IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

Survey on Sign Language Translation:Integrating Metahuman Avatars for Human Language Input Expression

Gangothry Nakirikanti[1], Venkata Vaibhavi Gedela[2], Pooja Pilla[3], Dr. D. Shravani[4]
Students, Department of Artificial Intelligence and Data Science[1,2,3]
Associate Professor, Department of Artificial Intelligence and Data Science[4]
Osmania University, Hyderabad, India

Abstract: *The deaf and mute population in the Indian subcontinent relies on Indian Sign Language (ISL) for communication, a comprehensive language employing hand gestures, facial expressions, and body postures. To address the communication gap, we present an end-to-end framework that recognizes spoken language and performs corresponding ISL gestures in real time. Leveraging Unreal Engine's depth sensing and motion capturing, we collected motion data for various ISL gestures. Unity3D facilitated animation setup for a diverse range of ISL expressions. Integrated into an application, our framework enables convenient, real-time conversations, fostering inclusivity. Beyond recognition, the system completes the communication loop by converting ISL to human language. Notably, Metahuman avatars are employed to present the ISL gestures, adding a human-like and expressive dimension to the communication process. This approach, leveraging technology for natural language translation and ISL interpretation through Metahuman avatars, stands as a significant step toward bridging the communication gap, promoting understanding, and enhancing accessibility for the hearing and speaking impaired community.*

Keywords: 3D Signing avatar, Animation framework, Indian Sign Language, Natural Language processing

I. INTRODUCTION

The intricate language of Indian Sign Language (ISL), comprising expressive hand gestures, facial cues, and body postures, forms the cornerstone of communication for the deaf and mute community in the Indian subcontinent. Recognizing the communication challenges faced by this community, and with the overarching goal of fostering inclusivity, we introduce an innovative end-to-end framework. This framework seamlessly combines spoken language recognition with the dynamic presentation of ISL gestures, using cutting-edge technologies such as Unreal Engine's depth sensing and motion capturing capabilities.

As a comprehensive solution, our framework is designed not only to facilitate real-time conversations between the hearing and speaking impaired and the wider population but also to serve as an educational tool for learning ISL. The utilization of Metahuman avatars adds a human-like and expressive layer to the presentation of ISL gestures, enhancing the learning experience.

Beyond its educational impact, our system contributes to breaking down communication barriers, promoting cultural exchange, and fostering a deeper understanding between communities. The integration of Metahuman avatars ensures that the nuances and intricacies of ISL are authentically conveyed, adding a unique human touch to the communication process. This holistic approach marks a significant stride towards a more inclusive, accessible, and empathetic future for the deaf and mute community.

In addition to its role in facilitating real-time communication and serving as an educational tool, our framework presents several advantages that extend its societal impact. By harnessing the immersive capabilities of Unreal Engine and Unity3D, the system creates a platform where individuals can actively engage and learn the art of ISL, nurturing a sense of community and cultural exchange. Moreover, the incorporation of Metahuman avatars introduces a unique form of visual storytelling, enabling the wider public to better comprehend the richness of ISL

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

as a distinct language. This not only enhances awareness but also contributes to breaking down societal stigmas surrounding hearing and speech.

II. LITERATURE REVIEW

1. Speech to Indian Sign Language Translation Systems:

The study conducted by Pankaj Sonawane et al. (2021) introduces a Speech to Indian Sign Language Translation System, leveraging Microsoft Xbox Kinect 360 and Unity3D for motion capture and animation. This research emphasizes the significance of technological solutions in bridging communication barriers for individuals with hearing impairments. The utilization of motion capture and animation technologies showcases the potential for real-time conversation between the hearing-impaired community and the wider population, addressing the critical need for effective communication tools in diverse linguistic contexts.

