

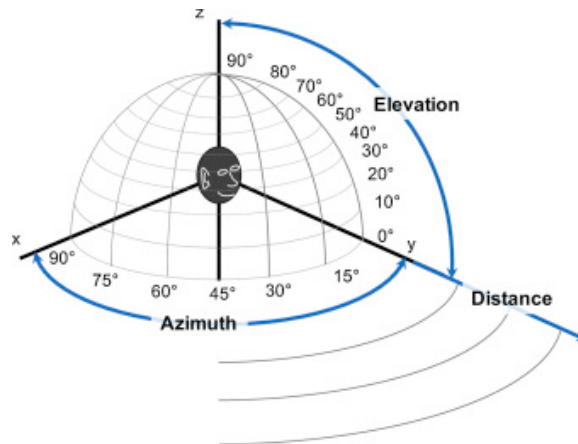
Localisation_sonore

December 17, 2021

1 Localisation sonore

1.1 Introduction

Le cerveau humain est capable grâce aux informations venant du système auditif de déterminer la direction et la distance approximative d'une onde sonore. En se plaçant en coordonnées sphériques ayant pour centre la tête, l'origine d'un son est déterminée par 3 paramètres : l'azimuth (angle sur le plan horizontal), l'élévation (angle sur le plan vertical) et la distance. Pour un être humain, la précision en azimuth est de $8,5^\circ$ autour de la vraie position. Cependant, repérer l'élévation et la distance d'un son est plus difficile [1]. Plusieurs indices sont utilisés par le cerveau afin de déduire l'origine d'un son. Tout d'abord, le décalage de l'onde sonore entre les deux oreilles permet la détermination de l'azimuth. La perte d'amplitude du signal permet d'estimer la distance. Enfin, de nombreux autres phénomènes plus complexes sont pris en compte comme le traitement de l'onde transmise par la tête d'une oreille à l'autre (fonction de transfert de la tête) ou bien l'atténuation différentes des fréquences en fonction de la distance [2].



1.2 Etat de l'art

Des systèmes de visio-conférence intelligents existent déjà [3]. Cependant, cela fonctionne par triangulation et corrélation de signal et nécessite une calibration ainsi que la présence de plusieurs micros à des endroits différents.

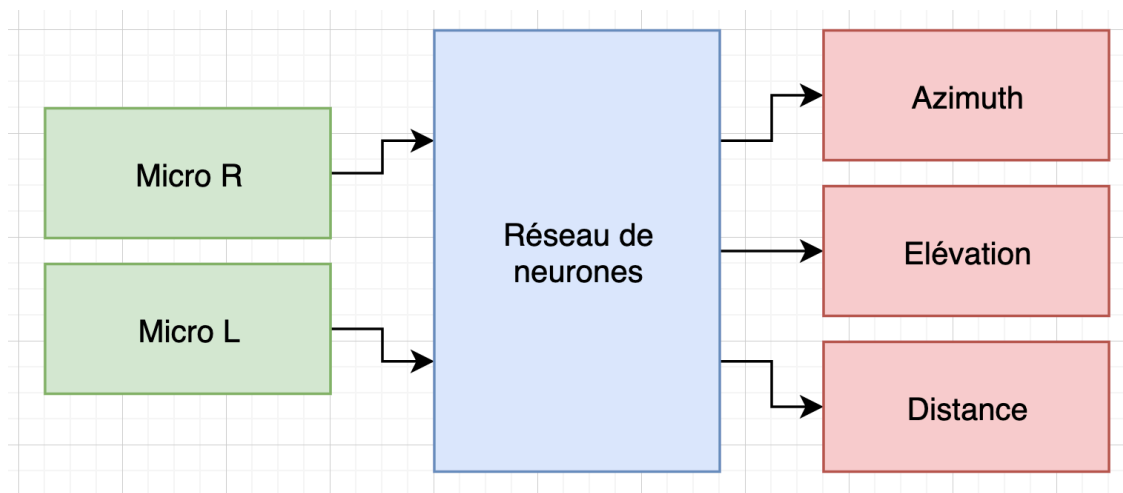
Une étude a été menée en utilisant un réseau de neurones profond à convolution [4]. Le nombre d'exemples utilisés est de 2,43 millions. Pour entraîner leur modèle sur 70 exemples, une epoch

met 24 h en utilisant un GPU NVIDIA Titan X.

1.3 Objectifs

Le premier objectif est de constituer un dataset de taille suffisante, idéalement dans différents environnements. Pour cela, une première idée est d'utiliser les caméras de motion capture présentes sur le campus et de déplacer une source sonore (dont la composition spectrale varie en fonction du temps). Ainsi, cela fourni 3 fichiers intéressants : la position de la source sonore, la position du récepteur et l'audio reçu par le récepteur (voir avec Renaud Séguier si cela est faisable). Dans un second temps, le premier objectif sera d'implémenter un algorithme d'IA afin de déterminer la position de la source sonore par rapport au récepteur (selon l'axe d'azimuth) puis ajouter la notion de distance et d'élévation. Il sera intéressant de voir à quel point l'algorithme sera robuste à un changement d'environnement sonore et de type de sons émis.

Voici un premier graphique décrivant l'architecture de l'algorithme :



1.4 Références

- [1] Accuracy-Precision Trade-off in Human Sound Localisation
- [2] Sound localization
- [3] Caméra de visio-conférence
- [4] Réseaux de neurones pour la localisation sonore

[] :