Vivekanand Education Society's Institute of Technology

Department of Computer Engineering

Group No.: 49

Date: 10th August 2024

Project Synopsis (2024 - 25) - Sem VI

Silent Cue: Sign Language Recognition for Deaf and Non Verbal

Mentor : Ms. Manisha Mathur
Assistant Professor, CMPN

Jiya Gangwani
VESIT
2021.jiya.gangwani@ves.ac.in

Nikhil Dhanwnai
VESIT
d2021.nikhil.dhanwani@ves.ac.in

Soham Panjabi
VESIT
2021.soham.panjabi@ves.ac.in

Chirag Santwani
VESIT
2021.chirag.santwani@ves.ac.in

Abstract

This project aims to develop an advanced system for recognizing and interpreting hand gestures, specifically designed to support communication for deaf and non-verbal individuals. The proposed system leverages a robust combination of Python programming, OpenCV for real-time image processing, and SIFT (Scale-Invariant Feature Transform) descriptors to capture and analyze hand gestures. By utilizing the Bag of Words model, the system creates a comprehensive vocabulary of gesture features, which are then classified using a Support Vector Machine (SVM) model trained with Scikit-Learn.

The core of the system involves capturing hand gestures through a camera interface, extracting key features with SIFT to ensure robustness against variations in scale and rotation, and representing these features in a quantifiable manner using the Bag of Words approach. The SVM model then classifies the gestures into predefined categories, enabling accurate and efficient recognition.

Additionally, the system integrates a user-friendly interface to display real-time gesture recognition results and provide contextual feedback. This integration aims to enhance the communication experience for users by translating hand gestures into actionable text or commands, thus bridging gaps in interaction and accessibility. The project leverages recent advancements in computer vision and machine learning to offer a practical tool for improving communication for deaf and non-verbal individuals.

Introduction

Effective communication is a fundamental aspect of human interaction, yet individuals who are deaf or non-verbal often face significant barriers in expressing themselves and engaging with others. To address these challenges, there is a growing need for advanced technological solutions that can facilitate seamless communication. This project focuses on developing a sophisticated hand gesture recognition system tailored to support and enhance communication for deaf and non-verbal individuals.

The proposed system integrates cutting-edge technologies including Python programming, OpenCV for real-time image processing, and SIFT (Scale-Invariant Feature Transform) descriptors. These technologies work together to enable precise recognition and interpretation of hand gestures. By leveraging the Bag of Words model, the system builds an extensive vocabulary of gesture features, which is crucial for accurate gesture classification.

Recent advancements in machine learning, particularly in the fields of computer vision and pattern recognition, have significantly enhanced the capabilities of gesture recognition systems. Innovations such as deep learning techniques and advanced neural network architectures have provided more accurate and efficient methods for feature extraction and classification. In this project, the integration of SIFT descriptors with the Bag of Words model and the use of Support Vector Machines (SVMs) trained with Scikit-Learn harness these advancements to improve the robustness and precision of gesture recognition.

Problem Statement

In today's world, effective communication is crucial for bridging gaps between individuals who are deaf or non-verbal and those who rely on verbal interactions. Despite advancements in technology, current methods for translating hand gestures into meaningful communication often fall short in accuracy and efficiency. Traditional systems for hand gesture recognition can be inconsistent, struggling with variations in gesture execution, scale, and orientation. This inconsistency creates barriers to clear and reliable communication, impacting the user experience for individuals who depend on these systems.

Moreover, many existing gesture recognition technologies lack real-time processing capabilities, making them less effective in dynamic communication scenarios. The challenge of accurately capturing and interpreting hand gestures in real-time can result in delays and miscommunications, further exacerbating the difficulties faced by deaf and non-verbal individuals.

Proposed Solution

To address the challenges in hand gesture recognition and enhance communication for deaf and non-verbal individuals, we propose the development of an advanced gesture recognition tool. This tool will leverage cutting-edge technologies to improve the accuracy, efficiency, and real-time processing of hand gestures. Key features of the proposed solution include:

- **Capture Gesture Data:**
    - Letter Detection: Identify and classify individual hand gestures as specific letters, enabling the recognition of alphabetic input.
    - Word Formation: Use detected letters to form complete words, facilitating more complex and meaningful communication.
    - Screen Axis Alignment: Adjust gestures based on the camera's perspective to account for variations in hand positioning and orientation.

