

Projet ingénieur 2015
**P14. Reconnaissance
vocale pour la maison
connectée**



Encadrants : C. LOHR - Département: Informatique et D. PASTOR - Département : Signal et Communication.

Partenaires : M. CISSE, Société ATHEMIUM.

Mots clés: Domotique, reconnaissance vocale, algorithme DTW, API Google, localisation de source, pilotage microphone

Résumé : La société Athémium possède actuellement un système domotique de reconnaissance vocale utilisant la reconnaissance du navigateur Chrome. L'objectif est d'améliorer cette solution en proposant un système de reconnaissance hors navigateur à travers l'API Google (interface pour la reconnaissance vocale). L'envoi de requête doit être ensuite limité par la reconnaissance d'un mot-clé. Deux nouvelles fonctionnalités ont été ajoutées : l'utilisation d'un micro déporté Bluetooth ainsi que la localisation de la source de voix afin, par exemple, d'allumer la lumière du salon ou de la cuisine en fonction de la position de l'utilisateur dans la pièce.

1. Présentation et contexte du projet.

La société Athemium, a choisi de développer une box permettant à ses usagers de contrôler et de commander sa maison en énonçant des ordres à voix haute. Par le biais de la reconnaissance vocale, la box applique cet ordre comme le contrôle des volets ou de la lumière. C'est l'attractivité des démonstrations de la société qui fait son efficacité. La box domotique N.A.D.I.A offre une solution utilisant la reconnaissance vocale Chrome et une interface pour les utilisateurs. L'objectif du projet est donc d'améliorer ces démonstrations grâce à l'ajout de nouvelles fonctionnalités comme la localisation de la source de voix ou encore la liaison sans fil entre le micro et la box. D'autre part, la solution du client nécessite le navigateur Chrome, nous avons donc proposé une solution fonctionnant hors navigateur par l'intermédiaire d'une interface Google.

2. Méthodologie développée pour aboutir

En premier lieu, nous avons dû comprendre et redéfinir le besoin du client Athémium, puis avons réalisé un cahier des charges fonctionnel afin que le groupe de projet et le client conviennent ensemble d'une direction commune, et s'entendent sur les différentes tâches qui seraient à accomplir ou non.

La décision des tâches à accomplir ayant été proposée par le groupe et validée par le client, nous avons décidé de séparer le groupe en quatre pôles : reconnaissance vocale, choix du micro, identification de la provenance de la parole et implémentation des requêtes pour l'API Google.

Le projet se termine sur une phase d'intégration visant à proposer au client un prototype permettant l'affichage d'un ordre donné par un utilisateur ainsi que la position de l'utilisateur dans la pièce.

3. Développement des différentes tâches et principaux résultats

3.1 Schéma du prototype

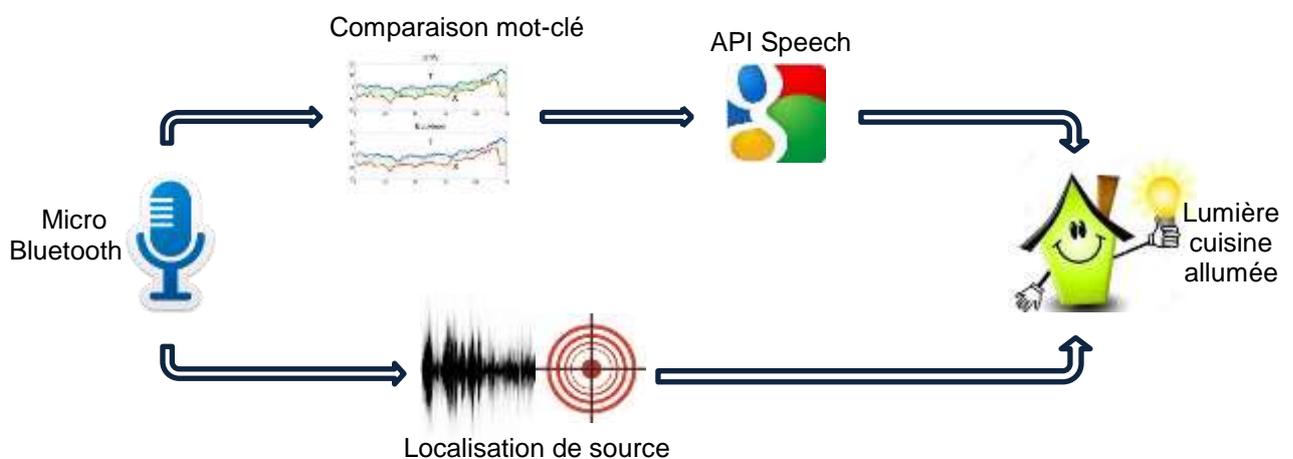


Figure 1 : Schéma du prototype

Le projet s'est articulé en deux principaux jalons : une phase de développement et une phase d'intégration.

3.2 Développement

3.1.1 Pilotage du micro et détection d'un mot clé

Le microphone doit écouter en continu dans la pièce et lorsque le mot-clé permettant de donner un ordre est entendu, une requête est envoyée à l'API. Cette étape nécessite le pilotage de la carte son (en Python) et donc l'installation de drivers sur une machine dédiée de l'école. La détection du mot-clé est ensuite réalisée par l'algorithme DTW de comparaison à un mot-clé référence.

3.1.2 Envoie de requête à Google

Une fois le mot-clé entendu, l'ordre de l'utilisateur est envoyé à l'API sous forme de fichier son FLAC et la réponse est la transcription écrite ainsi que la probabilité d'erreur. L'envoi de

requête est limité à 50 par jour, au-delà, une licence est nécessaire. Le temps de réponse de l'API est variable : entre 1 et 4 secondes selon la longueur de l'ordre.

3.1.3 Localisation de la source

Grâce à l'utilisation d'un réseau d'antennes, il est possible de déterminer d'où provient une source électromagnétique. Une simulation Matlab a permis d'obtenir des résultats satisfaisants. Ensuite, une transposition de fréquence a été nécessaire pour permettre l'utilisation de l'algorithme avec les ondes sonores.

3.1.4 Choix du micro

Un microphone en liaison Bluetooth avec dispositif anti-bruit a dû être acheté afin de répondre à l'impératif du micro déporté et à l'exigence de qualité de la reconnaissance vocale.

3.1 Intégration

Cette phase permet de proposer un prototype global comportant les solutions développées et intégrées et a nécessité la transposition de certains algorithmes vers Python ainsi que l'intégration des différents composants achetés.

4. Conclusions et perspectives.

Ce projet a permis de répondre à des besoins concrets de l'entreprise Athemium. Tout d'abord nous avons pu mettre en place un prototype qui utilise l'Api Speech de Google sans passer par Chrome. Deuxièmement, notre méthode de reconnaissance vocale permet de déclencher l'envoi de requêtes uniquement lorsqu'un mot clé a été prononcé. Cela évite l'envoi continu de données à l'API. Enfin nous avons voulu mettre en place un algorithme permettant de détecter la provenance de la parole. Ceci peut se révéler intéressant si l'on imagine un utilisateur présent dans une grande pièce composée d'un coin salon et d'un coin cuisine. La provenance de l'ordre permet de l'appliquer conformément à la position de l'utilisateur. La finalité de ce projet est de montrer qu'on a été capable de réaliser un prototype proposant de nouvelles fonctionnalités afin d'offrir au client de nouvelles perspectives pour sa box domotique.