
L'IA pour une interaction par les gestes

David Gómez
Enseignant-Chercheur en Informatique

Contenu

- Introduction
- Motivation
- Dispositifs pour l'acquisition des gestes
- Approche général
- Techniques, outils et challenges
- Projet de recherche: Rain of Music
- Perspectives

Introduction

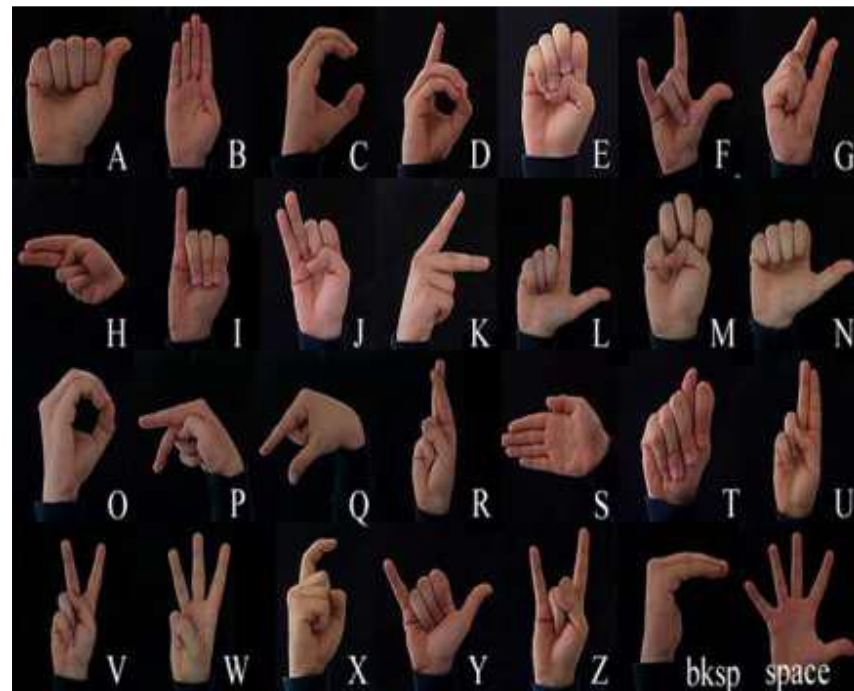
- Qu'est-ce que la communication gestuelle?
 - Une forme de communication non verbale.
 - Mouvement des mains, du visage ou d'autres parties du corps.
- Qu'est-ce que la reconnaissance des gestes?
 - Interpréter les gestes humains via des techniques d'intelligence artificielle (machine learning)
 - Les gestes comme entrée pour contrôler des appareils ou des applications.

Motivation pour l'interaction gestuelle

- Le geste est une forme de communication naturelle, universelle et puissante.
- Problèmes d'interaction avec les appareils traditionnels (par exemple, la souris)
 - Control limité ou difficile avec la souris (par exemple, le contrôle d'une souris dans un véhicule en mouvement ou le control des objets 3D dans un environnement VR)
- Fournir une interaction plus naturelle avec les machines (par exemple avec les robots)

Applications: langue de signes

Tout comme la reconnaissance vocale transcrit la parole en texte, certains systèmes d'interaction gestuelle peuvent transcrire les symboles représentés à travers de la langue des signes en texte.



