

M. Jottrand, A. Moinet, L. Couvreur et T. Dutoit

Faculté Polytechnique de Mons, TCTS Lab, Boulevard Dolez, 31, B-7000 Mons, Belgium

IRMA

Interface de Recherche Multimodale en contenu Audiovisuel

Projet subsidié par la région wallonne

Objectif

Concevoir et développer une interface multimodale innovante, sécurisée et économiquement viable pour la recherche dans une base de données audiovisuelles indexées

• Interface multimodale

- Clavier, souris, voix
- Mouvements des mains, des yeux

• Contenu Audiovisuel

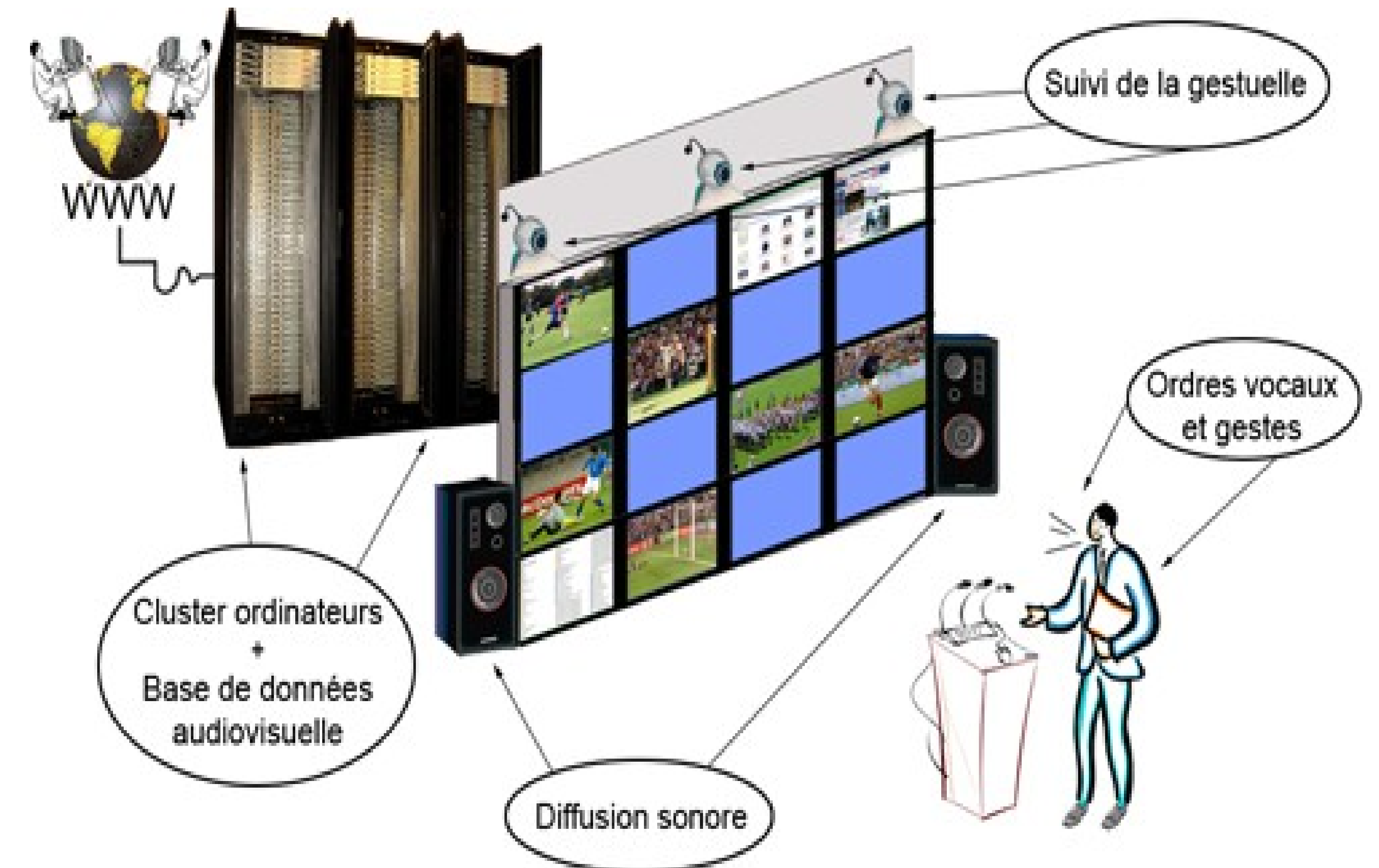
- Archives de broadcasters (journaux télévisés, journaux parlés)
- Vidéosurveillance (séquences de vidéosurveillance)
- Imagerie médicale

• Recherche

- Nécessite la construction d'un index

→ Faciliter la recherche d'une information, d'un document

→ Faciliter l'indexation des documents (indexation semi-automatique)



www.irmaproject.net

Partenaires

DICE (Crypto group)

- Authentification des utilisateurs
- Droits des utilisateurs
- Sécurité des données (DRM), chiffrement du contenu, communications sécurisées



