

République du Cameroun

*Paix – Travail – Patrie*

Université de Yaoundé I

*Sapientia – Collatitia – Cognitionis*

École Nationale Supérieure  
Polytechnique



Republic of Cameroon

*Peace – Work – Fatherland*

University of Yaounde I

*Sapientia – Collatitia – Cognitionis*

National Advanced School of  
Engineering

# SYSTÈME DE RECONNAISSANCE VOCALE ET D'EMPREINTE DIGITALE POUR LE CONTRÔLE D'ACCÈS ÉTAT DE L'ART

**Supervisé par :**

- Dr. CHANA LEMALE
- Dr. NGOUNOU

**ETUDIANT 4GI**

**2022 - 2023**

## Membres du groupe

- ✓ MENRA ROMIAL (CHEF)
- ✓ KEGNE CHATUE (SOUS-CHEF)
- ✓ ALFRED HETSRON YEPNJO
- ✓ DJOUMESSI GUEPI AUREL
- ✓ EWOKI EBOUELE ANGE
- ✓ MEVONGO ZE IVAN
- ✓ YEMKWA EMMANUEL
- ✓ ZIINAM DANATA

## Table des matières

INTRODUCTION.....	2
I. LES OBJECTIFS .....	2
II. MOTIVATIONS ET PRÉSENTATION DU PROJET .....	2
III. EMPREINTES DIGITALES.....	3
1. Principe de fonctionnement .....	3
2. Différentes techniques utilisées.....	4
a. Technique optique.....	4
b. La technique silicium .....	5
c. La technique ultrason .....	5
IV. RECONNAISSANCE VOCALE.....	6
• Principe de fonctionnement .....	6
V. QUELQUES EXEMPLES DE SYSTÈME DE RECONNAISSANCE VOCALE ET EMPREINTE DIGITALE.....	7
1. Le système de contrôle d'accès biométrique MYM7 [1] .....	7
2. Alibaba Cloud Interaction vocale intelligente .....	8
3. Transcription Amazon [2].....	9
4. Dragon Nuance .....	10
CONCLUSION .....	12
Bibliographie.....	13

## INTRODUCTION

Pour mener à bien le déroulement de notre travail qui consiste à concevoir et développer un système de reconnaissance vocale et d'empreinte digitale pour le contrôle d'accès, nous devons commencer comme dans toute démarche scientifique et rigoureuse par une étude de l'existant. Elle consistera à répertorier toutes les informations relatives aux techniques de reconnaissance vocale et d'empreinte digitale. Dans ce document nous présenterons le principe de fonctionnement de ces systèmes, les techniques employés et quelques exemples de systèmes.

### I. LES OBJECTIFS

L'objectif principal de notre projet est de répertorier les présences des étudiants au sein du campus de l'ENSPY afin de s'affranchir des listes d'appels classiques. La décision d'associer le module de reconnaissance vocale à celui de l'empreinte digitale a pour but d'augmenter la fiabilité du système.

### II. MOTIVATIONS ET PRÉSENTATION DU PROJET

Face aux menaces du terrorisme ou de la cybercriminalité, et face à l'évolution logique des réglementations internationales, de nouvelles solutions technologiques sont progressivement mises en œuvre. Parmi ces technologies, la biométrie s'est rapidement distinguée depuis 1997 comme la plus pertinente pour identifier et authentifier les personnes de manière fiable et rapide, en fonction de caractéristiques biologiques uniques.

La biométrie est la science qui porte sur l'analyse des caractéristiques physiques ou comportementales propres à chaque individu et permettant l'authentification de son identité. Au sens littéral et de manière plus simplifiée, la biométrie signifie la "mesure du corps humain". On distingue deux catégories de technologies biométriques : les mesures physiologiques, et les mesures comportementales. Les mesures physiologiques peuvent être morphologiques ou biologiques. Ce sont surtout les empreintes digitales, la forme de la main, du doigt, le réseau veineux, l'œil (iris et rétine), ou encore la forme du visage, pour les analyses morphologiques.