Applications: contrôleurs alternatifs des dispositifs



Applications: des jeux vidéo immersifs

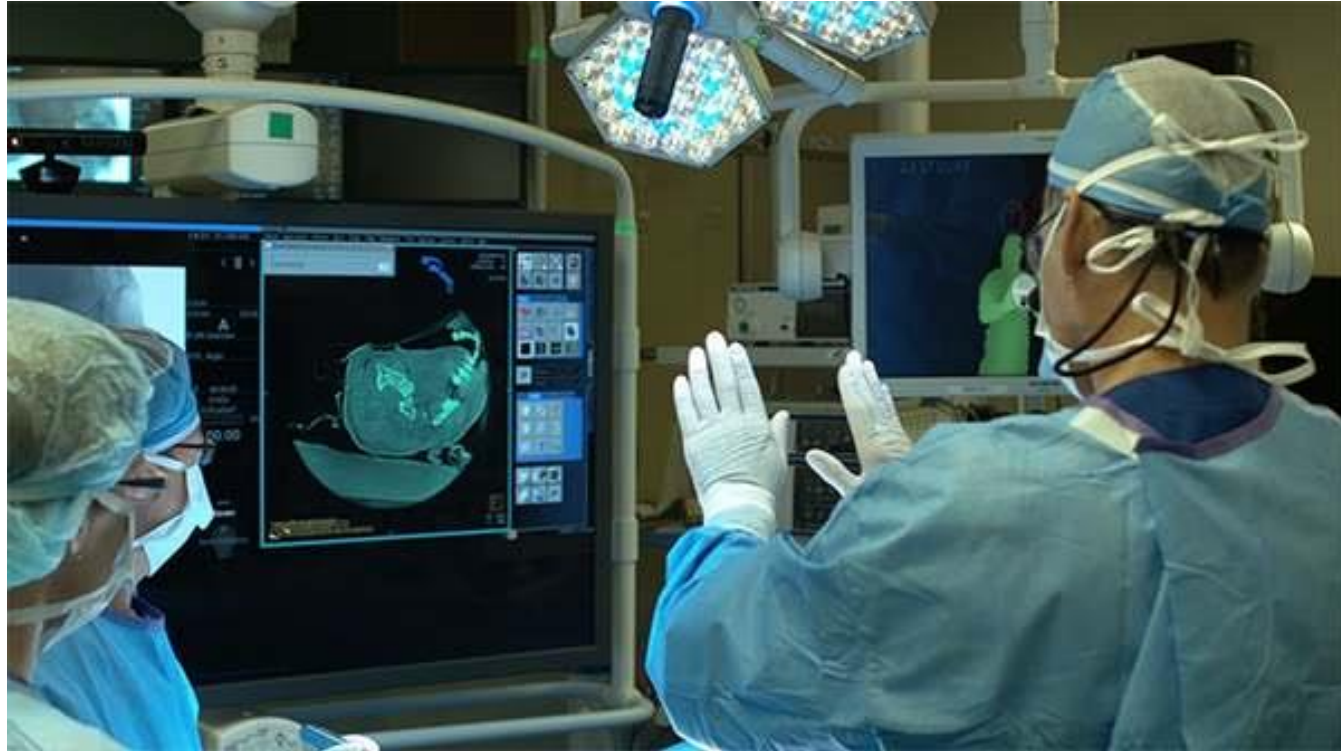


Applications: aide aux personnes handicapées

- Des fauteuils roulants avec des systèmes d'interaction gestuelle.
- Le mouvement de la main est utilisé comme un contrôleur de vitesse ainsi que de direction.



Applications médicales

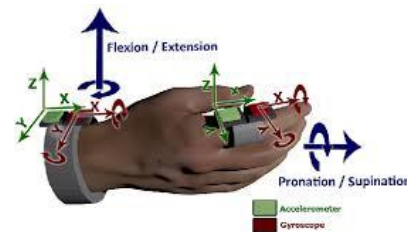
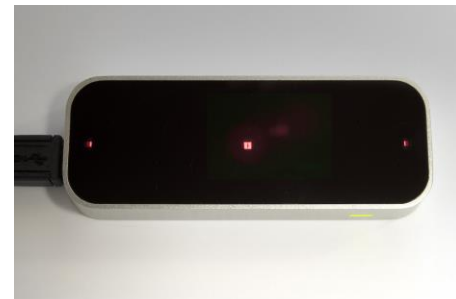


Interaction naturelle avec des robots

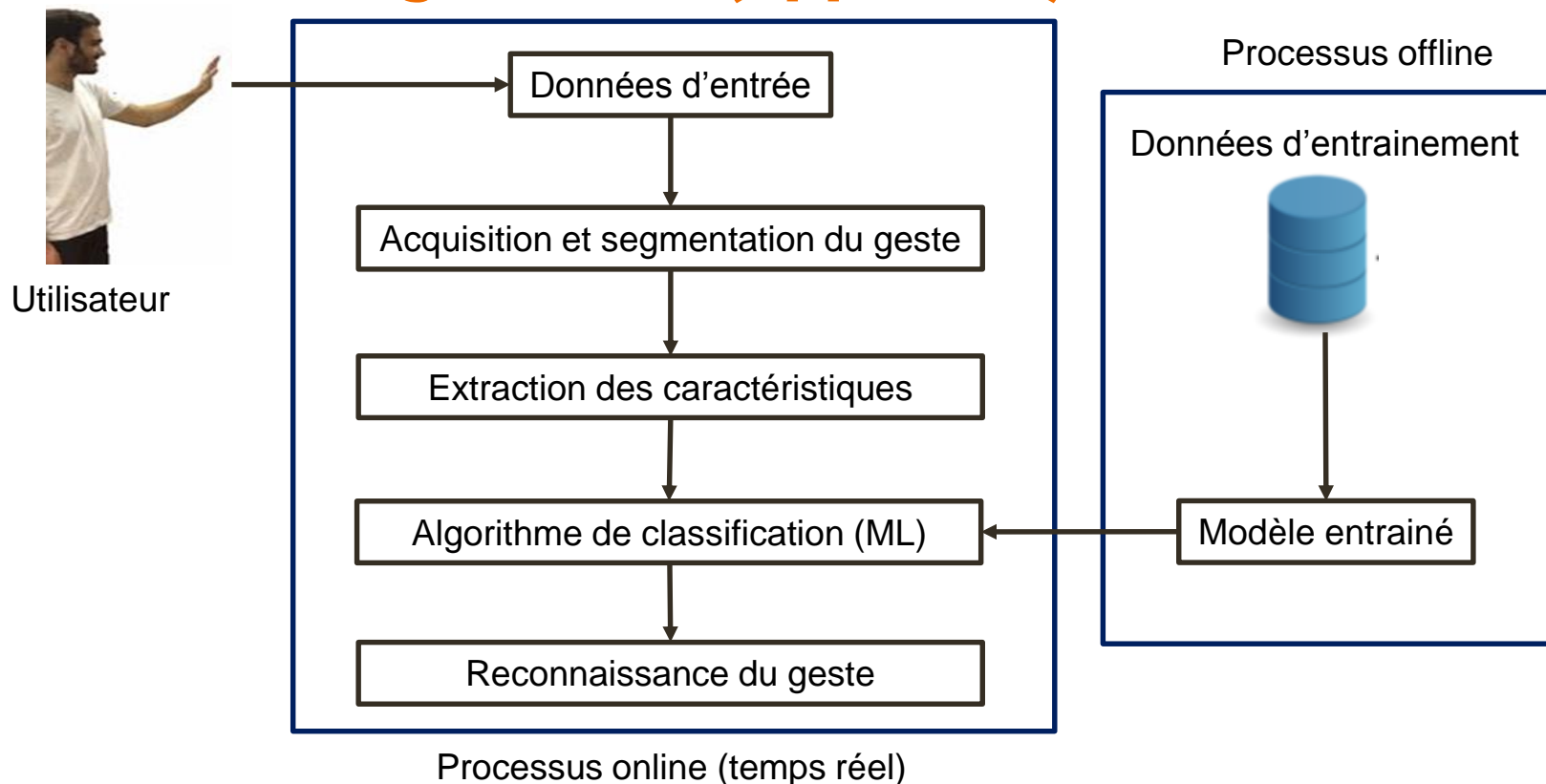


Dispositifs pour l'acquisition des gestes

- Gant de données
- Caméras de profondeur
- Capteurs infrarouges
- Capteurs inertiels
- Caméra monoculaire ou stéréo
- Capteurs musculaires (EMG)
- Systèmes de capture de mouvement



