

# Préface

Ce livre est le résultat d'un long travail de recherche fructueuse menée par l'équipe de traitement de la parole de Nancy. L'activité de cette équipe en reconnaissance de la parole a débuté vers 1970 sous la direction du professeur Jean-Paul Haton. C'est dans cette période, en fréquentant le Groupement des Acousticiens de Langue Française (GALF), que j'ai connu et que j'ai commencé à apprécier ce travail. Depuis, une soixantaine de thèses sur le sujet ont été soutenues dans l'équipe. De nombreux anciens de l'équipe de Nancy occupent aujourd'hui des positions élevées en traitement automatique de la parole dans le monde entier. Sans citer les noms pour éviter d'en oublier quelques-uns parce qu'ils sont tellement nombreux, on observe dans des publications majeures des anciens de Nancy, qu'ils travaillent dans des groupes de recherches d'importantes universités ou de sociétés telles que Microsoft, IBM, ATT, Texas Instruments, Apple, Panasonic. Les coauteurs de ce livre sont aussi des anciens doctorants du professeur Haton.

Les résultats de l'équipe s'inscrivent dans les lignes majeures de développement de la culture mondiale. Partant de la reconnaissance des formes, la production scientifique de l'équipe apparaît dans des domaines importants de l'intelligence artificielle, de la théorie des langages formels, de la phonétique, de la communication homme-machine, de la théorie de l'information, des réseaux de neurones, de la modélisation statistique, du traitement du signal, des nouvelles tendances sur les stratégies de génération et d'évaluation d'hypothèses.

Ce livre présente de façon claire et complète les théories et les principaux domaines d'application de la reconnaissance automatique de la parole.

Le premier chapitre introduit les principes du décodage du signal de parole. Après

une bref historique des problèmes, les méthodes d'analyse et paramétrisation du signal vocal sont introduites. Sont présentés clairement, dès le début, les liens entre les connaissances explicites des domaines et la modélisation stochastique des traits acoustiques et du langage. Comme il n'existe pas de système de reconnaissance automatique de la parole qui ne fasse pas d'erreurs, le problème de la robustesse des systèmes est d'importance fondamentale. Ce problème est aussi introduit dans le premier chapitre.

Le deuxième chapitre est dédié à la communication parlée avec des notions de production et de perception de la parole humaine, ainsi que des éléments d'anatomie et de physiologie. Des concepts de psychoacoustique, de phonologie du français et de prosodie sont aussi présentés. Cet aspect reflète le souci toujours présent dans l'école de Nancy de concevoir des modèles efficaces en utilisant le plus possible des connaissances spécifiques de différentes disciplines.

Le troisième chapitre décrit les méthodes d'analyse du signal de parole en relation avec les connaissances d'audition, de perception et de production de la parole humaine.

Le quatrième chapitre montre comment construire des modèles acoustiques avec des chaînes de Markov cachées et estimer les paramètres qui les caractérisent. D'autres modèles stochastiques alternatifs sont décrits, comme les modèles de trajectoires, les modèles multibandes, les réseaux bayésiens, les machines à vecteurs supports et les réseaux neuromimétiques.

Le cinquième chapitre introduit des techniques avancées pour l'amélioration des modèles et pour leur utilisation dans des systèmes à grand vocabulaire. Le chapitre contient aussi une description des mesures de confiance et des tests d'hypothèses qui peuvent être utilisés pour l'évaluation et la combinaison de systèmes.

Le sixième chapitre décrit la modélisation statistique du langage avec des contributions originales de l'équipe surtout pour ce qui concerne la modélisation de séquences. Une approche statistique à la traduction automatique y est aussi présentée.

Le septième chapitre discute de la compréhension automatique de la parole et des paradigmes d'évaluation.

Le huitième chapitre s'occupe de la robustesse des systèmes de reconnaissance automatique et du débruitage du signal de parole. On retrouve dans ce chapitre les résultats d'un grand travail de recherche de l'équipe avec des contributions scientifiques importantes.

Le neuvième chapitre est dédié au développement d'un système. Ce chapitre est très utile pour les professionnels qui développent des applications. On y décrit les procédures d'entraînement des modèles, la génération des différentes prononciations des mots, le choix du vocabulaire et la conception d'un lexique phonétisé. Le chapitre décrit également la construction de modèles de langage et leur utilisation dans un moteur de reconnaissance. Les problèmes d'adaptation au locuteur et à l'environnement sont aussi discutés en vue d'une utilisation dans une architecture générale.

Le dixième chapitre décrit les résultats de recherche de l'équipe concernant un cadre articulatoire pour la reconnaissance automatique de la parole.

Le onzième chapitre s'occupe des standards utilisés en reconnaissance automatique de la parole, des composants disponibles aujourd'hui dans des environnements de programmation, ainsi que des grands domaines d'application de la reconnaissance automatique de la parole.

Cet ouvrage est unique pour les apports originaux de l'équipe de Nancy et pour la profondeur avec laquelle les sujets les plus importants sont traités. L'ouvrage s'adresse autant aux chercheurs, étudiants et développeurs industriels de technologies vocales qu'au grand public scientifique intéressé à mieux connaître le domaine.

Je me permets de féliciter les auteurs de cette ouvrage pour leur excellent travail. Le professeur Jean-Paul Haton mérite une reconnaissance particulière pour les résultats de sa recherche scientifique et pour avoir été le guide et l'inspirateur d'une quantité impressionnante de chercheurs de valeur. Je souhaite à ce livre le grand succès que ses auteurs et son contenu méritent.

Renato De Mori

Professeur, Université d'Avignon et des Pays de Vaucluse