Air canvas: drawing in air using hand gestures

Written by: ASHISH PUNDIR

Abstract

Air writing is a developing field in the areas of image processing and pattern recognition. The latest technology, in its core, supports automation and develops human-computer interface applications. Some literature reviewed related research studies to present methods that would minimize processing time while maximizing the accuracy of recognition.

Object tracking is one of the important areas in computer vision. With video recording and hardware moving along rapidly, video analysis techniques became popularly adopted because automated analysis was possible. The traditional three stages of video analysis have been: object detection, object tracking between frames, and behavior analysis. The important problems in object tracking include the choice of relevant object representation, selection of features for detection, and management of the process of object tracking. Object tracking algorithms have been proved to be very crucial in many surveillance applications, video indexing, and vehicle navigation.

This project fills one of the gaps in current technology with a motion-to-text converter for air writing, that might be employed via smart wearable devices. In this system, computer vision tracks finger movements in the air and translates this into text. Such a generated text can be used, for instance, in any one of the messaging, email types, and it is actually quite an excellent communication tool especially for the people who suffer from hearing impairment. The system cuts down reliance on mobile devices and laptops too since no form of physical text input is required.

Keywords

Air Writing, Character Recognition, Object Detection, Real-Time Gesture Control, Smart Wearables, Computer Vision.

Introduction

The old practices are slowly being replaced by the digital age, which embraces new habits that provide very different ways of putting your thoughts and speech across. Air Writing is one of those modes of digital art whereby people can write in the air, thus eliminating the use of paper or actual physical instruments to practice the craft of writing. This technology uses the best that HCI developments can come up to provide for easy, natural, and fluid interaction between a human being and digital systems. For instance, air writing may help deliver a completely paperless form of writing, drawing, or even controlling of devices by tracking real-time movements of fingers and hands.

An air writing system is dependent on the concept of hand gesture recognition, thus the development of such an application can be best done with the help of machine learning algorithms. This project will serve to advance on the basis of an efficient hand gesture recognition system using the Python programming language in combination with OpenCV and NumPy libraries. Therefore, the focus will be on developing a model that could potentially capture and interpret hand movements in air and can further translate this into some form of text or command. This

development may revolutionize communication for deaf users, but how sheer this system is possible to be, can be already foreseen: a new technique of instant text production through writing in the air to produce messages, emails, and other such forms of written communication. Air writing would also minimize dependence on the traditional writing tools of pens and paper and keyboards for streamlined interactions with digital devices.

Literature Review

Air writing, as an area of research, has an enormous application potential in the contextualfields of HCI and digital communication. Thus, users can begin to write in the air through gesture recognition technologies without needing any writing instrument. Improving accuracy in gesture recognition and reducing processing time were some of the main research focuses in this area. As it is described in the paper by [Author et al., Year], it develops machine learning algorithms using deep techniques of CNNs for acquiring real-time recognition to translate hand-gesture to text. It does provide better efficiencies to the system for air writing to make this more practical and applicable for the real world. Object tracking and motion analysis go on to enhance gesture recognition systems in further ways by enabling more fluid and reliable hand tracking. Potential applications in this area include communication with deaf or hard-of-hearing people, hands-free device controls, or virtual reality environments. Further perfecting this technology, the research shall proceed towards a smooth, intuitive human-computer interaction—and here, libraries such as OpenCV and NumPy with Python, etc.

Problems Encountered

The development of the air writing system faced several challenges during its implementation. Some of the key issues that were encountered include:

1. **Detection of the Finger:**

   One of the primary challenges was accurately detecting the position of the finger in a standard RGB image. Since the system did not utilize any depth-sensing sensors (such as those used in depth-sensing cameras), detecting the fingertip in 3D space became inherently difficult. The absence of depth perception made it challenging to differentiate between fingers and background clutter, leading to interference and false detections. This issue became especially problematic when the background had similar colors or textures to the hand, or when the user's hand was in motion against a non-contrasting background.

2. **Tracking Pen Movements:**

   The air writing system was designed to capture the movement of the pen in the air; however, one major limitation of the system was its inability to determine whether the user's finger was in the "pen-up" or "pen-down" position. With only a single camera tracking the hand gestures, the system could not differentiate between a hovering hand and contact with an invisible surface. This limitation meant that the system could not track the drawing or writing process smoothly, resulting in mistakes while trying to write text linearly. Additionally, when the user attempted to change modes or commands using gestures, the

system sometimes failed to accurately recognize the intended command due to interaction ambiguity and interference.

3. **Real-time Control:**

   Real-time gesture recognition is a particularly challenging task, especially in dynamic environments where hand movements can be unpredictable. Achieving a highly precise and responsive system that operates with minimal latency and at a high frame rate is crucial for smooth user interaction. However, optimizing the code for such high precision and responsiveness was difficult, especially given that the system was primarily built on consumer-grade hardware (i.e., a regular webcam). Consequently, achieving optimal performance—where the lag between gesture recognition and text generation remained minimal—was challenging. Additionally, the system required the user to memorize and use a predefined set of gestures to interact with the system, which created a steeper learning curve. This reliance on gestures made the system less user-friendly, particularly for non-tech-savvy individuals.