Reconnaissance gestuelle (approche)



Algorithmes de classification pour la reconnaissance gestuelle

- Machine Learning
 - Hidden Markov Models (HMM)
 - Artificial Neural Networks (ANN)
 - K-nearest neighbor (KNN)
 - Naive Bayes (NB)
 - Support Vector Machine (SVM)
 - Decision Tree (DT)
 - Random Forest (RF)
 - Linear Discriminant Analysis (LDA)
- Deep learning:
 - Recurrent Neural Network (RNN)
 - Convolutional Neural Network (CNN)

Des exemples des outils pour la reconnaissance gestuelle

- The Gesture Recognition Toolkit (MIT, 2014)
 - une bibliothèque d'apprentissage machine C ++ multiplateforme et open source conçue pour la reconnaissance des gestes en temps réel.
- OpenCV
 - une bibliothèque open source de vision par ordinateur qui est disponible pour python et C++.
- Social Signal Interpretation (SSI)
 - Analyse et reconnaissance multimodal du comportement non-verbal et des gestes en temps réel.
- TensorFlow
 - un outil open source d'apprentissage automatique développé par Google.

Challenges

- Latence
 - Le traitement de l'image peut être considérablement lent, créant une latence inacceptable pour plusieurs applications qui demandent du temps réel.
- Manque de langage gestuel
 - Différents utilisateurs font des gestes différemment, ce qui rend difficile la reconnaissance des gestes.
- Robustesse
 - Manque de précision en raison des facteurs comme un éclairage insuffisant, des arrière-plan complexes, etc.
- Manque des données
 - Nombre faible de données d'entraînement pour une reconnaissance gestuelle précise.



Initiative d'excellence Arts et Sciences

RAIN



OF

MUSIC



université
de **BORDEAUX**

ESTIA
INSTITUTE OF TECHNOLOGY


Universidad
del País Vasco Euskal Herriko
Unibertsitatea

Rain of Music

- ✓ Un opéra interactif en direct conçu pour être interprété par 4 robots, 6 drones, un chanteur d'opéra et 2 musiciens..
- ✓ La musique sera projetée dans un espace sonore 3D distribué autour et au-dessus de la scène.



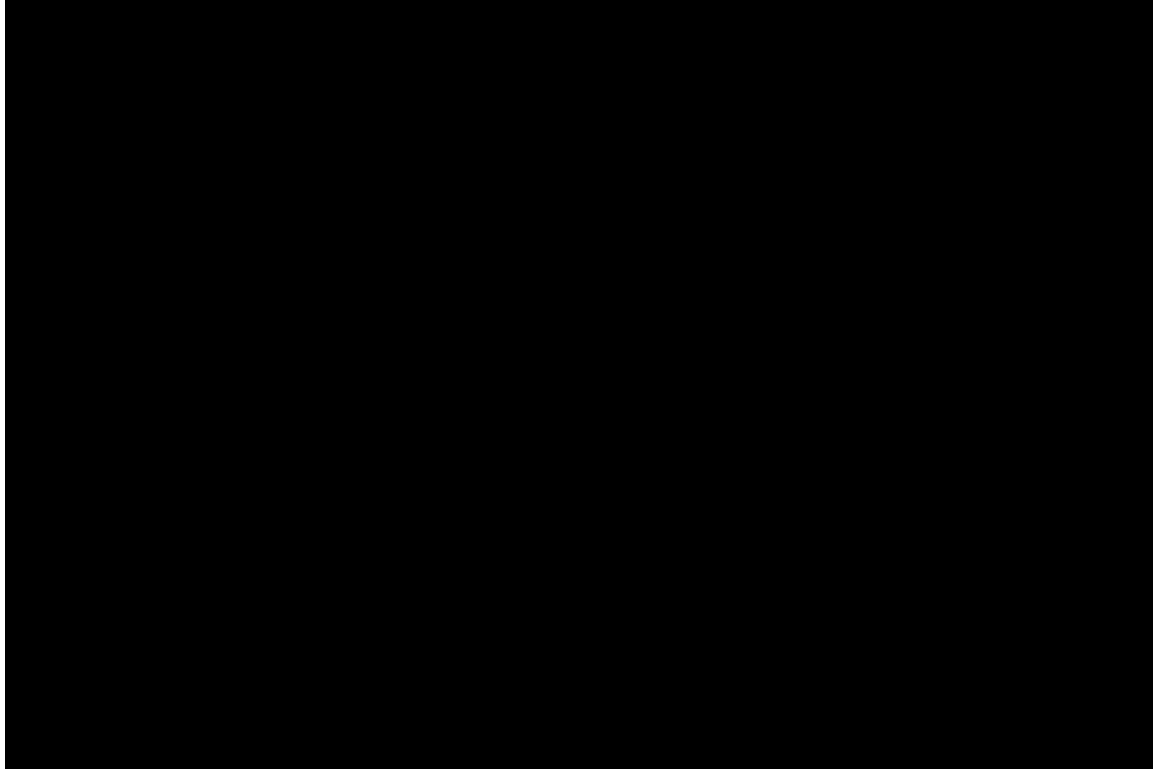
Rain of Music

COMMENT VOLER un essaim de drones?