Les mesures comportementales les plus répandues sont la reconnaissance vocale, la dynamique des signatures (vitesse de déplacement du stylo, accélérations, pression exercée, inclinaison), la dynamique de frappe au clavier d'un ordinateur, la façon d'utiliser des objets, la démarche, le bruit des pas, la gestuelle... Les différentes techniques utilisées font l'objet de recherches régulières, de développements et bien entendu, d'améliorations constantes. Toutefois, les différentes sortes de mesures n'ont pas le même niveau de fiabilité.

On estime que les mesures physiologiques ont l'avantage d'être plus stables dans la vie d'un individu. Par exemple, elles ne subissent pas autant les effets du stress, contrairement à l'identification par mesure comportementale.

La biométrie permet l'identification et l'authentification d'une personne à partir de données reconnaissables et vérifiables, qui lui sont propres et qui sont uniques.

L'identification consiste à déterminer l'identité d'une personne. Il s'agit de saisir une donnée biométrique de cette personne, en prenant par exemple une photo de son visage, en enregistrant sa voix, ou en captant l'image de son empreinte digitale. Ces données sont ensuite comparées aux données biométriques de plusieurs autres personnes qui figurent dans une base.

L'authentification, appelée également vérification, est le processus qui consiste à comparer les données caractéristiques provenant d'une personne, au modèle de référence biométrique de cette dernière (« template »), afin de déterminer la ressemblance. Le modèle de référence est préalablement enregistré et stocké dans une base de données, dans un équipement ou objet personnel sécurisé. On vérifie ici que la personne présentée est bien la personne qu'elle prétend être...

Ci-dessus est donc présenté les motivations de la réalisation d'un système de reconnaissance vocale et un bref aperçu du fonctionnement de celui-ci

### III. EMPREINTES DIGITALES

La technique des empreintes digitales est une des techniques les plus anciennes, elle a été développée vers la fin du 19ème siècle par Alphonse Bertillon, fondateur de la police scientifique en France. A cette époque et jusqu'à récemment, une tablette et un encreur sont le matériel utilisé pour la capture d'empreinte. Le premier système automatique d'authentification a été commercialisé au début des années 1960.

#### 1. Principe de fonctionnement

L'authentification par les empreintes digitales repose sur la concordance entre le fichier d'enregistrement, ou « signature », obtenu lors de l'enrôlement et le fichier obtenu lors de l'authentification.

Ces deux fonctions se décomposent chacune en plusieurs étapes :

##### ❖ Enrôlement

- Capture de l'image de l'empreinte. Les données d'un doigt sont en principe suffisantes à l'enrôlement, mais la plupart des systèmes enregistrent au moins deux doigts (un par main par exemple) pour parer l'indisponibilité résultant de petites blessures.
- Numérisation de l'image afin d'extraire les minuties, ou éléments caractéristiques.
- Enregistrement sur un support. (Carte à puce, disque dur...)

#### ❖ Authentification

- Capture de l'image de l'empreinte.
- Numérisation de l'image afin d'extraire les minuties, ou éléments caractéristiques.
- Comparaison entre l'échantillon et le gabarit « signature ».
- Prise de décision.

Lors de la capture de l'image, celle-ci est toujours constituée à partir des points de contact du doigt sur le capteur.

#### ❖ Etapes de traitement

- Lorsque la capture de l'image est réalisée, elle doit être convertie dans un format approprié. L'extraction des minuties est réalisée grâce à différents algorithmes. Il s'agit ensuite par une technique mathématique (segmentation) d'éliminer les informations non utiles au système : niveau de bruit trop élevé (image sale, doigt mal placé).
- L'image est numérisée. Afin de localiser précisément les terminaisons et les bifurcations, les crêtes sont affinées de 5 à 8 pixels à 1 pixel.

A ce stade, l'image a des distorsions et de fausses minuties, qui peuvent être dues à des cicatrices, de la sueur, un défaut de propreté du doigt comme du capteur. Les minuties vont être filtrées afin de ne conserver que les plus fiables.

Les avis divergent sur le rapport de proportion entre minuties extraites pour l'enrôlement et minuties suffisamment fiables pour la vérification. A partir de 31 minuties extraites, seulement 10 pourront correspondre lors de l'authentification.