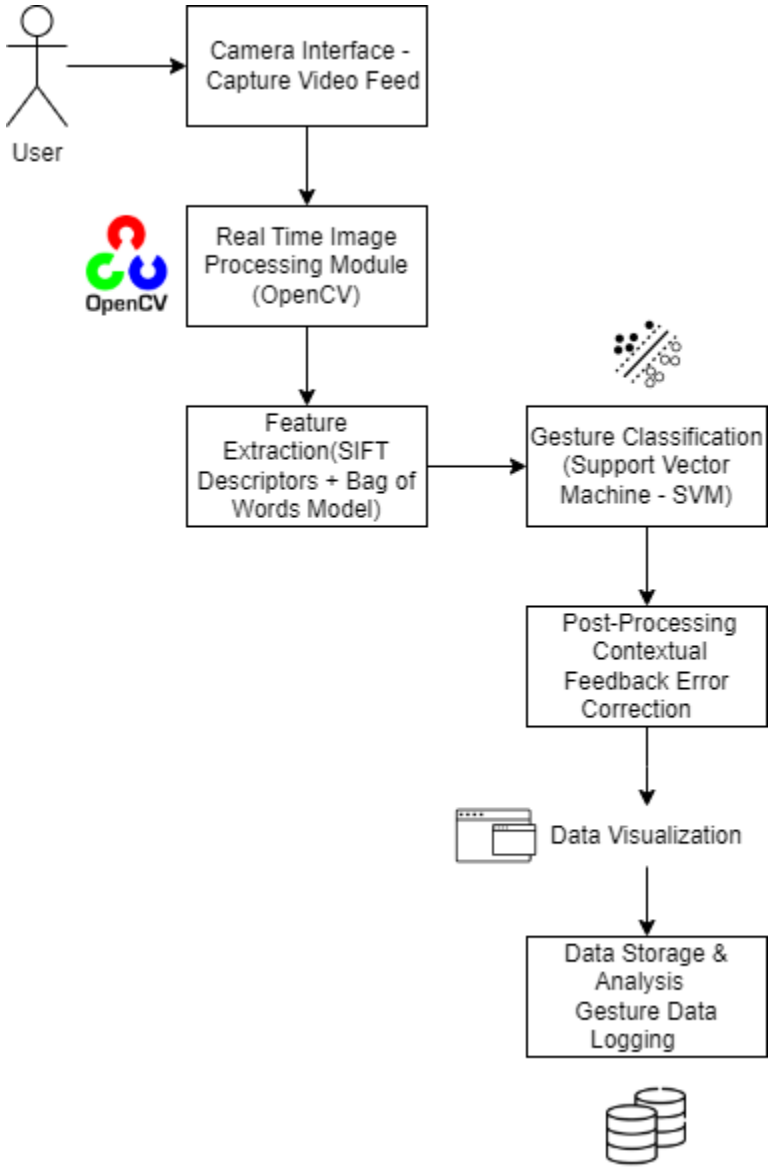
- **Enhance Communication:**
    - Sentence Formation: Build on word detection to enable the formation of complete sentences from sequences of gestures, allowing for more fluid and coherent communication.
    - Suggestions: Provide real-time suggestions based on detected gestures to help users refine their input and correct errors, improving the overall communication experience.

- **Support Diverse Needs:**
    - Sign Language Support: Include recognition for multiple sign languages, such as Indian Sign Language (ISL), American Sign Language (ASL), and British Sign Language (BSL), to cater to a broad range of users.
    - Multilingual Capabilities: Translate recognized gestures into text or speech in various spoken languages to enhance communication across different linguistic backgrounds.

- **Ensure High Performance:**
    - Accuracy: Implement advanced technologies to ensure high accuracy in gesture recognition, reducing errors and inconsistencies.
    - Fixed Gestures: Recognize and reliably interpret predefined gestures for common interactions, such as "hello" and "goodbye," to streamline communication for frequently used expressions.

By integrating these features, the proposed tool aims to deliver a comprehensive and effective solution for hand gesture recognition. The system's capabilities in letter detection, word and sentence formation, and multilingual support will significantly enhance communication for deaf and non-verbal individuals, making it a valuable tool for facilitating meaningful interactions.