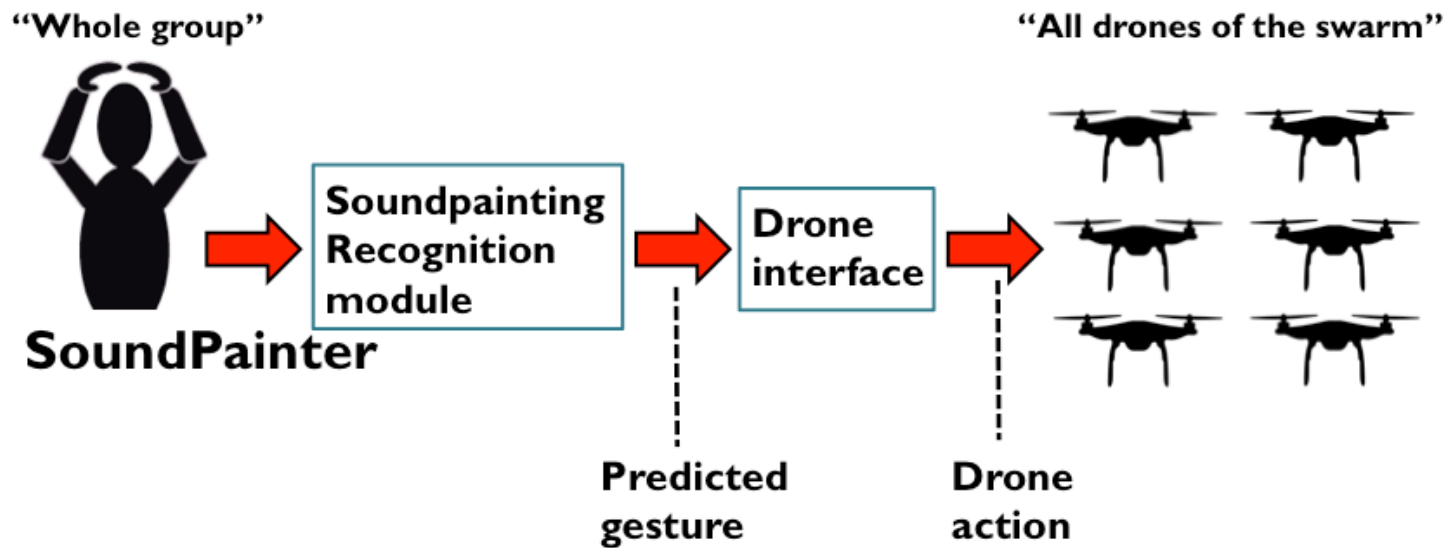
Utilisation du langage **Soundpainting**

$$s = g_{who}^+ \cdot g_{what}^+ \cdot g_{how} \cdot g_{when} \mid g_{who}^+ \cdot g_{what}^+ \cdot g_{when}$$

Langage gestuelle Soundpainting (Walter Thompson, 1974)

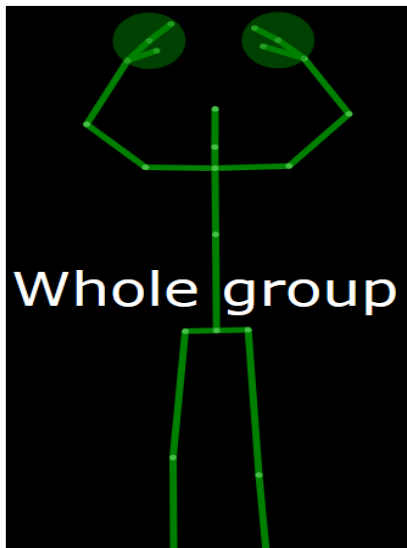


COMMENT VOLER un essaim de drones?

