

Application graphique pour l'extraction reconnaissance, affichage et prononciation de texte à partir d'images et vidéos complexes

CONTEXTE ET DESCRIPTION DU SUJET

De nos jours, les textes incorporés dans des images et des frames vidéo contenant de grandes quantités d'information sont de plus en plus disponibles dans des bases de données publiques et privées. L'extraction et reconnaissance de ces informations, peuvent être très utiles pour une meilleure compréhension sémantique de ce type de données.

La figure ci-dessous montre une variété des textes rencontrés dans ce cadre. Il est important de noter la différence entre les documents dits multimédias souvent avec arrière plans complexes (textes issus de photographies, de vidéos, de journal télévisé ...) et les documents usuels (livres, articles scannés).

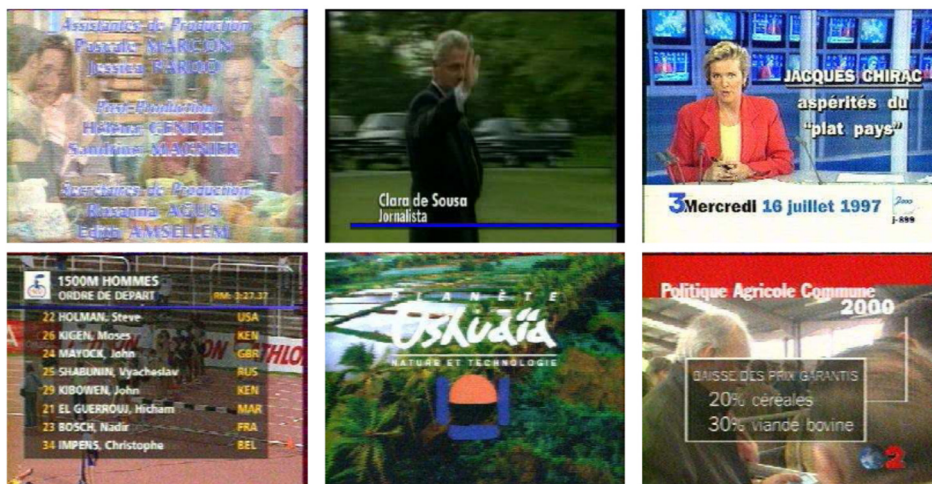


Figure 1. Texte en sur-impression : Images issues des bases de tests de Demarty et Wolf

Dans le cadre des travaux de recherche menés au sein du Laboratoire Informatique Gaspard Monge (LIGM), un système de reconnaissance de texte à partir des documents multimédias est développé. Il se compose d'un bloc de détection et d'extraction de texte basé sur la méthode de correction gamma et un bloc de reconnaissance optique de caractères (OCR). La figure suivante décrit le fonctionnement global du système.

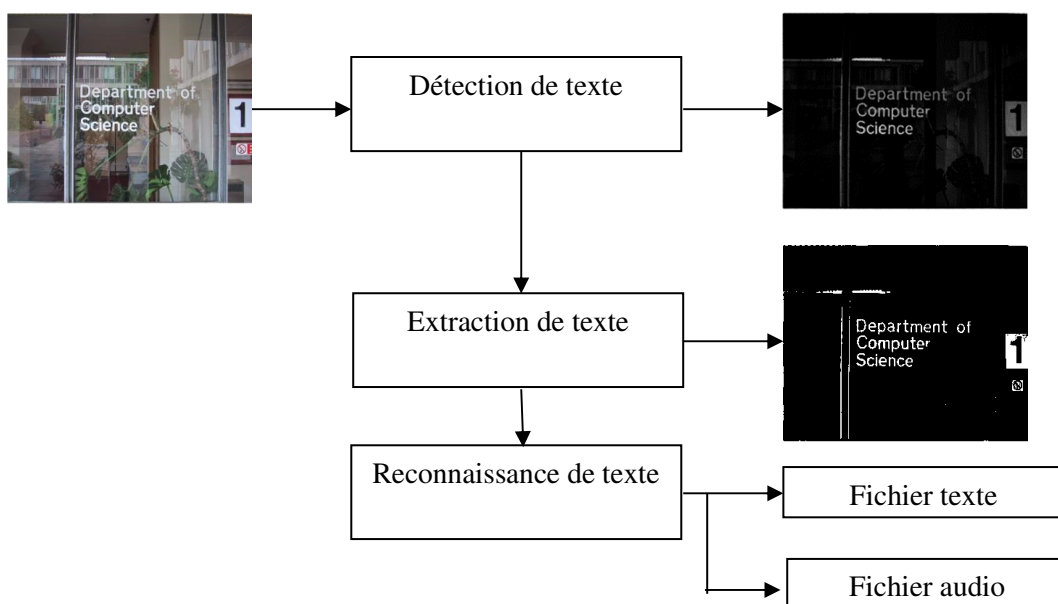


Figure 2. Système de reconnaissance de texte proposé

Dans ce projet on se propose de développer une interface graphique ludique et facile d'utilisation pour l'exécution et l'expérimentation du système de reconnaissance proposé. Les interfaces graphiques reposent sur deux notions clés : les widgets, ou composants graphiques de l'interface, et les événements, générés à la suite d'actions de l'utilisateur.

TRAVAIL A REALISER

Conception, développement, et test d'une interface graphique utilisateur pour le système de reconnaissance proposé (Outils et codes sources seront fournis).

Réalisations :

- Interface utilisateur graphique (GUI) constituée d'éléments :
 - de présentation d'information : fenêtre, ascenseurs, image
 - de saisie d'information : bouton, zone de saisie
 - mixtes (présentation/saisie) : liste déroulante (menu)
- Exécution des différentes fonctionnalités dans le cadre d'une fenêtre graphique
- Déroulement de l'exécution soumis aux actions de l'utilisateur qui peut interagir à tout moment avec l'application par le biais de la souris.
- Possibilités d'affichage très riches : Images des entées sorties de chaque bloc, la meilleure valeur de gamma, texte, ...

Outils de développement :

- Sous Visual Studio C#, C, C++

Contact :

- Rostom KACHOURI : rostom.kachouri@esiee.fr