

Reconnaissance d'Actions par Modélisation du Mouvement

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

*LIFL UMR CNRS 8022, Université de Lille1, TELECOM Lille1
IRCICA, Parc de la Haute Borne, 56950 Villeneuve d'Ascq
{yassine.benabbas,adel.lablack,thierry.urruty,chabane.djeraba}@lifl.fr

Résumé. Cet article propose une approche utilisant les modèles de direction et de magnitude de mouvement pour détecter les actions qui sont effectuées par des êtres humains dans des séquences vidéo. Des mélanges Gaussiens et de lois de von Mises sont estimés à partir des orientations et des magnitudes des vecteurs du flux optique calculés pour chaque bloc de la scène. Les paramètres de ces modèles sont estimés grâce à un algorithme d'apprentissage en ligne. Les actions sont reconnues grâce à une mesure qui se base sur la distance de Bhattacharyya et qui permet de comparer le modèle d'une séquence donnée avec les modèles créés à partir de séquences d'apprentissage. L'approche proposée est évaluée sur deux ensembles de vidéos contenant des actions variées exécutées aussi bien dans des environnements intérieur qu'extérieur.

1 Introduction

La reconnaissance des actions est un sujet particulièrement complexe dans le domaine de la vision par ordinateur. Cela consiste en la classification automatique des actions ou des activités réalisées par un individu dans une séquence vidéo. La reconnaissance des actions est cruciale dans de nombreux domaines comme la vidéo-surveillance, l'interaction homme-machine et l'indexation des vidéos.

L'objectif d'un système de reconnaissance d'actions est de reconnaître des actions simples de la vie courante dans une vidéo (ex : marcher, répondre au téléphone, sauter...) à partir de vidéos de référence. Ces actions répondent à des modèles de mouvements simples effectués par une seule et même personne durant un laps de temps court. Certaines approches détectent les actions à partir d'images fixes, tandis que d'autres ont recours à des vidéos stéréoscopiques ou à des maillages 3D (Ganesh et Bajcsy, 2008). Dans cet article, nous traitons les séquences vidéo enregistrées par des caméras monoculaires car elles permettent de détecter des actions en combinant des informations spatiales et temporelles (Johansson et al., 1994). Cet intérêt pour les vidéos monoculaires résulte du fait qu'elles sont couramment utilisées, moins gourmandes en ressources et plus économiques.

Cet article présente une méthodologie qui permet de reconnaître des actions en se basant sur l'analyse du mouvement des sujets. Il est organisé comme suit, les travaux antérieurs sont passés en revue dans la Section 2. Nous décrivons ensuite notre approche dans la Section 3 en détaillant ses deux phases principales qui sont la *construction de modèles* et la *reconnaissance*