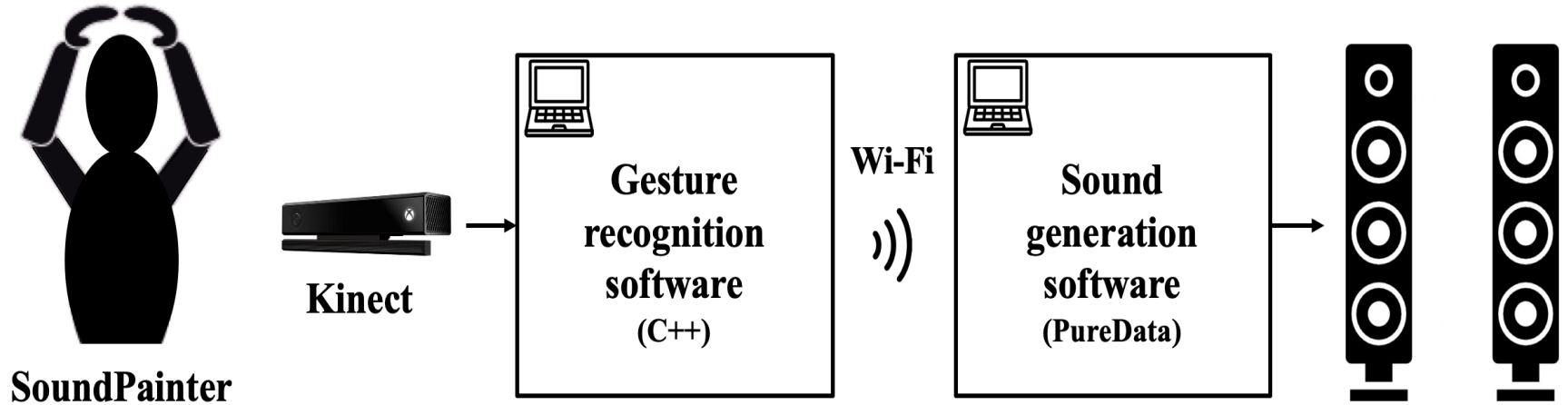


Génération du son avec des gestes

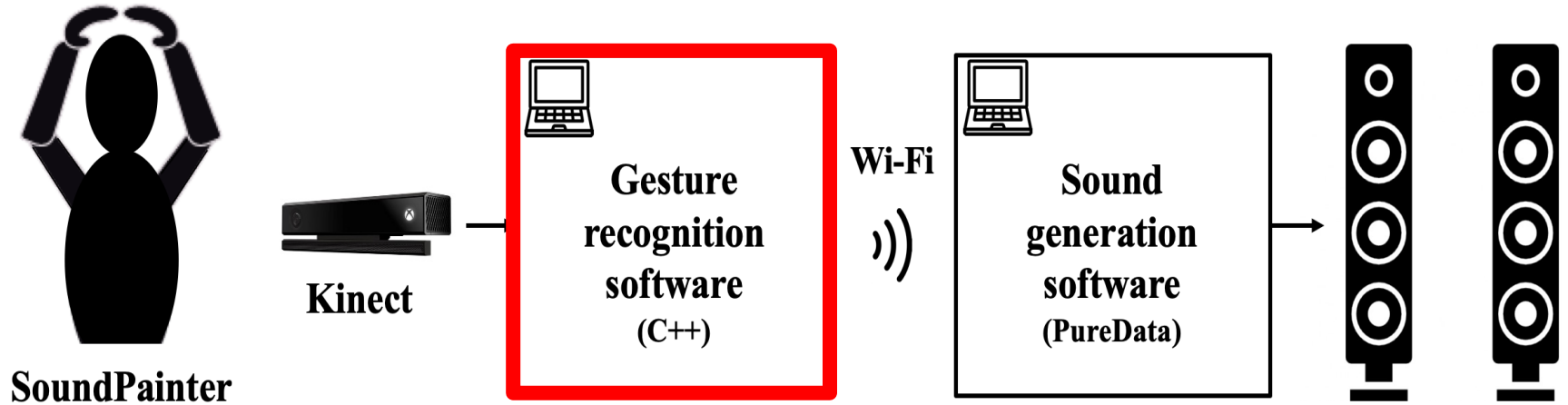
- Un système capable de reconnaître les gestes de **Soundpainting** à partir d'un capteur Kinect pour produire des sons de musique électronique.



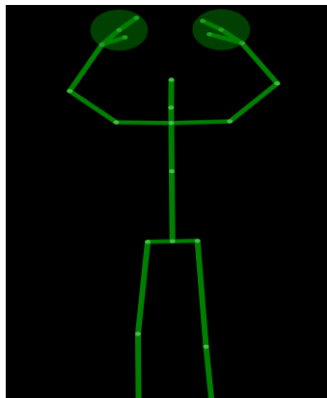
Implémentation du système



Implémentation du système

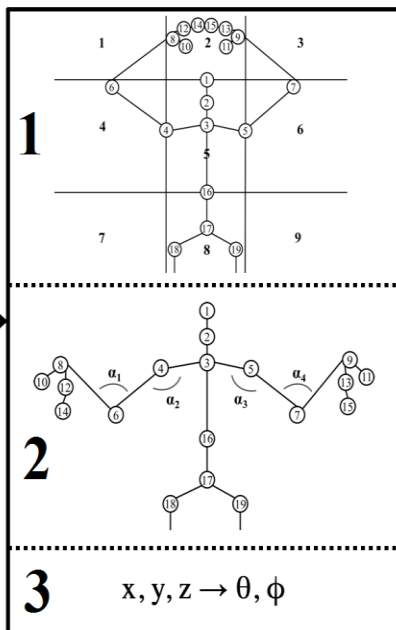


Systeme de reconnaissance gestuelle

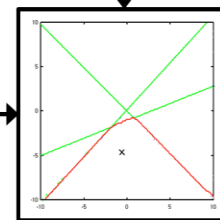
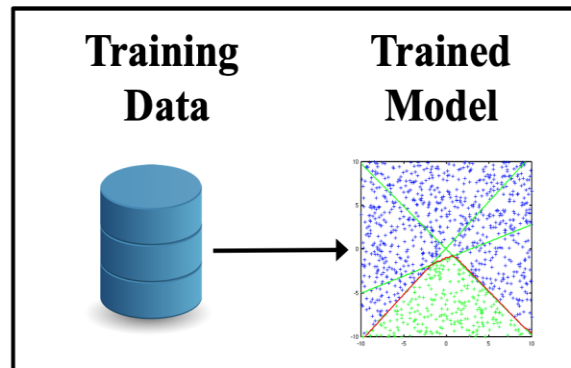