2. Speech/Text to Indian Sign Language Using Natural Language Processing:

Anannya Priyadarshini Neog and Arunabh Kalita (2022) contribute to the field by developing a system that translates speech/text into Indian Sign Language (ISL) using Natural Language Processing (NLP) techniques. Focusing on the challenges associated with the diversity of sign languages, the study explores the potential of NLP-based systems to address communication barriers. The successful translation of speech/text into ISL signs using Deep Neural Network and NLP concepts demonstrates the efficacy of the developed model in facilitating dynamic and accurate sign language translation.

3. Translating Speech to Indian Sign Language Using Natural Language Processing:

Purushottam Sharma et al. (2022) presents a system for translating speech to Indian Sign Language (ISL) using natural language processing. This research aims to bridge the communication gap between hearing-impaired individuals and the broader society. The study involves converting audio or text input into ISL videos based on grammar rules and a database of ISL videos, showcasing the potential of NLP techniques in enhancing accessibility and communication for the deaf and hard-of-hearing community.

4. 3D Signing Avatar for Sinhala Sign Language:

M. Punchimudiyanse and R. G. N. Meegama (2015) contributed to the field with the development of a 3D signing avatar and animation system for Sinhala Sign Language (SSL). The system's unique approach utilizes Blender software for animation, offering flexibility and independence from expensive motion capture hardware. This innovative system provides a comprehensive solution for animating sign language gestures, demonstrating its potential applicability to various sign languages beyond Sinhala.

5. A Review of the Hand Gesture Recognition System:

Noraini Mohamed et al. (2021) provide a comprehensive review of vision-based hand gesture recognition systems, with a specific focus on sign language recognition. The review encompasses 98 articles published from 2014 to 2020, highlighting the active research conducted in this area. The study emphasizes the challenges faced in developing robust signer-independent recognition systems and sets the stage for future research in efficient sign language recognition systems.

6. Hand Gesture Recognition for Sign Language Using 3DCNN:

Muneer Al-Hammadi et al. (2020) delve into the topic of hand gesture recognition for sign language using a 3DCNN approach. Motivated by the growing deaf and hearing-impaired population, the research proposes an efficient deep learning approach that addresses spatial and temporal features in sign language analysis. The study's experimental results showcase high recognition rates for both signer-dependent and signer-independent modes, contributing valuable insights to the field of hand gesture recognition.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

7. Neural Decoding of Chinese Sign Language:
Wang et al. (2021) investigate the neural decoding of Chinese sign language using electroencephalography (EEG) signals and machine learning for brain-computer interfaces (BCIs). The study explores the potential of EEG-based decoding for sign language, providing insights into the neural characteristics associated with sign language. The findings suggest promising applications for BCI in communication, rehabilitation, and human-computer interaction.

8. Development of an End-to-End Deep Learning Framework:
B. Natarajan et al. (2022) introduce the H-DNA framework, a novel deep learning approach for sign language recognition, translation, and video generation. This innovative framework combines Neural Machine Translation (NMT), MediaPipe, and Dynamic Generative Adversarial Network (GAN) models to produce high-quality sign gesture videos for spoken sentences. The study's detailed technical explanations and evaluations showcase the framework's effectiveness in handling sign language recognition and translation tasks, representing a significant advancement in the field.

9. Metahuman Motion Detection:
Gianmarco Thierry Giuliana's paper (2022) titled "Metahuman Creator: A Semiotic Analysis of Realistic Digital Faces" explores the features and perceived realism of the software "Metahuman Creator." While not explicitly focused on sign language, the paper introduces an analysis of realistic digital faces and highlights the semiotic processes that contribute to their perceived realism. Although not directly related to sign language, the concept of Metahuman motion detection could potentially play a role in future sign language recognition systems. The paper delves into the technological features of the software and the characteristics of the artificial faces it produces. While the primary focus is on facial expressions and realistic digital faces, the concept of motion detection within the Metahuman Creator framework could be adapted or extended to incorporate gesture and motion-based sign language recognition in future developments. This addition underscores the interdisciplinary nature of research, where insights from fields such as computer graphics and animation can potentially inform advancements in sign language recognition technology.

