



universidade de aveiro  
theoria poiesis praxis

DETI - IM'18/19

# Relatório Trabalho Prático Fusão

## SPOTIFY - Interação Por Voz e/ou Gestos

**Realizado por:**

Filipe Rocha - 67432

Pedro Silva - 72645

# Índice

Introdução .....	2
SPOTIFY .....	2
Cenário .....	2
Arquitectura .....	3
Funcionalidades .....	3
Utilização .....	4
Diálogos .....	4
Gestos .....	7
Redundância .....	8
Complementaridade .....	8
Considerações .....	8
Conclusão .....	8

## Introdução

Este trabalho prático tem como objetivo permitir a interação por voz (em português) e/ou por gestos com o spotify web player (recorrendo ao navegador Chrome).

Apresenta-se neste relatório os diferentes comandos por voz e por gestos disponíveis para o utilizador. Onde é possível interagir por comandos de voz, ou por comandos de gestos com recurso ao kinect e por comandos que requerem voz e gesto (complementares).

As principais características referentes aos comandos para este projecto escolhidas são: a sua forma intuitiva mas com garantia de autenticidade na gramática e nos gestos disponíveis, assim como o feedback do sistema auditivo e também visual apresentado ao utilizador.

## SPOTIFY

É um serviço de streaming de música e podcasts, disponibilizado em diferentes sistemas e dispositivos. Fornece conteúdo protegido de conteúdo provido de restrição de gestão de direitos digitais de estúdios musicais e empresas da área.

Para este projeto foi decidido incidir sobre o Web Player disponibilizado de forma gratuita. Mas devido a alterações recentes por parte da empresa SPOTIFY sobre as APIs disponibilizadas não foi possível utilizarmos uma API específica disponibilizada por esta empresa, devido a esta nova estar em fase Beta e requerer conta *premium* (nenhum dos elementos do grupo possui conta paga e como tal não é possível utilizar a nova e agora única funcional API deste sistema). Como forma de resolução desta questão optámos por utilizar o acesso ao Web Player com recurso a drivers do navegador Chrome como meio de permitir a interação com o respectivo sistema da SPOTIFY.

## Cenário

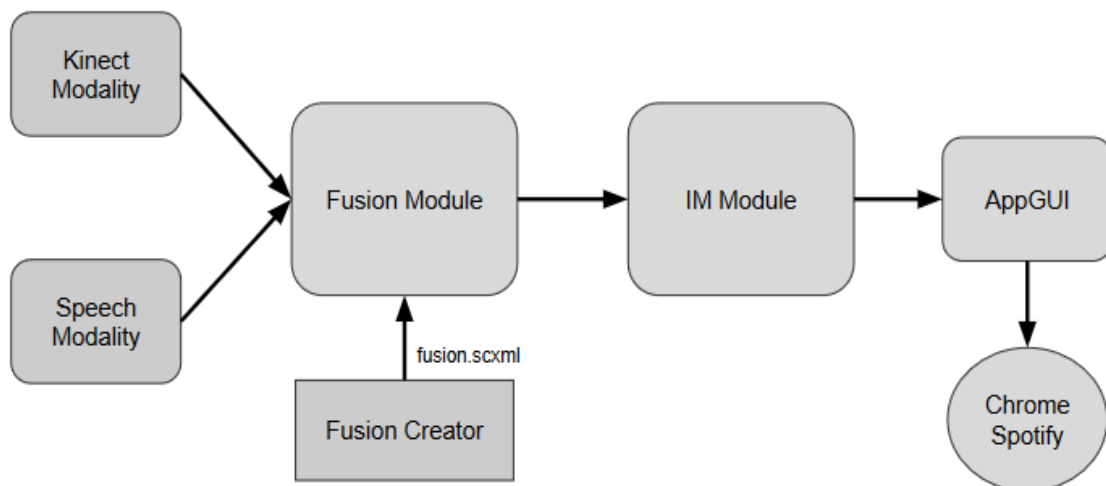
Iniciando uma sessão de utilizador no Web Player do SPOTIFY, e recorrendo ao uso do “Kinect for Windows V2”, deverá de ser possível interagir com esta aplicação através de fala natural e gestos.

Exemplo:

- O utilizador pretende ouvir a música Anzol dos Rádio Macau.
- Para tal, diz “Anzol” em voz audível para o microfone do kinect e de seguida realiza o gesto associado a “reproduzir música”.
- A aplicação irá dar feedback auditivo e realizar a ação pretendida.