CRID (Centre de Recherches Informatiques et Droit)

- Analyse juridique
- Analyse d'IRMA (pour l'utilisation de données et le respect de la vie privée)



Département image

- Analyse des vidéos (détection de scènes, détection d'événements)
- Intérêt pour la vidéosurveillance



TELE (télécommunication et télédétection)

- Promoteur du projet
- Traitement d'images
- Conception de l'interface



ULg, LENTIC

- Analyse économique d'IRMA et des différents scénarios d'utilisation

Institut de Radioélectricité et Cinématographie

- Archivage audiovisuel
- Contact direct avec l'industrie



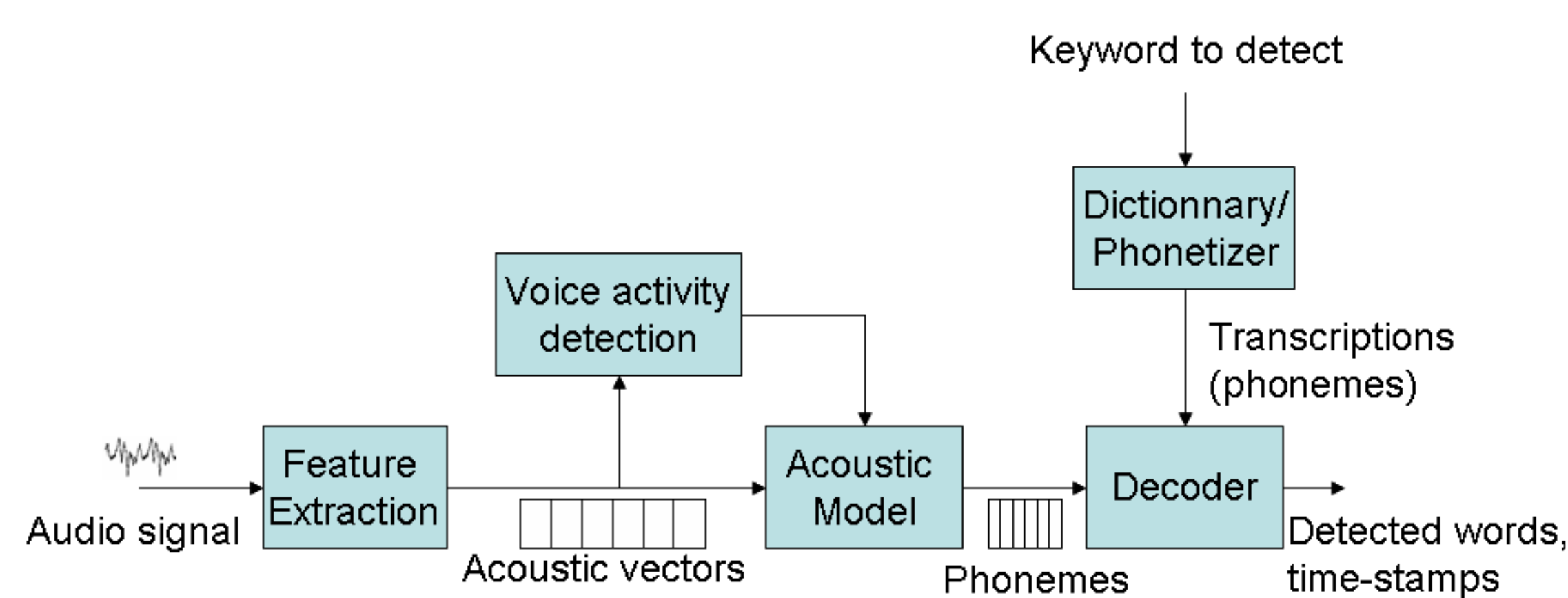
TCTS (speech group)

ASR

- Transcription du contenu audio
- Détection de mots-clés pour l'indexation
- Modalité audio dans l'interface



Indexation semi-automatique



ASR (Reconnaissance automatique de la parole):

- Voice Activity Detection : Supprimer les parties de l'audio inutiles pour la reconnaissance vocale (musique, bruit, parole bruitée)

Réseau de neurones artificiel pour la classification zones utiles/non-utiles

- Keyword spotting sur les zones utiles, utilisation de STRUT (Speech Training and Recognition Unified Tool) un outil développé par TCTS et Multitel

Sélection automatique des mots-clés:

Analyse de textes contemporains des documents audio (vidéo) à indexer (pour les broadcast news, utilisation des flux RSS de sites web de journaux en lignes)

Sélection des mots-clés dans ces textes:

Le choix est guidé par:

- L'information sémantique contenue dans les mots sélectionnés
 - La mesure a priori de la faculté de détection du mot
- Noms propres, mots composés

Création de liens entre les documents textes et les documents audio-visuels

→ possibilité de rechercher d'abord dans le texte, et ensuite de présenter à l'utilisateur les documents audio-visuels liés au document texte trouvé

