

# Extraction de la région d'intérêt d'une personne sur un obstacle

Adel Lablack\*, Thierry Urruty\*, Yassine Benabbas\*, Chabane Djeraba\*

\*Université des Sciences et Technologies de Lille 1  
59655 Villeneuve D'Ascq, France  
{adel.lablack,thierry.urruty,yassine.benabbas,chabane.djeraba}@lifl.fr

## 1 Introduction

La direction du regard d'une personne apporte une information utile pour l'accomplissement des tâches visuelles dans plusieurs domaines tels que la réalité virtuelle, les interaction homme-machine, etc. Dans cet article, nous présentons un système qui permet d'extraire l'endroit où se porte l'attention visuelle d'un utilisateur sur un obstacle à partir d'un flux vidéo. Un obstacle (ou scène cible) est une région d'intérêt définie pour être analysée (e.g. un écran plasma large, une image projetée sur un mur, un poster publicitaire, un linéaire dans un magasin, ou la vitrine d'un magasin).

## 2 Méthodologie

Deux facteurs contribuent à la formation du regard d'une personne : l'orientation de la tête et la position des yeux. Nous décrivons brièvement dans ce qui suit la méthodologie utilisée pour extraire la région d'intérêt d'une personne sur un obstacle à partir d'un flux vidéo.

**Estimation de l'orientation de la tête et de la position des yeux** Nous avons utilisé le système proposé par [(Valenti et al., 2009)] afin d'extraire l'orientation de la tête d'un utilisateur, la position de ses yeux ainsi que sa distance par rapport à la caméra. Le système utilise un modèle cylindrique pour l'estimation de l'orientation de la tête et son suivi, et une méthode basée sur les isophotes pour localiser la position des yeux.

**Estimation du champ visuel** Le champ visuel est l'ensemble de l'espace vu par les yeux. Le volume de perception est représenté par la pyramide  $OABCD$  dont le point de départ  $O$  est associé aux yeux (voir Figure 1). La longueur du champ visuel et sa hauteur sont connues en fonction de  $d$  (distance de l'obstacle),  $\alpha$  et  $\beta$  (les angles d'ouvertures horizontale et verticale qui sont égaux à  $120^\circ$  et  $60^\circ$  respectivement selon la vision binoculaire humaine). Les coordonnées des points du volume de perception à une certaine distance  $d$  changent car la personne se déplace (translation) et effectue divers mouvements de la tête (rotations selon les 3 degrés de liberté de la tête). Afin d'adapter le champ visuel à l'orientation de la tête, nous avons utilisé les rotations matricielles selon la convention " $x - y - z$ ".