



Pitch Piano

Computer Audition and Generative Music

Sistemas Digitais Interativos

Mestrado em Multimédia

Joana Cerqueira Maia 202108835

Contexto do trabalho

Nos dias de hoje, a integração de tecnologia e música tem gerado novas formas de interação e expressão, especialmente com o uso de dispositivos interativos que possibilitam uma experiência musical intuitiva e imersiva. A ideia de experimentar e tocar música de forma simples e acessível desperta o interesse de pessoas que desejam aprender ou se conectar com a música sem a necessidade de instrumentos físicos complexos. Com essa ideia em mente, desenvolvi um sistema interativo que permite aos utilizadores tocar e explorar notas musicais usando apenas a voz.

Este projeto, inspirado pelo conceito de interação contínua em loop, utiliza *Computer Audition* e *Generative Music* para criar uma experiência fluida e envolvente. *Computer Audition* é um campo da inteligência artificial e do processamento de sinal digital que lida com a capacidade dos computadores e algoritmos ouvirem e interpretarem o som. *Generative Music* refere-se à criação de música através de algoritmos que produzem sons ou composições musicais de forma autónoma ou semi-autónoma.

Através de reconhecimento de voz, o sistema identifica as notas musicais que o utilizador deseja tocar ao simplesmente ouvir o comando falado e as reproduz automaticamente, como a sequência “Dó, Ré, Mi” ou apenas uma nota, “Sol”. Destacando também as teclas correspondentes num piano virtual. Isso permite que qualquer pessoa, independente do conhecimento musical prévio, experimente a criação de sons e até de pequenas melodias, criando um ambiente lúdico e de aprendizagem.

Concluindo, este projeto une esses conceitos de inteligência artificial e visualização interativa, representando uma interseção entre tecnologia e música e oferecendo uma experiência imersiva para explorar a música de forma intuitiva. Dessa forma, o utilizador pode criar sons e melodias no conforto da sua casa, explorando o universo musical de maneira acessível e estimulante.