## 2. Différentes techniques utilisées

### a. Technique optique

C'est, après l'encre, la technique la plus ancienne et qui a fait ses preuves. Le principe physique utilisé est celui de « la réflexion totale frustrée » : Le doigt est placé sur un capteur éclairé par une lampe. Une caméra CMDs (Charge Modulation Device) avec CCD (Charged Coupled Device en français : DTC : Dispositif à Transfert de Charge ) convertit l'image, composée de crêtes foncées et de vallées claires, en un signal vidéo retraité afin d'obtenir une image utilisable. Nous pouvons différencier les terminaux en lumière visible à fenêtre sèche et à fenêtre à film liquide (la fenêtre est l'emplacement où l'utilisateur pose le doigt). Dans ce dernier cas, la fenêtre est nettoyée avant chaque prise de vue par un mélange d'eau et d'éthanol injecté sous le doigt. Des terminaux à image infra-rouge par capteur linéaire intégré sont parfois utilisés, mais présentent les mêmes inconvénients que ceux à lumière visible.

Avantages	Inconvénients
Son ancienneté et sa mise à l'épreuve	Il est possible que l'empreinte d'utilisateurs précédents reste latente, d'où une

	possibilité de dégradation de l'image par sur-impression.
Sa résistance aux changements de température, jusqu'à un certain point.	Apparition possible de rayures sur la fenêtre.
Son coût abordable	D'autre part, le dispositif CCD peut s'user avec le temps et devenir moins fiable.
Sa capacité à fournir des résolutions de plus de 500 dpi.	Problèmes de contrastes (doigt propre et sec devient trop clair tandis qu'un doigt humide et recouvert d'un film gras devient très foncé), problème résolu grâce au film liquide mais système mal accepté.

#### b. La technique silicium

Cette technique est apparue à la fin des années 90. Le doigt est placé sur un capteur CMDs. L'image est transférée à un convertisseur analogique-numérique, l'intégration se faisant en une seule puce. Cette technique produit des images de meilleure qualité avec une surface de contact moindre que pour la technique optique. Les données fournies sont très détaillées. Elle possède une bonne résistance dans des conditions non-optimales. Cette technique est adaptée à un développement de masse, notamment par ses coûts réduits.

Avantages	Inconvénients
Coût assez bas.	Capteur vulnérable aux attaques extérieures fortuites ou volontaires.

#### c. La technique ultrason

Très peu utilisée à ce jour, elle repose sur la transmission d'ondes acoustiques et mesure l'impédance entre le doigt, le capteur et l'air. Cette technique permet de dépasser les problèmes liés à des résidus sur le doigt ou sur le capteur. Cette technique peut aussi être utilisée par des scanners à ultrasons qui construisent par échographie une image ultra sonore. Elle est considérée comme la plus fiable.

Avantages	Inconvénients
Facilité d'usage avec de grandes plaques	Aucun inconvénient technique significatif n'a pu être identifié à ce jour au travers des textes et des témoignages des experts.
Capacité à surmonter des conditions de lecture non optimales (les poussières sont souvent transparentes aux ultrasons)	Coût élevé

## IV. RECONNAISSANCE VOCALE

La reconnaissance de la voix n'est pas intrusive pour la personne et n'exige aucun contact physique avec le lecteur du système. Le logiciel de reconnaissance peut être centralisé et la voix transmise par le réseau, d'où un impact de réduction des coûts. Le dispositif nécessite un micro en source de capture. Les systèmes d'identification de la voix sont basés sur les caractéristiques de voix, uniques pour chaque individu. Ces caractéristiques de la parole sont constituées par une combinaison des facteurs comportementaux (vitesse, rythme, etc...) et physiologiques. (Tonalité, âge, sexe, fréquence, accent, harmoniques, ...).

### Principe de fonctionnement

Pour être stockée, la voix est numérisée puis segmentée par unités échantillonnées. Les méthodes sont basées sur des algorithmes mathématiques (Shannon). Les systèmes d'identification de la voix utilisent soit un texte libre, soit un texte imposé, les mots devant être lus devant un micro.

