Pour les articles homonymes, voir ASR.

La reconnaissance automatique de la parole (souvent improprement appelée reconnaissance vocale) est une technique informatique qui permet d'analyser la voix humaine captée au moyen d'un microphone pour la transcrire sous la forme d'un texte exploitable par une machine.

La reconnaissance de la parole, ainsi que la synthèse de la parole, l'identification du locuteur ou la vérification du locuteur, font partie des techniques de traitement de la parole. Ces techniques permettent notamment de réaliser des interfaces homme-machine (IHM) où une partie de l'interaction se fait à la voix : « interfaces vocales ».

Parmi les nombreuses applications, on peut citer les applications de dictée vocale sur ordinateur où la difficulté tient à la taille du vocabulaire et à la longueur des phrases, mais aussi les applications téléphoniques de type serveur vocal interactif, où la difficulté tient plutôt à la nécessité de reconnaître n'importe quelle voix dans des conditions acoustiques variables et souvent bruyantes (téléphones mobiles dans des lieux publics).

Dans Parole et dialogue homme-machine, W. Minker et S. Bennacef expliquent que la reconnaissance automatique de la parole est un domaine complexe, car il existe une différence importante entre le langage formel, qui est compris et utilisé par les machines, et le langage naturel, que les humains utilisent. Le langage formel est structuré par des règles syntaxiques strictes et sans ambigüité. À l'inverse, dans le langage naturel, des mots ou des phrases peuvent avoir plusieurs sens selon l'intonation de l'énonciateur ou le contexte par exemple.

La reconnaissance de la parole peut se rattacher à de nombreux plans de la science : traitement automatique des langues, linguistique, théorie de l'information, traitement du signal, réseaux de neurones, intelligence artificielle, etc.

Les travaux sur la reconnaissance de la parole datent du début du XXe siècle. Le premier système pouvant être considéré comme faisant de la reconnaissance de la parole date de 1952.

Ce système électronique, développé par Davis, Biddulph et Balashek aux laboratoires Laboratoires Bell, était essentiellement composé de relais et ses performances se limitaient à reconnaître des chiffres isolés (voir référence). La recherche s'est ensuite considérablement accrue durant les années 1970 avec les travaux de Jelinek chez IBM (1972-1993). La société Threshold Technologies fut la première à commercialiser en 1972 un système de reconnaissance d'une capacité de 32 mots, le VIP100. Aujourd'hui, la reconnaissance de la parole est un domaine à forte croissance grâce à la déferlante des systèmes embarqués.

Une évolution rapide :

Depuis 2024, de nombreux logiciels de transcriptions utilisent l'intelligence artificielle[7].

Une phrase enregistrée et numérisée est donnée au programme de reconnaissance automatique de la parole (RAP). Dans le formalisme RAP (ASR en anglais), le découpage fonctionnel est le suivant

Enfin, en suivant le principe que les propriétaires d'une machine sont les destinataires des fruits de la production d'une machine automatisée, on peut en déduire que l'automatisme permet un esclavage mécanique qui libère petit à petit l'humain de toute forme de domination de l'homme sur l'homme et de l'homme sur la nature.

La robotique industrielle est officiellement définie par l'Organisation Internationale de Normalisation (ISO) comme étant un système commandé automatiquement, multi-applicatif, reprogrammable, polyvalent, manipulateur et programmable sur trois axes ou plus.

Les applications typiques incluent les robots de soudage, de peinture et d'assemblage. L'avantage de la robotique industrielle est sa rapidité d'exécution et sa précision ainsi que la répétition de cette précision dans le temps[2].

Un automate programmable industriel

Cet article possède un paronyme, voir Bigdatha.

Mégadonnées

Cet article sur l'informatique doit être recyclé (novembre 2020).