Posture = Cartesian 3D coordinates of 25 articulations

Feature Extraction



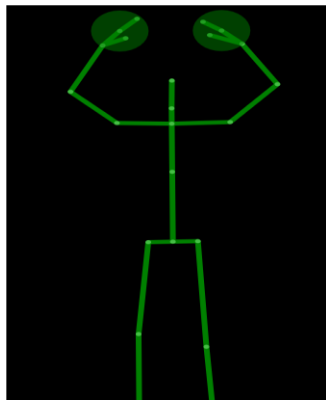
Learning



Classification

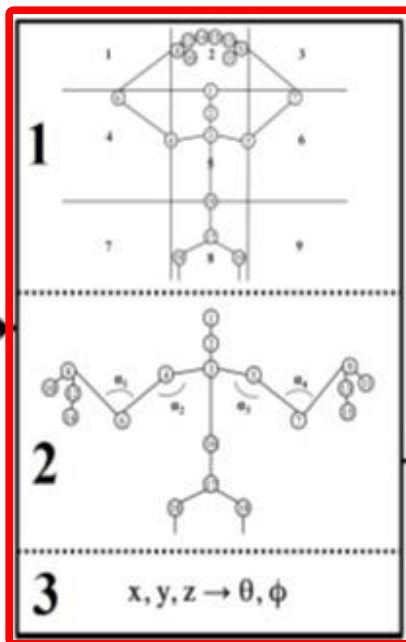
Gesture recognized

Systeme de reconnaissance gestuelle

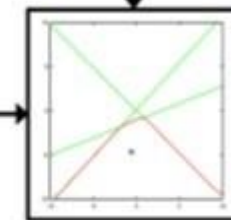
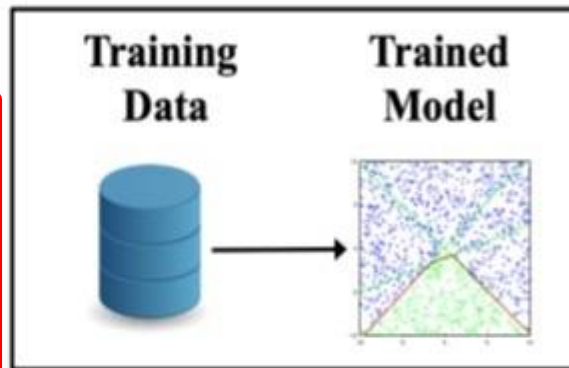


Posture = Cartesian 3D coordinates of 25 articulations

Feature Extraction



Learning



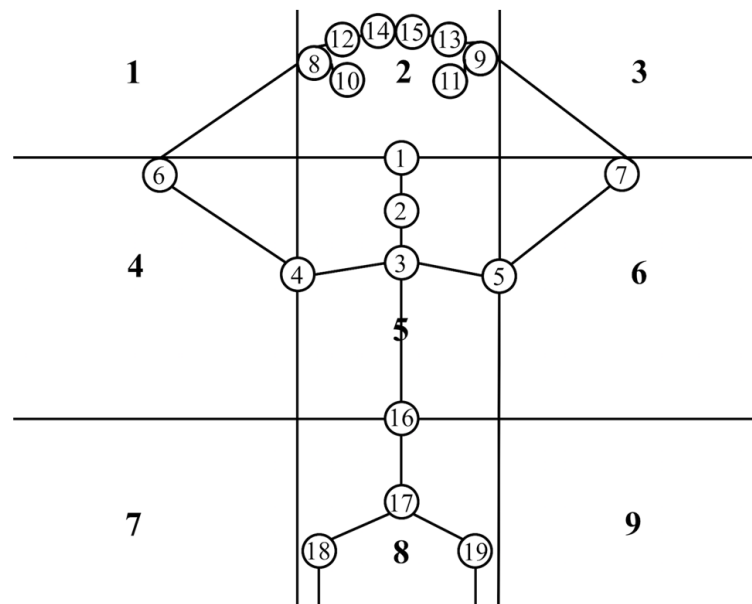
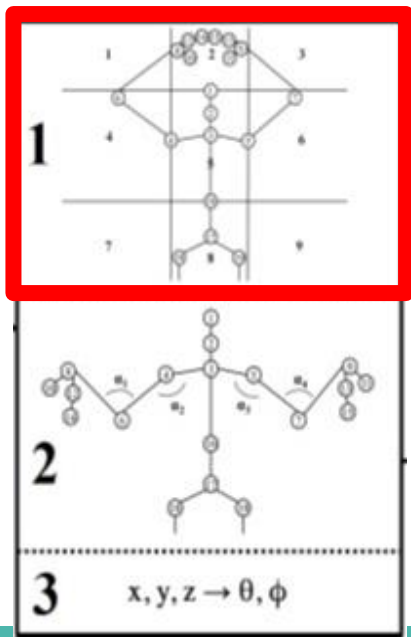
Classification

Gesture recognized!

Premier caractéristique extraite

Positions des poignets

- Les parties du corps qui contiennent les plus de variations.
- Division de l'espace corporel en 9 quadrants définis par 4 lignes.