Comme avantages nous avons :

- Disponible via le réseau téléphonique.
- Les imitateurs utilisent les caractéristiques vocales sensibles au système auditif humain, mais ne sont pas capables de récréer les harmoniques de la voix, servant de base à l'identification. Il est quasi impossible d'imiter la voix stockée dans la base de données.
- Non intrusif.

Les inconvénients sont :

- Fraude possible en utilisant un enregistrement de la voix de la personne autorisée, facilitée dans le cas de système basé sur la lecture d'un texte fixe.
- L'utilisation d'un micro nécessite un dispositif adapté présent sur l'environnement.
- Sensibilité à l'état physique et émotionnel d'un individu.
- Sensibilité aux conditions d'enregistrement du signal de parole : bruit ambiant, parasites, qualité du microphone utilisé, qualité de l'équipement, lignes de transmission.

Les inconvénients signalés montrent que ce système est vulnérable et doit être utilisé couplé avec un système d'identification.



## V. QUELQUES EXEMPLES DE SYSTÈME DE RECONNAISSANCE VOCALE ET EMPREINTE DIGITALE

### 1. Le système de contrôle d'accès biométrique MYM7 [1]



*Système de contrôle d'accès biométrique*

Le MYM7 contrôle d'accès avec empreintes digitales est un système de contrôle d'accès de porte professionnel développé pour les petites et moyennes entreprises afin d'assurer une sécurité physique des lieux. Il intègre l'identification d'empreinte digitale, le RFID, le suivi de présence et les fonctions de contrôle d'accès de sécurité etc., avec un aspect élégant et une qualité fiable. Il est multi langues et dispose d'une interface conviviale et une puissante fonction de communication adaptée à la gestion des données dans divers environnements.

Le MYM7 est livré avec un logiciel de gestion fonctionnelle, compatible avec différents types de bases de données et prend en charge le fuseau horaire et le contrôle d'accès de groupe pour un niveau de sécurité élevé est polyvalent, pratique et multifonctionnel. Il possède les caractéristiques suivantes :

- Design ergonomique et élégant
- Il supporte d'identification sous différents angles
- Il supporte plusieurs langues
- Commande de verrouillage direct sur porte ouverte
- Capteur de porte ouverte
- Écran couleur TFT de 2,8 pouces RS485,
- Disque USB et communication TCP / IP Carte MF optionnelle,
- Module carte RFID.



## 2. Alibaba Cloud Interaction vocale intelligente

Vue d'ensemble : Alibaba, la grande entreprise chinoise du cloud, utilise des technologies telles que la synthèse vocale, la reconnaissance vocale et la compréhension du langage naturel pour créer son offre d'interaction vocale intelligente. Il est actuellement accessible dans les langues suivantes : chinois cantonais, chinois mandarin, japonais, anglais, français, coréen et indonésien, avec d'autres langues en cours de route.

Principales caractéristiques : Les principales caractéristiques d'Alibaba Cloud Intelligent Speech Interaction incluent :

**Haute précision** : bien que l'entreprise ne révèle pas le niveau de précision exact, la plateforme peut s'auto-apprendre.

**Capacités de transcription** : Il peut traiter des transcriptions multilingues en temps réel et à partir de fichiers préenregistrés.

