[SIGHTSTEP]

Computer-Vision Karaoke Web Application


Stephanie Mullins


Robert Kelly


Athulya Ganesh

Computer Science | Class of 2024 | The Vision Vanguards


Dr. Jillian Aurisano

UNIVERSITY OF Cincinnati

Background

Karaoke is a massively popular activity among people, young and old alike. Similarly, dancing is a universal language among many. Currently, there are no mainstream free-of-cost products that adequately merge the two in a fun and interactive way.

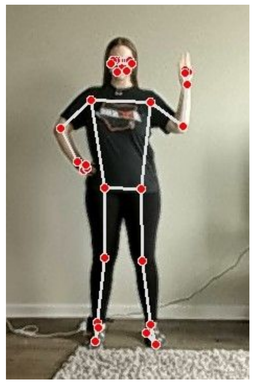
Objective

Our project, initially focused on the song "Single Ladies," merges the thrill of karaoke with dance and posing for an immersive interactive experience. Users sing along, execute generated poses, and are scored, allowing them to see where they stand on a leaderboard against their peers. Designed as a proof of concept, this initiative aims to broaden its scope with future enhancements.

Accomplishments

- Established a scalable web application framework
- Enhanced audio processing for improved sound interaction
- Integrated advanced pose estimation for accurate movement tracking
- Merged pose estimation with audio processing for seamless user experience
- Resolved numerous bugs for system stability and performance
- Addressed a high volume of software issues, ensuring reliability under heavy use
- Streamlined user interface design for intuitive navigation

Pose Modeling



- Used MediaPipe along with OpenCV
- Used pose landmarks to identify joints/ connecting lines
- Found angle between three points to hardcode poses and determine if they are accomplished

Audio Processing

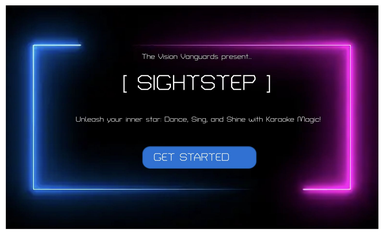
- Utilized SpeechRecognition to transcribe audio files to text for both user audio and song lyrics
- Implemented a lyrics comparator to compare sung lyrics to actual lyrics
- Used Librosa to extract and compare user pitches to song pitches, accounting for different tones and octaves
- Implemented threading to be able to play audio while recording user audio synchronously

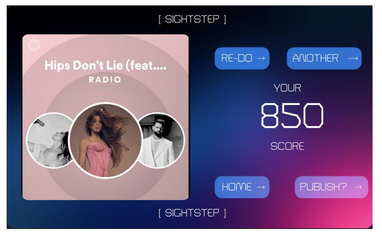
Technologies Used

OpenCV, python, pyAudio Analysis, MediaPipe

Web Interface

- **Landing Page**: An engaging welcome to the app
- **Authentication**: Straightforward process powered by Firebase
- **Dashboard**: Offers a leaderboard and navigation bar containing settings, and option to start a performance
- **Game Interface**: Shows lyrics, and the camera if the user is following gestures
- **Progress Review**: Tracks users performance and progress post-performance





Challenges

- Difficult to implement adequate pose models for specific songs
- Problems with Spotify API led to refactoring of audio retrieval

Future Work

- Enhance the game interface to improve user interaction and experience
- Introduce a wider variety of character poses to increase game realism
- Incorporate a more extensive collection of fully developed songs to enhance gameplay