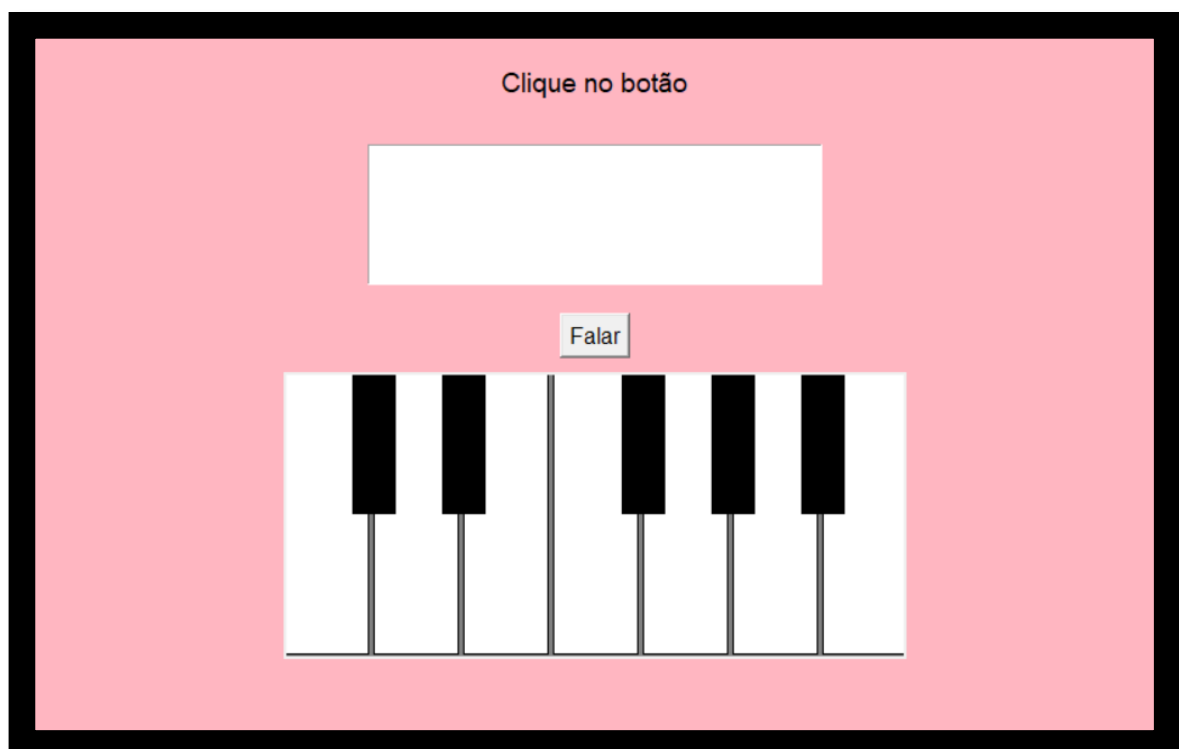


Figura 1 - Pitch Piano Project

Framework Loop Interativo

Bongers

O framework loop interativo de Bongers é um modelo que descreve como as experiências interativas podem ser organizadas e compreendidas. Esse ciclo enfatiza a interação entre o utilizador e a interface e é composto por dois processos: Controlo e Feedback.

No meu projeto, o framework pode ser observado e aplicado na sequência de interação entre o utilizador e o sistema e o seu loop pode ser observado nos seguintes parâmetros:

- **Sensores:** O microfone captura o input de voz do utilizador de modo a que o sistema reconheça as notas que este deseja tocar.
- **Processamento:** O comando de voz captado pelo microfone é processado para ser traduzido numa nota musical específica.
- **Feedback:** O sistema reproduz o som da nota musical que o utilizador escolheu e a sua tecla correspondente é destacada no piano virtual.
- **Atuadores:** As funcionalidades do sistema produzem o som e destacam visualmente as teclas do piano.