**Entraînement du modèle IA et ML** : les utilisateurs peuvent entraîner le modèle pour réduire les erreurs de 20 %.

**Support développeur** : Il propose une large gamme d'API et un guide du développeur.

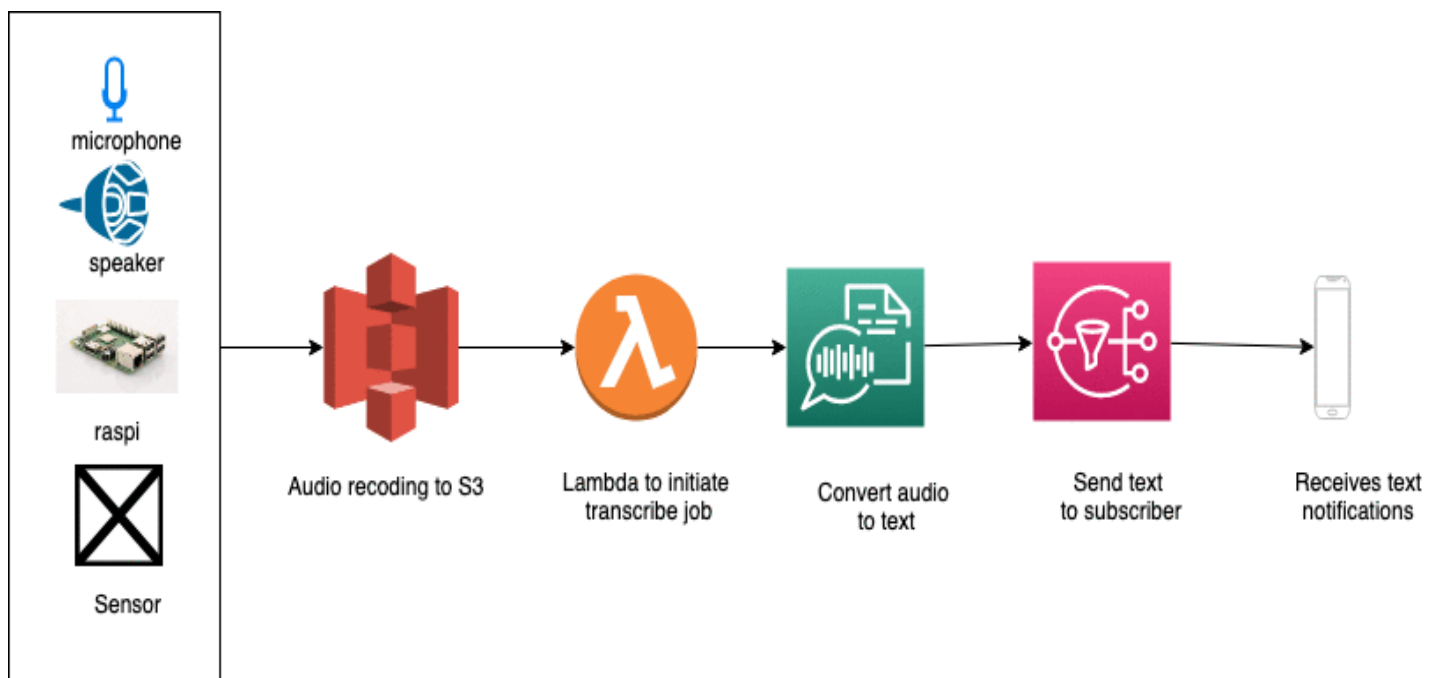
Préparation à l'entreprise : il dispose de solutions d'entreprise prédéfinies pour le service client, le sous-titrage en temps réel et l'analyse des appels de service.

**USP** : Alibaba Cloud Intelligent Speech Interaction utilise une technologie de décodage innovante à faible fréquence d'images (LFR). Cela réduit considérablement le temps de réponse sans compromettre la précision.

**Prix** : Le prix commence à 1,00 \$ de l'heure pour les fichiers enregistrés et à 1,40 \$ de l'heure pour la reconnaissance vocale en temps réel.

**Commentaires éditoriaux** : La plateforme est riche en fonctionnalités et convient à la reconnaissance de phrases courtes. Cependant, la courbe d'apprentissage peut être abrupte pour les entreprises qui découvrent l'environnement cloud d'Alibaba.

### 3. [Transcription Amazon \[2\]](#)



*Transcription amazon*

**Présentation** : Amazon Transcribe est un logiciel de reconnaissance vocale d'Amazon Web Services (AWS). Il est simple d'ajouter des capacités de synthèse vocale via le traitement du langage naturel. Ses capacités vous permettent de prendre en charge l'entrée audio, de créer des transcriptions faciles à lire et à réviser, de filtrer le matériel pour préserver la confidentialité des clients et d'augmenter la précision grâce à la personnalisation. Transcribe est une plateforme de transcription basée sur le cloud.

