other; coping with user culture, fears, and other value considerations(Sheridan, T. B. (2016). Human–robot interaction: status and challenges. Human factors, 58(4), 525-532.)
- Human factor like situation awareness, workload, Task performance affects the success of a human robot interaction System(Bray-Miners, J., Ste-Croix, C., & Morton, A. (2012). Human-robot interaction literature review. HUMANSYSTEMS INC GUELPH (ONTARIO).)
- GOMS can be used to compare different interfaces for human-robot interaction and is useful in determining the user's workload, for example when introducing different displays for sensors(Jill L. Drury, Jean Scholtz and David Kieras (2007). The Potential for Modeling Human-Robot Interaction with GOMS, Human Robot Interaction, Nilanjan Sarkar (Ed.), ISBN:978-3-902613-13-4)

Robotics and Control Systems

- Kinematics of popular manufacturing robots were surveyed, and it was found that research is largely driven by the electronics and automotive industries, but as cobots become cheaper and easier to integrate into workcells, we can expect SMEs from a wide range of industrial applications to lead their adoption.(Matheson, E., Minto, R., Zampieri, E. G., Faccio, M., & Rosati, G. (2019). Human–Robot Collaboration in Manufacturing Applications: A Review. Robotics, 8(4), 100.)
- The implementation of multi-modal human-robot interfaces presents a big advantage in addressing complex interaction scenarios and an appropriate trade-off between the complexity and completeness of human sensorimotor behaviour and intention modelling must be defined.(Ajoudani, A., Zanchettin, A. M., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., & Khatib, O. (2018). Progress and prospects of the human–robot collaboration. Autonomous Robots, 42(5), 957-975.)

Interaction Modalities

- The use of remote-control manipulators was likely due to the unstructured nature of operational environments for SUS, for example, in search and rescue missions. When systems cannot know the remote environment a priori, manual or computer aided teleoperation becomes a default operation mode. Different types of advanced visualization aid in direct telemanipulation, with a majority of systems relying on some form of real-world imagery through a live video feed.(Young, S. N., & Peschel, J. M. (2020). Review of Human–Machine Interfaces for Small Unmanned Systems With Robotic Manipulators. IEEE Transactions on Human-Machine Systems, 50(2), 131-143.)
- A novel approach was presented for multimodal interactions between humans and industrial robots using direct voice input, eye gaze and soft buttons in an assembly process. (Bannat, A., Gast, J., Rehrl, T., Rösel, W., Rigoll, G., & Wallhoff, F. (2009, July). A multimodal human-robot-interaction scenario: Working together with an industrial robot. In International conference on human-computer interaction (pp. 303-311). Springer, Berlin, Heidelberg.)
- Alsharif and colleagues configured eye gaze movement, eye blinks and winks to the 7 degrees of freedom of a robotic arm and evaluated performance of the system with 10 participants including one person with motor impairment for a block rearrangement task. (Alsharif S., Kuzmicheva O. and Gräser A., Gaze Gesture-Based Human Robot Interface, Zweite transdisziplinär Konferenz,Technische Unterstützungssysteme, die Menschen wirklich wollen, 2016)
- Stiefelhagen and colleagues investigated direct voice input, pointing gestures and head orientations and reported results on accuracy of each individual modality.(Stiefelhagen, R., Ekenel, H. K., Fugen, C., Gieselmann, P., Holzapfel, H., Kraft, F., & Waibel, A. (2007). Enabling multimodal human–robot interaction for the Karlsruhe humanoid robot. IEEE Transactions on Robotics, 23(5), 840-851.)
- Eye tracking is traditionally used for analyzing visual perception, eye gaze movement and making visual perception models(Biswas P. and Robinson, Modelling Perception using Image Processing Algorithms, 23rd British Computer Society Conference on Human-Computer Interaction (HCI09))
- Eye-gaze controlled user interfaces have also been explored for able bodied users in situational impairments like drivers and pilots, whose hands are engaged with the primary task of driving or flying(Biswas P. and Jeevithashree DV, Eye Gaze Controlled MFD for Military Aviation, ACM International Conference on Intelligent User Interfaces (IUI) 2018)
- Srinivasan, M. A., & Basdogan, C. (1997). Haptics in virtual environments: Taxonomy, research status, and challenges. *Computers & Graphics*, *21*(4), 393-404.
- Srinivasan, M. A. (1995). What is haptics?. Laboratory for Human and Machine Haptics: The Touch Lab, Massachusetts Institute of Technology, 1-11.

Contents

Research Gap and Research Problem

Research Gap

- Fusion of Multimodal Interactions system have not been explored exhaustively
- HRI research requires an internally valid setup
- Most of the systems have not been validated externally on actual users
- Adoption of multimodal HRI systems is still complex and highly interdependent
- Eye Tracking to manipulate a robotic arm has not been investigated thoroughly

Research Problem

- Investigation and comparative analysis of available modalities of interaction
- Investigation and evaluation of possible fusion of multiple modalities
- Exploration of possible multimodal human robot interaction use cases and applications
- Finding out the suitable modalities for use-cases and validate and its effectiveness
- Exploration of suitable modalities for AR/VR/MR based applications

https://asegrad.tufts.edu/academics/explore-graduate-programs/human-robot-interaction