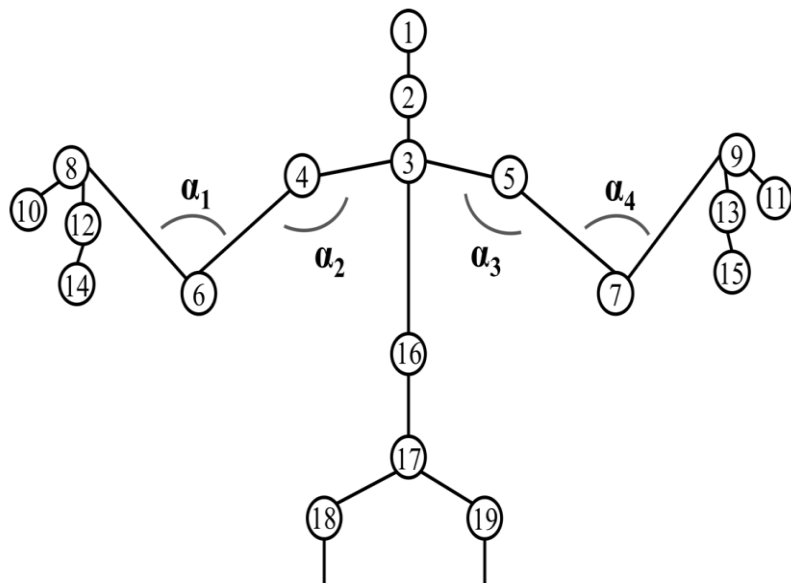


Chan et al. IFAC 2009

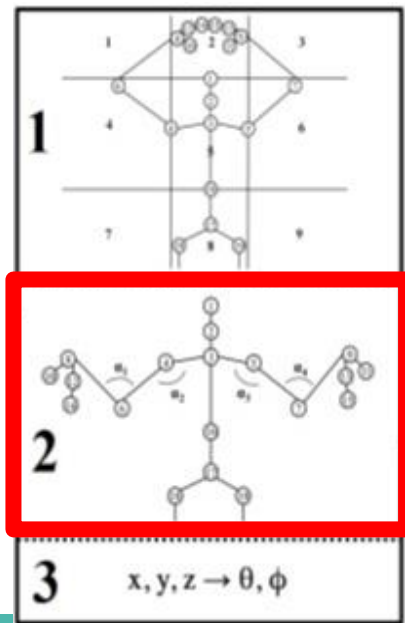
Deuxième caractéristique extraite

- Quatre angles générés par les bras.

- Invariants aux changements d'échelle et à la translation.
- Indépendants de la morphologie.
- Les angles des jambes ne sont pas pris en compte.



Zhang et al. WACV 2017

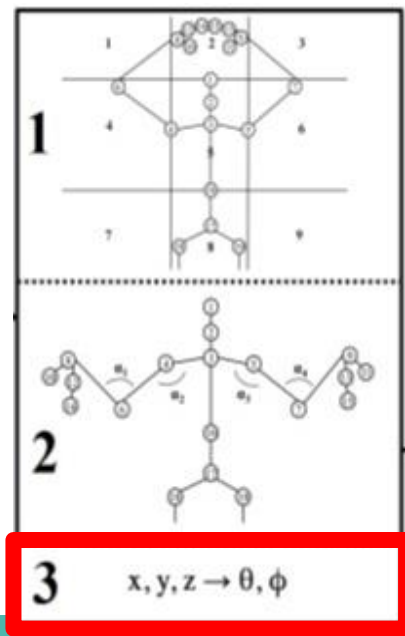


Troisième caractéristique extraite

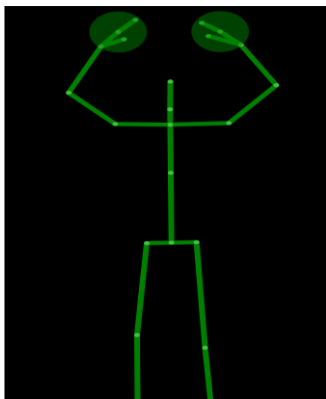
- Conversion des coordonnées cartésiennes en coordonnées sphériques.
 - Ne pas considérer la distance radiale qui dépend de la morphologie.

$$x, y, z \rightarrow \theta, \phi$$

Taha et al. ICIT 2015

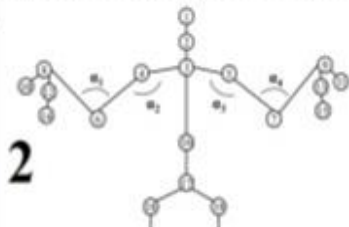
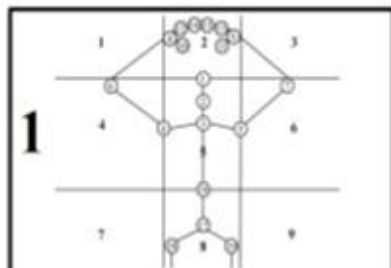


Systeme de reconnaissance gestuelle



Posture = Cartesian 3D coordinates of 25 articulations

Feature Extraction



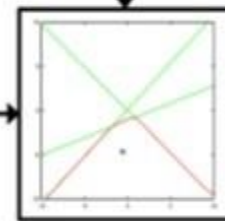
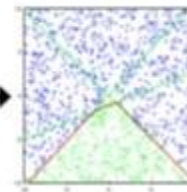
3 $x, y, z \rightarrow \theta, \phi$

Learning

Training Data



Trained Model



Classification

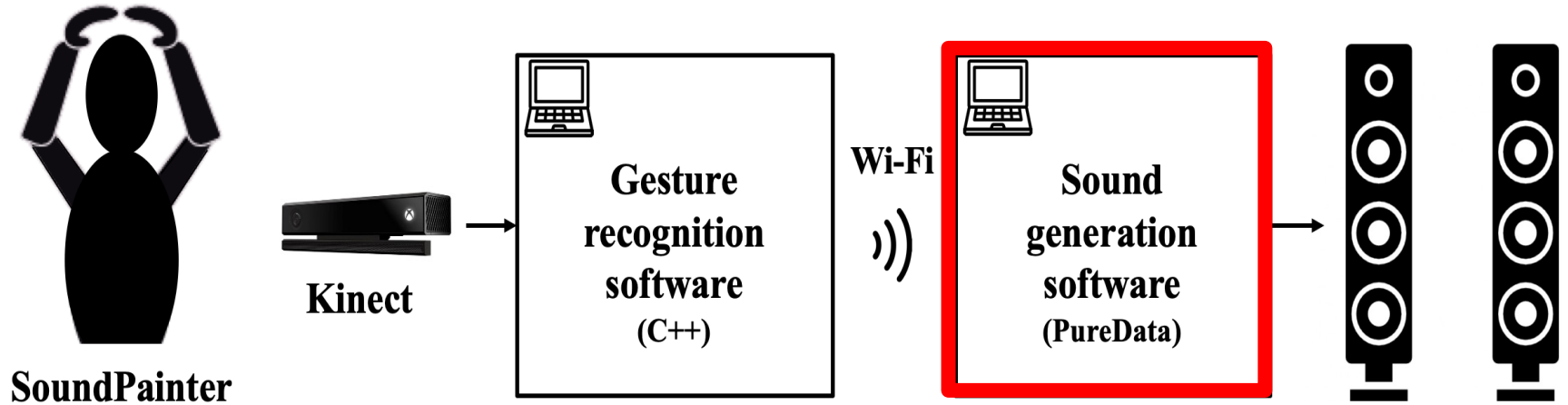
Gesture recognized

Entraînement du modèle

- Une bibliothèque d'apprentissage automatique conçue pour la reconnaissance des gestes en temps réel → Gesture Recognition Toolkit.
- Six gestes de **Soundpainting** (postures clés) ont été appris.
 - ☛ Exécution d'une posture clé devant le capteur Kinect pendant une minute.