III. PROPOSED SYSTEM

In our groundbreaking communication framework, the intricate actions, and motions inherent in Indian Sign Language (ISL) are meticulously captured using Unreal Engine's advanced depth sensing and motion capturing capabilities. These captured actions are then intricately mapped to corresponding words, phrases, or sentences within Unity3D, forming a comprehensive library of expressive ISL gestures.

Embarking on the forefront of inclusive communication solutions, our pioneering framework endeavors to bridge the gap between the hearing and speaking impaired community and the wider population. Leveraging advanced technologies, we meticulously capture the intricacies of Indian Sign Language (ISL) through Unreal Engine's depth sensing and motion capturing capabilities. These actions are artfully mapped to linguistic elements within Unity3D, fashioning a dynamic library of expressive ISL gestures. The user interface, seamlessly integrated into an Android application, serves as a gateway to a novel and immersive communication experience. Upon receiving user input, whether in spoken language or text, our system dynamically interprets the input, triggering corresponding ISL actions presented through Metahuman avatars. Beyond real-time communication, our system stands as a versatile educational tool, offering an interactive platform for actively learning the nuances of ISL. Bidirectional interpretation capabilities ensure that our framework facilitates inclusive and comprehensive communication, marking a transformative stride toward a more connected and empathetic society.

User Interaction:
Upon receiving user input, whether in the form of spoken language, text, or sign language gestures, our system dynamically interprets the input and triggers the corresponding ISL action. The integration of Metahuman avatars adds a human-like and expressive layer to this process, ensuring a more engaging and relatable communication experience.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

Seamless Integration:
The seamless integration of our system into an application enhances accessibility and usability, allowing users to engage in real-time conversations effortlessly. The Android application acts as a user- friendly interface, bridging the gap between the hearing and speaking impaired and the wider population.

Educational Aspect:
Beyond its role in facilitating real-time communication, our system functions as a valuable educational tool for learning ISL. The pre- captured actions, intricately mapped to linguistic elements, provide users with a dynamic and immersive environment for actively learning the nuances of sign language.

Sign Language Interpretation:
An additional pivotal feature of our system is its ability to read sign language gestures input by the user and interpret them into human language. This bidirectional capability ensures that communication is not only initiated by spoken or text input but also by the rich and expressive medium of sign language.

Advantages:

- Real-time Responsiveness: The system offers immediate and real- time responsiveness, ensuring a fluid and natural communication experience for users.
- Expressive Communication: Metahuman avatars enhance the expressiveness of ISL, making the communication process more engaging and relatable.
- Usability and Accessibility: The Android application ensures that the system is accessible to a wide range of users, contributing to its usability and reach.
- Learning Platform: With its mapped actions and educational focus, the system becomes a valuable resource for individuals looking to learn and understand ISL.
- Cultural Inclusivity: By providing an immersive representation of ISL, our system contributes to breaking down cultural barriers, fostering a more inclusive and empathetic society.
- Innovative Technological Bridge: The system acts as an innovative bridge, merging technology, language learning, and expressive communication for the benefit of the hearing impaired and speech impaired community.
- Enhanced Social Inclusion: The system contributes to enhanced social inclusion by providing a platform for meaningful and expressive communication, fostering connections between individuals irrespective of their linguistic abilities.
- Global Outreach: With its intuitive design and technological sophistication, the system opens avenues for global outreach, enabling users worldwide to engage, learn, and communicate in Indian Sign Language, promoting cross-cultural understanding.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

IV. BLOCK DIAGRAM

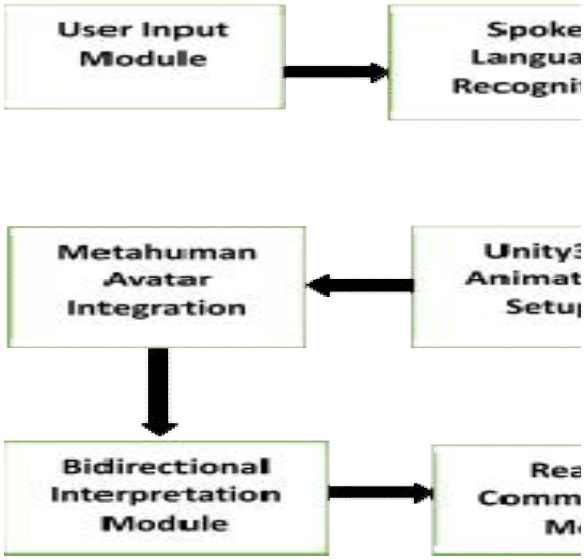

Fig. 1 Model of System Architecture.