## Arquitectura

- **speechModality** - Reconhecimento de comandos de Voz;
- **KinectGestureBasics** - Reconhecimento de comandos por Gestos;
- **AppGui** - Responsável por receber mensagens do IM e executar respectivas acções no Spotify Web Player;
- **Fusion Creator** - Responsável por criar o ficheiro “fusion.scxml” usado no módulo FusionEngine runtime;
- **FusionEngine runtime** - Responsável por classificar dados recebidos das aplicações de reconhecimento de Voz e Gestos como redundantes, complementares ou singulares e enviar a respectiva informação para o módulo IM;
- **IM runtime** - Responsável por receber dados vindos do módulo FusionEngine e efetuar a comunicação com a aplicação AppGui;



## Funcionalidades

- Reproduzir
- Parar
- Retroceder tempo na música
- Retroceder a música
- Avançar a música
- Repetição
- Activar modo aleatório
- Criar playlist
- Controlar volume
- Navegar entre páginas
  - Páginas pré-definidas
- Pesquisar
  - Pesquisas pré-definidas
- Tocar músicas
  - Músicas pré-definidas
- Fechar aplicação

## Utilização

Aplicações requeridas: Visual Studio 2017; Pacotes MSSpeech 64bits; Kinect for Windows V2 SDK; navegador Chrome.

Na aplicação “AppGui”, ficheiro “SpotifyWorker.cs”, onde se inicia a *web\_driver* [presente no SpotifyWorker()], deverá alterar o caminho para o *chromedriver.exe* agora no seu computador.

Na aplicação “KinectGestureBasics”, ficheiro “GestureDetecture.cs”, deverá editar a variável “string gestureDatabase” para o respectivo path no seu computador ou para o path @”Database\reproduzirSolutionGBD.gbd”.

Deverá executar na linha de comandos o módulo IM [“java -jar .\mmiframeworkV2.jar”] e o módulo Fusion [“java -jar .\FusionEngine.jar”]. De seguida deverá iniciar a aplicação “speechModality”, a aplicação “KinectGestureBasics” e a aplicação “AppGui” que irá apresentar uma janela de navegação Chrome com a página de login da SPOTIFY. Deverá então introduzir as credenciais de Utilizador e respectiva Password e efetuar assim o login.<sup>1</sup>

De seguida, interaja por voz com a fala: “Página inicial” ou “Menu principal” ou “Vai para o início”, ou com o gesto referente a ir para a “Página inicial”. E poderá agora continuar a interação por voz e/ou por gestos com base nas funcionalidades apresentadas neste projeto.

## Diálogos

- Reproduzir
  - Reproduzir spotify
  - Toca música
  - Reproduz a música
- Parar
  - Parar spotify
  - Parar música
  - Pára a música
  - Pára a reprodução
  - Mete em pausa
- Retroceder tempo na música
  - Retrocede na música spotify
  - Retrocede nesta música
- Retroceder a música
  - Voltar à música anterior
  - Quero voltar à música anterior
  - Ir para a música anterior
  - Reproduz a música anterior
- Avançar a música
  - Avançar spotify
  - Reproduz a música seguinte
  - Avança esta música

---













<sup>1</sup> Se não possuir um conta de utilizador spotify pedimos que nos envie e-mail, onde procederemos ao envio de credenciais para que possa aceder. Contactos: [filipe.esr@ua.pt](mailto:filipe.esr@ua.pt) ou [pedro.mfsilva@ua.pt](mailto:pedro.mfsilva@ua.pt).

- Repetição
  - Repetir
    - Repetir spotify
    - Repetir música
    - Repete a música
  - Não Repetir
    - Não repetir spotify
    - Não repetir mais a música
    - Desligar repetir
  - Repetir uma vez
    - Quero repetir spotify
    - Quero repetir a música
    - Quero repetir uma vez
  - Não Repetir uma vez
    - Não quero repetir spotify
    - Não quero repetir mais a música
    - Quero desligar repetir
- Activar modo aleatório
  - Baralha as músicas
  - Mete em aleatório
  - Activa o modo reprodução aleatória
  - Desactiva o modo reprodução aleatória
- Navegar entre páginas
  - Menu Principal
    - Menu principal
    - Página inicial
    - Vai para o início
  - Relevantes
    - Navegar para relevantes
  - Podcast
    - Navegar para podcast
  - Charts
    - Navegar para cartas
  - Género
    - Navegar para género
  - Novos lançamentos
    - Navegar para novo
  - Descobrir
    - Navegar para descobrir
    - Quero descobrir música
  - Biblioteca
    - Navegar para biblioteca
    - Quero ir para a biblioteca
  - Playlists
    - Mostra-me as minhas playlists
  - Feito para mim
    - Mostra-me as músicas feitas para mim

