

Reconhecimento de Gestos para Controle de Volume de Áudio com Inteligência Artificial

Projeto Final - Robótica Inclusiva - 2024-2



Equipe:
Lucas Gurgel
Maria Clara Soares
Pedro Filipe

Resumo

Este projeto desenvolve um sistema de reconhecimento de gestos para controlar o volume de áudio do sistema operacional. Utilizando a biblioteca MediaPipe para detecção das mãos e integração com a API PyCaw para manipulação de áudio, o sistema permite controlar o volume do som apenas com a quantidade de dedos levantados, em tempo real. O código é projetado para oferecer uma interação acessível e intuitiva, principalmente para usuários com dificuldades de mobilidade.

Palavras-chave: Reconhecimento de Gestos

1 Introdução

Este trabalho visa desenvolver uma solução interativa para controle de volume utilizando gestos manuais, uma tecnologia que pode ser aplicada em cenários de acessibilidade, proporcionando maior autonomia para usuários com deficiência motora. A integração entre a visão computacional, por meio do MediaPipe, e o controle de áudio via PyCaw oferece uma interface simples e intuitiva, utilizando apenas a câmera.

2 Objetivos

O objetivo principal deste projeto é criar um sistema funcional que permite ao usuário ajustar o volume de áudio de seu dispositivo com base na quantidade de dedos levantados em um gesto manual. Para isso, serão utilizados:

- Detecção de mãos em tempo real utilizando a biblioteca MediaPipe.
- Controle de volume de áudio via a API PyCaw.
- Interface simples e intuitiva de interação com a máquina.

3 Metodologia

O sistema foi desenvolvido utilizando três componentes principais:

- **Captura de Vídeo:** O OpenCV é utilizado para capturar imagens da câmera em tempo real.
- **Detecção de Mãos:** A biblioteca MediaPipe é responsável por identificar as mãos e seus pontos de referência no vídeo.
- **Controle de Áudio:** A biblioteca PyCaw é utilizada para ajustar o volume do sistema de acordo com o número de dedos levantados.

A seguir, são descritas as etapas do processo de detecção e controle:

1. O vídeo da câmera é processado para detectar os pontos de referência das mãos usando o MediaPipe.
2. O número de dedos levantados é identificado com base nas posições desses pontos de referência.
3. O volume do sistema é ajustado proporcionalmente ao número de dedos levantados: 0 dedos para mudo, 1 dedo para 20.
4. O controle do volume é feito utilizando a API PyCaw, que permite alterar o volume do sistema.

3.1 Algoritmo de Reconhecimento de Gestos

4 Resultados

O sistema foi testado com diferentes condições de iluminação e posicionamento da câmera. Durante os testes, o reconhecimento de gestos foi eficiente, e o volume foi ajustado corretamente com base na quantidade de dedos levantados. O código demonstrou estabilidade e baixo tempo de resposta, operando em tempo real.

Algorithm 1 Processamento de Gestos para Controle de Volume

Require: Imagem capturada pela câmera

Ensure: Volume ajustado com base no número de dedos levantados

```
1: Processar imagem com MediaPipe para detectar as mãos
2: if mãos detectadas then
3:   Contar o número de dedos levantados
4:   Ajustar volume de acordo com o número de dedos:
5:   if 0 dedos then
6:     volume = 0 %
7:   end if
8:   if 1 dedo then
9:     volume = 20 %
10:  end if
11:  if 2 dedos then
12:    volume = 40 %
13:  end if
14:  if 3 dedos then
15:    volume = 60 %
16:  end if
17:  if 4 dedos then
18:    volume = 80 %
19:  end if
20:  if 5 dedos then
21:    volume = 100 %
22:  end if
23: end if
```

5 Conclusões

Inicialmente, nossa proposta de projeto final envolvia o desenvolvimento de um sistema de reconhecimento de gestos em LIBRAS, com o objetivo de criar uma ferramenta de comunicação acessível para a comunidade surda. No entanto, ao longo de várias semanas de trabalho intenso, enfrentamos desafios técnicos significativos, especialmente com a biblioteca TensorFlow, que dificultaram o progresso da ideia original.

Diante disso, decidimos reformular o projeto e buscar uma solução alternativa, mas igualmente relevante e inovadora. Optamos por criar um sistema de reconhecimento de gestos para controle de volume de áudio, utilizando as bibliotecas MediaPipe para a detecção de mãos e PyCaw para o ajuste do volume do sistema. Essa nova abordagem, embora diferente da proposta inicial, manteve nosso compromisso com a acessibilidade e inclusão, oferecendo uma solução prática e intuitiva para pessoas com mobilidade reduzida, que podem agora controlar o volume do computador de forma simples e natural, utilizando apenas gestos com as mãos.

6 Trabalhos Futuros

Trabalhos futuros podem incluir:

- Aumentar a precisão do sistema com o uso de aprendizado de máquina para reconhecer mais gestos e realizar controles mais complexos.
- Aplicação do sistema em outros dispositivos, como sistemas de automação residencial ou robôs assistivos.
- Inclusão de feedback visual ou auditivo para melhorar a experiência do usuário.