V. METHODOLOGY

- Sign Language Translation: The process of converting sign language gestures into another form, often spoken or written language, to facilitate communication between individuals who use different modes of expression.
- Indian Sign Language (ISL): A natural language used by the deaf and mute community in India, employing hand gestures, facial expressions, and body postures for communication.
- Metahuman Avatars: Digital representations of humans created using advanced technologies, enhancing the expressiveness and relatability of virtual communication.
- Real-time Communication: Immediate and instantaneous exchange of information between users, allowing for fluid and natural conversations without delays.
- Spoken Language Recognition: Technology that interprets and understands spoken language inputs, often used as a component in systems facilitating communication with deaf or mute individuals.
- Text Input Processing: The handling and interpretation of textual input, preparing it for further processing or mapping in a communication system.
- Gesture Mapping: The association of specific gestures with linguistic elements, such as words, phrases, or sentences, forming a mapping for expressive communication.
- Unity3D Animation: Utilization of the Unity3D platform for creating dynamic and interactive animations, often employed in virtual environments or applications.
- Unreal Engine: A powerful game development engine that supports advanced graphics, physics, and immersive experiences, used in this context for capturing sign language gestures.
- Educational Tool: A resource or system designed to facilitate learning, in this case, providing an interactive platform for individuals to actively learn and understand sign language.
- Cultural Inclusivity: Promoting the understanding and inclusion of diverse cultural and linguistic communities, ensuring that communication tools are accessible and relatable to all.
- Natural Language Processing (NLP): A field of artificial intelligence focused on enabling machines to understand, interpret, and generate human language, often used in speech and text recognition.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

- Depth Sensing: Technology that measures the distance of objects, often used in gesture recognition systems to capture the three-dimensional aspects of movements.
- Motion Capturing: The process of recording movements of objects or people, typically for animation or analysis, used here to capture sign language gestures.
- Bidirectional Interpretation: The ability to interpret and translate communication in both directions, allowing for seamless interaction between different language modalities.
- Inclusive Communication: Ensuring that communication tools and technologies are accessible and accommodating to individuals with diverse communication needs.
- Deaf and Mute Community: The community of individuals who are deaf or mute, often facing communication challenges that technologies like this aim to address.
- Communication Accessibility: Making communication tools and platforms accessible to individuals with varying abilities and linguistic preferences.
- Human-Computer Interaction: The study and design of interactions between humans and computers, focusing on creating user-friendly and intuitive interfaces.
- Expressive Communication: Communicating in a way that conveys emotions, intentions, and nuances, often achieved through expressive gestures or language.
- User Interface: The visual and interactive elements of a system that users interact with, encompassing screens, pages, and elements facilitating communication.
- Linguistic Bridge: A tool or system that serves as a bridge between different languages or communication modalities, fostering understanding and connection.
- Dynamic ISL Gestures: Animated and varied hand movements, facial expressions, and body postures used in Indian Sign Language to convey dynamic and expressive messages.

These keywords encapsulate the core components, features, and aspects of your proposed system, covering technological, educational, and cultural dimensions.