Implémentation du système



Logiciel de génération de sons

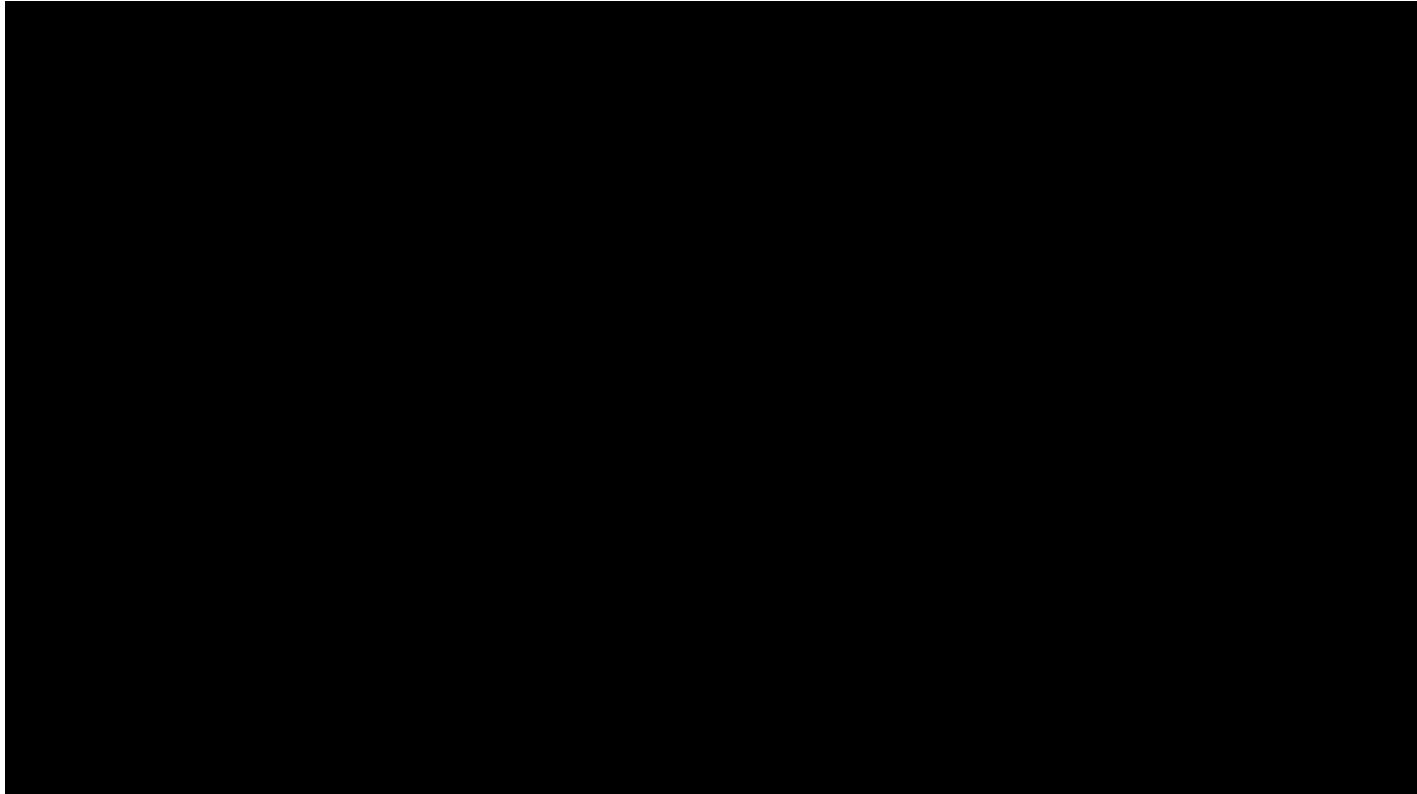
- Le logiciel Pure Data est utilisé pour la génération de musique.
- Chaque geste **Soundpainting** est associé à un code et envoyé au logiciel Pure Data (via une socket).

Gesture	Code	Actions
Whole Group	1	Range of six ascending and descending notes
Rest of the Group	2	Set of simultaneous joint sounds
Wait	3	Increase in volume
Scanning	4	Harmonic series of musical sounds
Silence	5	Decrease in volume
Off	6	Stop sounds

Démonstration artistique publique



Démonstration artistique (Festival Haizebegi)



Perspectives de l'IA pour l'interaction gestuelle

- Des études estiment que l'interaction gestuelle sur le marché devrait augmenter de 27,54% de 2018 à 2023.
- Des applications telles que les systèmes de contrôle à domicile, les systèmes de santé, les technologies de jeu, les automobiles, les téléviseurs, les automatismes domestiques et la robotique devraient pouvoir utiliser l'interaction gestuelle dans les années qui viennent.
- Selon les chercheurs de l'Université de Purdue, les nouveaux avancements de l'IA rendront les systèmes de reconnaissance des gestes beaucoup plus efficaces.

MERCI DE VOTRE ATTENTION!