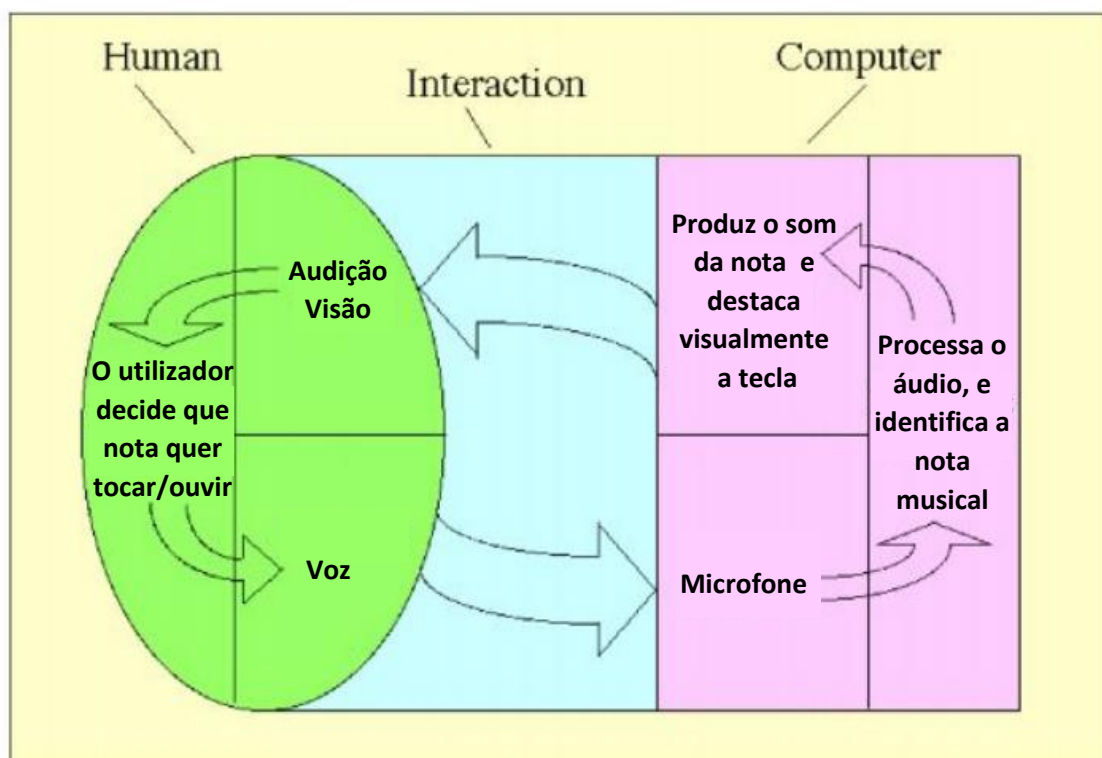


Figura 2 - Framework loop interativo de bongers

Artistas e Trabalho relacionado

Perfect Pitch Challenge 2, Tik Tok

O filtro Perfect Pitch Challenge 2 é a última tendência a tornar-se viral no TikTok. Trata-se de controlar uma bola que deve subir através de buracos nas paredes marcados com as notas musicais: "do", "re", "mi", "fa", "so", "la", "ti" e "do". No entanto, a diferença é que a voz, ou melhor, o tom determina se a bola sobe ou desce.

Tanto o meu projeto, Pitch Piano, como o projeto Perfect Pitch Challenge 2 utilizam o reconhecimento de áudio para identificar notas musicais e darem feedback instantâneo de modo a manter o utilizador interessado. Ambos permitem a interação musical acessível através da voz, e oferecem uma experiência educativa e divertida. No entanto, o meu projeto não se foca apenas na aprendizagem do tom correto de cada nota mas também em mostrar a nota visualmente.

Este filtro do TikTok foi uma fonte de inspiração porque mostrou que a tecnologia pode tornar a aprendizagem musical mais acessível e interativa. Com este aprendi que reconhecimento de voz pode ser uma ferramenta poderosa para interações musicais, mesmo para pessoas sem experiência prévia.



Figura 3 - Perfect Pitch Challenge 2



Figura 4 - Very Nervous System

Very Nervous System, David Rokeby

O projeto Very Nervous System é a terceira geração de instalações sonoras interativas criadas por David Rokeby. Neste sistema são utilizadas câmeras de vídeo, processadores de imagem, computadores, sintetizadores e sistemas de som para criar um espaço no qual os movimentos do corpo criam som e/ou música. Foi apresentado principalmente como uma instalação em galerias, mas também foi instalado em espaços públicos ao ar livre e foi utilizado em diversas performances.

Assim como no projeto Very Nervous System, o meu projeto utiliza o corpo (no caso, a voz) para criar sons interativos. Ambos os sistemas traduzem um input humano, gesto (Very Nervous System) e voz (Pitch Piano), numa resposta sonora, criando uma experiência em que o utilizador interage com o som através de ações físicas.

Este projeto serviu de inspiração no que toca a reforçar a ideia de que o utilizador pode ser um "compositor" sem precisar de habilidades técnicas prévias.

Arquitetura Técnica

Utilizando a linguagem Python e as bibliotecas Tkinter, SoundDevice e SpeechRecognition, o sistema permite que o utilizador selecione notas musicais através dos seus nomes. Estas são, então, reproduzidas em áudio e destacadas visualmente numa interface gráfica em forma de piano.

Ao iniciar o programa, o utilizador depara-se com um piano, um botão "Falar" e uma caixa branca. Este piano é constituído por teclas brancas e pretas desenhadas com o Tkinter Canvas que estão associadas às respetivas notas musicais, com coordenadas e cores distintas para facilitar a visualização.

Através de SpeechRecognition, quando o utilizador clica no botão "Falar", o sistema ativa o microfone para captar a voz e processa o áudio para reconhecer o nome da nota. O reconhecimento de voz converte, então, o comando falado na nota musical, usando o mapeamento específico de palavras/notas em português para notas em inglês. Por exemplo, a palavra "dó" é convertida para "C", e assim por diante.