Problem Definition

The air writing system is designed to address multiple societal challenges, particularly in the areas of communication, sustainability, and safety. Some of the key problems the system aims to solve include:

1. **Communication for the Hearing Impaired:**

   One of the primary applications of air writing is in facilitating communication for people with hearing impairments. For the deaf and hard-of-hearing community, sign language is a commonly used form of communication. However, sign language can be challenging for those who are unfamiliar with it, and there is a shortage of trained interpreters. Air writing provides a simple and direct means of communication that can be easily interpreted by anyone, regardless of their familiarity with sign language. The system allows deaf users to communicate with others without the need for a physical writing instrument, enhancing accessibility and reducing the dependency on sign language interpreters.

2. **Moving Away from Mobile Device Addiction:**

   Mobile device addiction has become a growing concern in modern society. People are increasingly dependent on smartphones for various aspects of daily life, which has been linked to issues such as excessive screen time, road accidents, and mental health problems. Air writing offers an alternative to traditional mobile devices by enabling users to type text using hand gestures in the air. By reducing reliance on smartphones and traditional input devices such as keyboards, the system minimizes the time spent in front of screens and promotes healthier, safer communication methods.

3. **Paper Use Preservation:**

   The environmental impact of paper production is significant, with deforestation, energy consumption, and waste generation being key concerns. Air writing eliminates the need for

physical tools such as paper and pens, offering a paperless method of writing and communication. By reducing the use of paper, the system helps mitigate the negative environmental effects associated with paper production. Additionally, it reduces the generation of waste and the energy consumption associated with the disposal and management of electronic devices that rely on paper documents.

System Methodology

The air writing system is fundamentally based on fingertip detection. It tracks hand movements and converts them into text. The methodology follows these key steps:

1. **Dataset Creation for Fingertip Recognition:**

   The development of the fingertip detection model began with creating a training dataset. Videos of hand movements performing various gestures were captured to form a comprehensive dataset. These videos were segmented into frames, from which 2000 images were selected for analysis. The frames were manually labeled with fingertip positions to ensure accurate training of the model.

   Dataset Details:
   The dataset aimed to capture different hand shapes, angles, and lighting conditions to ensure that the model could generalize well across various profiles. By including a diverse range of hand poses and environmental factors, the system was designed to minimize errors in fingertip detection.

2. **Training the Fingertip Recognition Model:**

   After annotating the dataset, it was split into a training set and a validation set for model evaluation. Deep learning techniques were applied to train the recognition model using algorithms such as Single Shot Detector (SSD) and Faster R-CNN (Region-based Convolutional Neural Networks). These models were chosen due to their proven performance in object detection tasks. Upon evaluation, Faster R-CNN outperformed SSD in terms of both speed and accuracy, providing faster inference times and more reliable fingertip detection.

3. **Gesture Classification and Text Generation:**

   Once the fingertip positions were detected, the system classified the hand gesture into predefined commands or letters. By tracking the trajectory of the fingertip in real-time, the system recognized specific gestures associated with writing or deleting text. For instance, a circular motion could correspond to the letter "O," while a leftward swipe could trigger the backspace function.

System Algorithm

The air writing system offers several key features that enhance user interaction through intuitive hand gestures:

1. **Writing Mode:**
   The system tracks the movement of the fingertip and translates this into text. By following the trajectory of the fingertip, the system recognizes specific hand gestures and maps them to corresponding characters, which are displayed on the screen.
2. **Color Mode:**
   Users can change the color of the text by making specific hand gestures. This feature allows for personalization, enabling users to switch between different text colors during the writing process.
3. **Backspace Mode:**
   A backspace gesture allows users to delete the last character that was written. This feature offers an intuitive way to correct mistakes while writing in the air.

Conclusion

This air writing system represents a novel and efficient alternative to traditional writing tools. It offers potential applications for individuals with hearing impairments, helping bridge communication gaps. Additionally, its paperless nature contributes to environmental sustainability by reducing paper consumption. The system also provides a safer and more efficient way to interact with digital devices, offering a new approach to human-computer interaction (HCI).

While the system is promising, it is still in its prototype stage, and there are areas for improvement. For instance, replacing the current character recognition model with more advanced handwriting recognition techniques could enable faster text input. Moreover, integrating more complex object detection algorithms, such as YOLO v3, could further improve the accuracy and reliability of fingertip recognition, enhancing the overall performance of the system. As the technology evolves, air writing has the potential to become a seamless part of everyday communication, reducing reliance on traditional writing tools and contributing to the advancement of HCI.