- Músicas
  - Mostra-me músicas
- Álbuns
  - Mostra-me álbuns
- Artistas
  - Mostra-me artistas
- Podcasts
  - Mostra-me podcasts
- Playlist
  - Criar
    - Criar Playlist
  - Adicionar Música
    - Adicionar música à biblioteca
    - Adicionar música à fila de espera
    - Adicionar música às minhas playlists
- Volume
  - Aumenta aí o volume.
  - Baixa o volume.
  - Quero silêncio se faz favor.
  - Quero o volume como estava.
  - Minimiza o volume.
  - Maximiza o volume.
  - Mete o volume a trinta.
  - Mete o volume a cinquenta.
  - Mete o volume a oitenta.
- Pesquisar
  - Pesquisa por Contentores
  - Pesquisa por Xutos
  - Pesquisa por In The End
  - Pesquisa por Linkin Park
  - Pesquisa por There Will Be Time
  - Pesquisa por Never Gonna Give You Up
  - Pesquisa por Bon Jovi
- Tocar músicas
  - Anzol dos Rádio Macau
    - Anzol
    - Toca a música Anzol dos Rádio Macau
    - Quero ouvir Rádio Macau Anzol
    - Mete a dar a música Anzol dos Rádio Macau
  - Sete Mares dos Sétima Legião
    - Toca a música Sete Mares dos Sétima Legião
    - Quero ouvir Sete Mares
    - Mete a dar Sete Mares
  - Canção de Engate do António Variações
    - Canção do Engate
    - Toca a Canção do Engate do António Variações
    - Quero ouvir a Canção de Engate por António Variações
    - Mete a dar a Canção de Engate por António Variações

- Cavalos de Corrida dos UHF
  - Toca Cavalos de Corrida dos UHF
  - Quero ouvir UHF com Cavalos de Corrida
  - Mete a dar UHF com Cavalos de Corrida
- Circo de Feras dos Xutos e Pontapés
  - Circo de Feras
  - Toca a música Circo de Feras dos Xutos e Pontapés
  - Quero ouvir Circo de Feras pelos Xutos e Pontapés
  - Mete a dar Circo de Feras pelos Xutos e Pontapés
- Paixão dos Heróis do Mar
  - Toca Paixão dos Heróis do Mar
  - Quero ouvir Heróis do Mar Paixão
  - Mete a dar Heróis do Mar com a música Paixão
- Fechar aplicação
  - fecha o spotify
  - sair do spotify
  - fecha a janela

## Gestos

Gestos		Funcionalidade	
Abrir os braços como um "livro"		Ir para o menu principal	
Push frontal com a mão direita aberta		Reproduzir música	
Push frontal com mão esquerda aberta		Parar música	
Swipe com braço esquerdo para a direita		Ir para a música anterior	
Swipe com braço direito para a esquerda		Ir para a música seguinte	
Braços levantados sobre a cabeça		Tirar o som	



## Redundância

Neste projecto, existem redundâncias de comandos por voz e por gestos. Todos os comandos por gestos, à excepção do comando de gesto “reproduzir”, é possível executar quer por voz quer por gesto.

## Complementaridade

Existem casos de complementaridade em que será necessário recorrer a comando por voz e a comando por gesto para executar uma determinada acção.

Estes casos referem-se à reprodução de três músicas específicas. “Anzol” dos Rádio Macau, “Circo de Feras” dos Xutos e Pontapés e “Canção do Engate” de António Variações. Estes são os casos exemplo para se utilizar complementaridade e reproduzir uma destas três músicas de acordo com o pretendido pelo utilizador.

### **Interacção:**

O utilizador deve falar o comando por voz (por exemplo) “Canção do Engate” e de seguida executar o gesto referente a “reproduzir” e só assim a aplicação irá reproduzir esta música.

## Limitações

As principais dificuldades que presenciamos no decorrer deste projecto foram:

- Timing entre acção de fala e gesto para comando complementar provou-se um desafio;
- Tratamento de recepção de gestos idênticos seguidos;
- Detecção de gestos, dependente do ambiente e utilizador para influência da taxa de sucesso.

É de referir que projectos de interacção requerem especial atenção ao ambiente e tipo de utilizadores para os quais a aplicação tem como mercado alvo pois existe uma grande quantidade de variáveis que pode influenciar o desempenho da aplicação.

## Conclusão

Os objetivos que foram inicialmente propostos foram alcançados e achamos que conseguimos desenvolver uma aplicação interativa com base em voz e gestos indo de encontro aos conteúdos leccionadas na unidade curricular.

No final, queremos agradecer ao professor António Teixeira e ao professor Miguel Oliveira Silva pela sua disponibilidade e por nos ter ajudado no desenvolvimento deste projecto.