

Méthodologie 3-way d'extraction d'un modèle articulatoire de la parole à partir des données d'un locuteur

Martine Cadot, Yves laprie

LORIA

martine.cadot@loria.fr, yves.laprie@loria.fr,

<http://www.loria.fr/~cadot>

<http://www.loria.fr/~laprie>

Résumé. Pour parler, le locuteur met en mouvement un ensemble complexe d'articulateurs : la mâchoire qu'il ouvre plus ou moins, la langue à laquelle il fait prendre de nombreuses formes et positions, les lèvres qui lui permettent de laisser l'air s'échapper plus ou moins brutalement, etc. Le modèle articulatoire le plus connu est celui de Maeda (1990), obtenu à partir d'Analyses en Composantes Principales faites sur les tableaux de coordonnées des points des articulateurs d'un locuteur en train de parler. Nous proposons ici une analyse 3-way du même type de données, après leur transformation en tableaux de distances. Nous validons notre modèle par la prédiction des sons prononcés, qui s'avère presque aussi bonne que celle du modèle acoustique, et même meilleure quand on prend en compte la co-articulation.

1 Introduction

Construire un modèle articulatoire de la parole, c'est être capable d'indiquer les mouvements des articulateurs (mâchoires, lèvres, etc.) à l'origine de celle-ci (voir figure 1, à gauche). Les applications pratiques d'un tel modèle sont nombreuses ¹.

Nous exposons ici comment nous avons extrait un modèle articulatoire à partir de données recueillies auprès d'un locuteur. Ce travail se situe dans la lignée des travaux initiés par Maeda (1990). Il a construit son modèle articulatoire (voir figure 1, à droite) au moyen d'analyses en composantes principales sur des données de même type. Puis il l'a évalué de façon acoustique en comparant les sons réels aux sons produits par un synthétiseur de sons piloté par son modèle. La nouveauté de notre démarche consiste en l'utilisation d'une méthode d'analyse 3-way pour extraire le modèle, et de méthodes d'apprentissage supervisé pour le valider. Notre évaluation se fait en comparant de façon phonétique les sons prédits aux sons réels. L'acoustique intervient de surcroît dans notre évaluation car nous mettons en parallèle les performances de notre modèle et celles du modèle acoustique formé des coefficients *cepstraux*².

1. Pour plus de détails, voir <http://parole.loria.fr/>.

2. Ces coefficients représentant le signal de parole ont été extraits à l'aide de la bibliothèque R de Ligges (2011)