Com a ajuda do NumPy, uma onda sonora sinusoidal é gerada para cada nota, baseada na frequência específica de cada uma. Uma vez gerada esta onda, o SoundDevice reproduz o som através dos altifalantes do computador, permitindo ao utilizador ouvir a nota pretendida.

Para complementar o feedback auditivo, o sistema altera temporariamente a cor da tecla da nota ativa, mudando para uma cor de destaque enquanto a nota está a ser tocada⁽¹⁾. Caso o utilizador queira reproduzir uma sequência de notas como "Dó, Ré, Mi", estas ficarão todas selecionadas mostrando a respetiva sequência. Quando este clica novamente no botão "Falar", estas notas deixam de estar destacadas e o mesmo processo se repetirá.

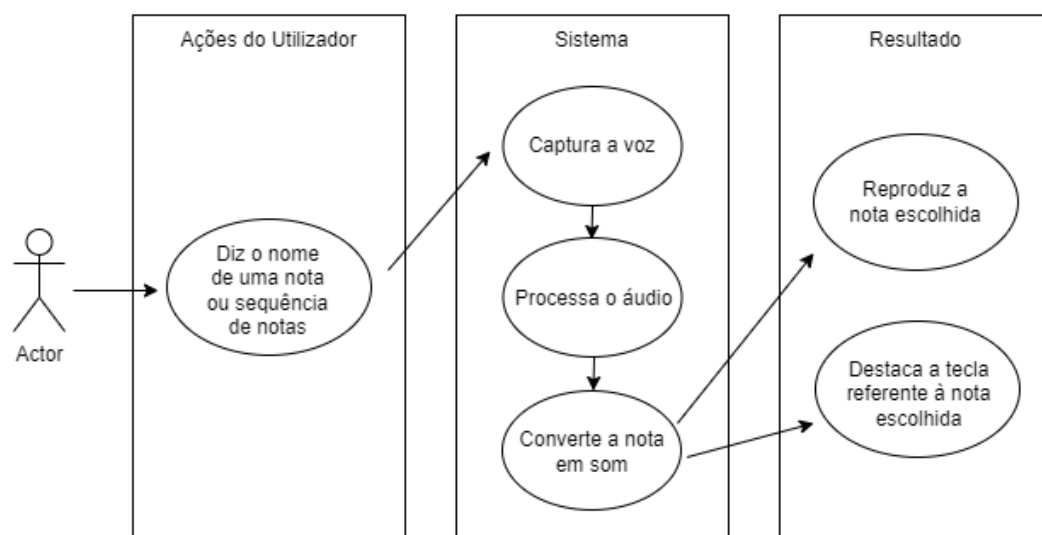


Figura 5 - Diagrama sobre o algoritmo do projeto

(1) O sistema está programado para as teclas aparecerem destacadas ao mesmo tempo que o som da nota é produzido. No entanto, visualmente as teclas só ficam destacadas quando a última nota é produzida. De modo a confirmar que o código está correto, testei no terminal através de mensagens complementares (prints) e de facto a função `destacar_tecla()` está a fazer o suposto, apesar do resultado visual não corresponder ao que está programado. Comprovo o mesmo no 'vídeo demo'.

Como correr o projeto?

1. Instalar a linguagem python: <https://www.python.org/downloads/>
2. Abrir o terminal 'Linha de Comandos'
3. Instalar as bibliotecas necessárias: `pip install numpy sounddevice speechrecognition opencv-python`
4. Abrir a pasta onde se encontra o ficheiro "PitchPiano" num IDE (recomendo VScode)
5. Correr o programa no terminal através do comando: `python PitchPiano.py`
6. Fechar o programa naturalmente no botão X

Referências

- Perfect Pitch: <https://www.dexerto.com/tiktok/how-to-get-the-viral-perfect-pitch-filter-on-tiktok-2907098/>
- Very Nervous System: <https://www.davidrokeby.com/vns.html>