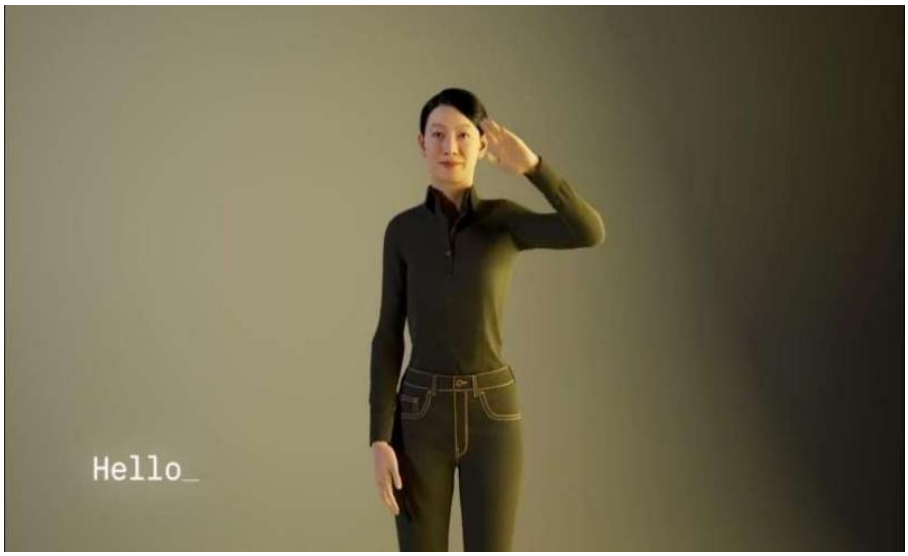

Fig 2 Metahuman showing word hello

VI. DISCUSSIONS

The proposed system opens important discussions on several fronts. First and foremost, it holds significant potential for enhancing communication access for the deaf and mute community. By facilitating real-time interaction and understanding between individuals using different modes of communication, the system can bridge longstanding gaps in communication. Additionally, the integration of an educational component is noteworthy, offering a promising avenue for fostering learning and understanding of Indian Sign Language (ISL).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

This dual-purpose system acts as a communication facilitator and an educational tool, system acts as a communication facilitator and an educational tool, catering to the needs of both the deaf and mute community and those interested in learning sign language. Furthermore, the incorporation of cultural elements within the virtual communication experience contributes to cultural inclusivity, promoting a more understanding and inclusive society. The technological innovation demonstrated using advanced technologies such as Unreal Engine, Unity3D, and Metahuman avatars showcases the potential for groundbreaking solutions to communication challenges.

Moreover, the system's adaptability to various platforms and devices, such as smartphones, tablets, and VR headsets, makes it accessible to a wide range of users.The user-friendly interface and intuitive design further enhance the accessibility and usability of the system, ensuring that it can be easily adopted by people with varying levels of technological proficiency.In conclusion, this innovative system holds significant promise for improving communication access and fostering educational opportunities for the deaf and mute community. By leveraging advanced technologies and emphasizing cultural inclusivity, it has the potential to make a meaningful impact on individuals' lives and contribute to a more inclusive and understanding society.

VII. CHALLENGES

Despite its promise, the proposed system faces several challenges that warrant careful consideration. One primary challenge lies in ensuring the accuracy of gesture recognition. Variability in gestures among users and potential environmental factors may impact the system's accuracy, requiring robust recognition algorithms. Adapting the system to cater to different sign languages requires careful consideration of linguistic variations. Designing an intuitive and user-friendly interface that accommodates the diverse needs of users poses a challenge, particularly for individuals with varying levels of technological familiarity. Additionally, integrating the proposed system with existing communication platforms and technologies requires addressing compatibility issues for seamless adoption. Ensuring the privacy and security of user data, given the sensitive nature of communication, becomes a priority. Users, especially those new to technology or sign language learning, may face a learning curve in effectively using the system, necessitating the design of user-friendly tutorials and support features. Achieving a dynamic and natural representation of ISL gestures through Metahuman avatars requires careful consideration of capturing the nuances of facial expressions, hand movements, and body postures to convey meaning authentically. Adapting the system for global use and ensuring cultural sensitivity in the representation of sign language gestures across diverse communities is a complex task that demands meticulous attention. Finally, the technological requirements of the system may make it resource-intensive, necessitating efforts to ensure accessibility on various devices and network conditions.