Methodology / Block Diagram



Hardware, Software and Tools Requirements

Hardware Requirements:

- Computer with a modern processor (Intel i5/Ryzen 5 or higher recommended), at least 8 GB of RAM and sufficient storage space, preferably an SSD.
- NVIDIA GPU with CUDA support is recommended for computation(recommended)
- High-definition camera (720p or higher) for capturing hand gestures in real time.

Software Requirements:

- Windows 10 or later, macOS Mojave or later, or Linux (Ubuntu 18.04 or later).
- Python 3.8 or later.
- Computer Vision and Machine Learning:
- OpenCV for real-time image processing.

- Scikit-Learn for implementing machine learning models (SVM).
- Deep learning framework (TensorFlow or PyTorch) for advanced model training and evaluation.
- SIFT (Scale-Invariant Feature Transform) for feature extraction.
- Bag of Words model for feature representation.
- Flask for web-based interfaces or React Native for mobile applications.
- Any modern IDE like PyCharm, VS Code, or Jupyter Notebook for Python development.
- pyttsx3 or gTTS for converting recognized gestures into spoken language.
- Matplotlib, Seaborn, or Plotly for creating visualizations of gesture recognition data and performance metrics.
- Libraries or APIs for translation to support multilingual capabilities.

Proposed Evaluation Measures

- Recognition Accuracy: Evaluate how accurately the system identifies and classifies hand gestures, focusing on error rates and overall precision in gesture detection and interpretation.
- Word and Sentence Formation Quality: Determine how effectively the system constructs words and sentences from detected gestures, ensuring they are coherent, contextually accurate, and grammatically correct.
- User Satisfaction: Collect detailed feedback from users regarding the tool's ease of use, effectiveness, and overall satisfaction through surveys and usability tests. This will help gauge the user experience and identify areas for improvement.
- Real-Time Processing Speed: Measure the time required for the system to process and recognize gestures in real-time, ensuring it operates with minimal latency and provides timely feedback.
- Processing Efficiency: Assess the overall efficiency of the system in terms of computational resources and time required for feature extraction, classification, and gesture-to-text conversion. This includes evaluating the system's performance in various conditions and with different types of gestures.
- Multilingual Support: Evaluate the tool's performance in recognizing and translating gestures across multiple sign languages (e.g., ISL, ASL, BSL) and spoken languages. This includes checking for accuracy, language coverage, and effective translation of gestures into text or speech.

Conclusion

The proposed hand gesture recognition system leverages advanced technologies such as OpenCV, SIFT descriptors, and machine learning models to deliver a highly accurate and efficient solution for enhancing communication for deaf and non-verbal individuals. By integrating features such as letter detection, word and sentence formation, and real-time suggestions, the system significantly improves the clarity and fluidity of communication.

The inclusion of screen axis alignment ensures consistent gesture recognition, regardless of variations in hand positioning. Multilingual support and recognition of various sign languages, including ISL, ASL, and BSL, make the system versatile and accessible to a diverse user base.

References

[1] V. R. P. K. T. Suba and M. S. Patil, "Hand Gesture Recognition Using Convolutional Neural Networks," Proceedings of the IEEE International Conference on Image Processing (ICIP), 2017.

[2] J. G. C. Garcia, R. Hernandez, and D. F. Munoz, "Hand Gesture Recognition for Deaf People: Interfacing with Real-Time Applications," International Journal of Computer Applications, vol. 47, no. 2, pp. 1-6, 2012.

[3] R. A. Alomari, M. A. H. Alzubaidi, and S. M. Khan, "Hand Gesture Vocalizer for Dumb and Deaf People," Journal of Computing and Information Technology, vol. 28, no. 1, pp. 55-70, 2020.

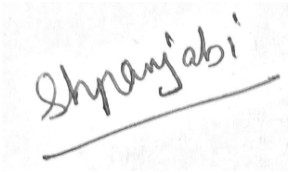
[4] G. Yang, Z. Song, and Z. Liu, "Hand Gesture Recognition Using Deep Learning with Multimodal Data," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

Signatures

Mentor

Mrs. Manisha Mathur
(Assistant Professor, CMPN)

Students



Chirag Santwani Jiya Gangwani Nikhil Dhanwani Soham Panjabi