**Fonctionnalités clés** : les fonctionnalités clés d'Amazon Transcribe incluent :

**Haute précision** : Le logiciel offre des niveaux de précision d'environ 80 %.

**Capacités de transcription** : Il produit des transcriptions faciles à lire et à intégrer dans les applications professionnelles.

**Entraînement des modèles d'IA et de ML** : Il fournit dix transcriptions alternatives pour chaque phrase et apprend de vos entrées, en prenant en charge votre modèle de langage personnalisé (CLM).

**Assistance aux développeurs** : elle est extrêmement conviviale pour les développeurs, avec une formation sur l'utilisation de la plate-forme.

**Préparation à l'entreprise** : il est conforme aux réglementations d'entreprise telles que la loi HIPAA (Health Insurance Portability and Accountability Act) et prend en charge la rédaction automatique du contenu.

**USP** : Amazon Transcribe se concentre sur la confidentialité, la sécurité et la conformité. Cela signifie que des mesures spéciales pour les secteurs traitant des données sensibles comme les soins de santé sont en place.

**Tarification** : Amazon Transcribe est gratuit pendant 60 minutes par mois pendant un an et coûte 0,00780 USD par minute.

**Commentaires éditoriaux** : Transcribe offre un degré élevé de personnalisation. Cependant, son intégration dans vos systèmes peut nécessiter des efforts considérables.

#### 4. [Dragon Nuance](#)



*Dragon Nuance*

**Vue d'ensemble** : Ce logiciel de reconnaissance vocale a été développé pour la première fois en 1997 et a été acquis par de nombreuses entreprises jusqu'à ce qu'il appartienne à Nuance Communications et finalement à Microsoft. Il propose des solutions ASR pour divers cas d'utilisation, y compris des applications professionnelles et individuelles, des équipes d'entreprise, des professionnels du droit, des forces de l'ordre et un usage domestique, couvrant des applications pour les environnements Windows et mobiles.

**Haute précision** : Il offre jusqu'à 99 % de précision.

Capacités de transcription : les utilisateurs peuvent bénéficier d'un logiciel de transcription prêt à l'emploi et de l'édition de documents par commande vocale.

Entraînement des modèles d'IA et de ML : la prise en charge de la personnalisation est limitée, mais vous pouvez définir des commandes vocales personnalisées.

**Assistance aux développeurs** : elle offre de nombreuses ressources aux développeurs pour aider à créer des chatbots, des systèmes de messagerie et d'autres applications de reconnaissance vocale.

**Prêt pour l'entreprise** : les utilisateurs de l'entreprise peuvent installer le logiciel sur leur bureau et commencer à l'utiliser immédiatement.

**USP** : Nuance Dragon est simple à utiliser et à mettre en œuvre. Il est idéal pour les utilisateurs professionnels. Il prend également en charge Citrix, d'autres environnements virtualisés et un centre d'administration centralisé.

**Prix** : Il commence à 200 \$ pour Dragon Home pour Windows, et l'édition professionnelle commence à 150 \$ pour les abonnements annuels.

**Commentaires éditoriaux** : Nuance Dragon est un leader reconnu dans le segment des logiciels d'IA et de reconnaissance vocale. Cependant, les utilisateurs ont noté que le logiciel a parfois du mal avec la ponctuation.

## CONCLUSION

Au terme de notre devoir, il était question pour nous de présenter l'état de l'art d'un dispositif de reconnaissance vocale et d'empreinte digitale pour le contrôle d'accès. Il en ressort que, jusqu'à présent, ces différents dispositifs ont été développés séparément et présente chacun de multiples failles. L'intérêt de les associer est donc d'augmenter la fiabilité tout en utilisant les technologies appropriées et en optimisant les coûts.

## Bibliographie

- 1 :Amazon. (2022, Novembre 04). *aws.amazon*. Récupéré sur amazon:  
<https://aws.amazon.com/fr/blogs/machine-learning/category/artificial-intelligence/amazon-transcribe/page/4/>
- 2 :Sodishop. (2022, Novembre 04). *Sodishop*. Récupéré sur Sodishop:  
<https://www.sodishop.com/shop/securite-electronique/controle-acces/controle-dacces-a-empreinte-injes/>