Addressing these discussions and challenges will be crucial in refining and optimizing the proposed system, making it a robust and inclusive solution for effective communication in diverse linguistic and cultural contexts.

Fig. 3 process showing speech to text.

VIII. CONCLUSION

In conclusion, the envisioned communication framework, leveraging cutting-edge technologies and innovative approaches, presents a transformative solution to the challenges faced by the deaf and mute community. The

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

system not only acts as a bridge for real-time communication but also serves as a comprehensive educational tool for learning Indian Sign Language (ISL). The incorporation of Metahuman avatars brings a human-like and expressive layer to the communication experience, enhancing engagement and cultural inclusivity.

The dual-purpose nature of the system addresses both immediate communication needs and contributes to the broader goal of fostering a more understanding and empathetic society. By seamlessly integrating educational components, the project becomes a valuable resource for individuals looking to learn ISL, thereby promoting linguistic inclusiveness.

However, as with any pioneering project, challenges in accuracy of gesture recognition, global adaptation to diverse sign languages, and user interface design must be diligently addressed. The commitment to privacy, security, and

and accessibility remains paramount to ensure user trust and widespread adoption. Overcoming these challenges will be pivotal in realizing the full potential of the system on a global scale.

In essence, the proposed project not only signifies a technological breakthrough in communication accessibility but also embodies a commitment to educational inclusivity and cultural sensitivity. As we move forward, continual refinement, collaboration with the deaf and mute community, and an unwavering dedication to user experience will be essential in creating a lasting and impactful solution that transforms the way we communicate and learn.

IX. REFERENCES

[1]. Indian Sign Language text generation from English/Hindi Text, Pawan Kumar, Savita Khatri et al. International Journal of Recent Research Aspects (IJRRA) ISSN: 2349- 7688, Special Issue: Conscientious and Unimpeachable Technologies 2016, pp. 30-33

[2]. National Sample Survey Organization. Disabled persons in India. NSS 58th round (July December 2002) Report no.485 (58/ 26/ 1). New Delhi: National Sample Survey Organization, Ministry of Statistics and Program Implementation, Government of India, 2003.

[3]. M. Suresh Anand, A. Kumaresan, Dr.N. Mohan Kumar, An Integrated Two Way ISL (Indian Sign Language) Translation System a New Approach, Volume 4, No. 2, Jan-Feb 2013 International Journal of Advanced Research in Computer Science.

[4]. Hee-Deok Yang, Stan Sclaroff, and Seong-Whan Lee, Sign Language Spotting with a Threshold Model Based on Conditional Random Fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 7, pp.1264- 1277, July 2009.

[5]. Archana S. Ghotkar, Rucha Khatal, Sanjana Khupase, Surbhi Asati Mithila Hadap, Hand Gesture Recognition for Indian Sign Language,2012 International Conference on Computer Communication and Informatics (ICCCI - 2012), Jan. 10 12, 2012, Coimbatore, INDIA.

[6]. Yanan Xu, Yunhai Dai, Review of Hand Gesture Recognition Study and Application, Contemporary Engineering Sciences, Vol. 10, 2017, no. 8, 375 384.

[7]. Ananya Choudhury, Anjan Kumar Talukdar, Kandarpa Kumar Sarma, A Conditional Random Field Based Indian Sign Language Recognition System under Complex Background,2014 Fourth International Conference on Communication Systems and Network Technologies.

[8]. Gajanan K. Kharate, Archana S. Ghotkar, Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction.

[9]. Dr.C. Chandrasekar, G.T. Shrivakshan, A Comparison of various EdgeDetection Techniques used in Image Processing, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.

[10]. Samta Gupta, Susmita Ghosh Mazumdar, Sobel Edge Detection Algorithm, International Journal of Computer Science and Management Research.

[11]. Lei Yang, Xiaoyu Wu, Dewei Zhao, Jun Zhai, Hui Li, An Improved Prewitt Algorithm for edge detection based on noised image, 2011 4th International Congress on Image and Signal Processing.