Gesture based human-computer interface for 3D design

Vincent Spruyt

Supervisor(s): Wilfried Philips, Alessandro Ledda

I. Introduction

Hand tracking, hand segmentation and hand modeling are amongst the most important fields of interest in current computer vision research. However, traditional hand recognition systems can only operate in constrained environments using coloured gloves or static backgrounds and do not allow for 3D object manipulation. The goal of this research is to develop real-time camera based solutions to control 3D modeling applications using natural hand gestures.

II. Hand tracking

Deterministic tracking systems such as the Camshift algorithm are unable to cope with the uncertainty caused by noise or changing lighting conditions in real-life video. Therefore, Bayesian filtering techniques such as the particle filter have been gaining attention for the past several years. Several theoretical improvements to the traditional particle filter have been proposed lately, such as auxiliary particle filtering or unscented particle filtering. This research aims to integrate these improvements with a novel observation model to obtain a robust 2D hand tracker, thereby maintaining its real-time characteristics.

An important aspect in developing such an observation model, is colour constancy. To allow for lighting invariant tracking, an algorithm is being developed which is capable of maintaining colour constancy in dynamic lighting environments. Furthermore, features such as optical flow and texture descriptors are merged with the probabilistic colour model to obtain an accurate and robust representation of the human hand.

III. Hand segmentation

Based on the hand location and bounding box as estimated by the particle filter, the hand can be segmented fairly accurate using level sets or active contour techniques. Such an approach allows for the incorporation of prior knowledge about the hand shape. Based on the segmented hand, several simple feature detectors can be used to obtain a rough description of finger position and hand orientation.

IV. 3D Hand modeling

The raw pixels of the region of interest resulting from the segmentation phase, together with the detected fingertip features, can be used to obtain a relative depth estimate by first training a Markov Random Field (MRF) using a training set containing both hand images and depth images obtained with a range scanner. Once the relative depth of the fingers and handpalm is known, a 3D hand model can be fitted to this data resulting in a 3D hand tracker.

V. Conclusions

This research aims to develop a solution for 3D hand tracking based on monocular video resulting from a standard low resolution webcam. The resulting algorithm should be lighting invariant and should be able to operate in dynamic environments.

V. Spruyt is with the Department of Telecommunication and Information Processing