

Abstract

The present thesis deals with the study of vision techniques for the detection of human pose based on the analysis of a single image, as well as the tracking of these poses along a sequence of images.

It is proposed to model the human pose by four kinematic chains that model the four articulated extremities. These kinematic chains and head remain attached to the body. The four kinematic chains are composed by three keypoints. Therefore, the model initially has a total of 14 parts.

All methods proposed in this work are implemented, validated and analyzed using the public CAD60 dataset, dataset with images in RGB and depth, and other dataset created expressly to complement the CAD60 dataset.

In this thesis it is proposed to modify the technique called Deformable Parts Model (DPM), adding the depth channel. Initially, the DPM model was defined over three RGB channel images. While in this thesis it is proposed to work on images of four RGBD channels, so the proposed extension is called 4D-DPM. The experiments performed with 4D-DPM demonstrate an improvement in the accuracy of pose detection with respect to the initial DPM model, at the cost of increasing its computational cost when treating an additional channel.

On the other hand, it is defined to reduce the previous computational cost by simplifying the model that defines the human pose. The idea is to reduce the number of variables to be detected with the 4D-DPM model, so that the suppressed variables can be calculated from the detected variables using inverse kinematics models based on dual quaternions. The experiments show that the

combination of these two techniques allows, by reducing the computational cost of the original DPM method, to improve the accuracy of the pose detection due to the extra depth channel information.

In addition, it is proposed to use a particle filter models to continue improving the accuracy of detection of human poses along a sequence of images.

Considering the problem of detection and monitoring of human body pose along a video sequence, this thesis proposes the use of the following method.

1. Camara calibration. RGBD image processing. Subtraction of the image background with the MSER method.
2. 4D-DPM: method used to detect the keypoints (variables of the pose model) within an image.
3. Particle filters: this type of filter is designed to track the keypoints over time and correct the data obtained by the sensor.
4. Inverse kinematic modeling: the control of kinematic chains is performed with the help of dual cuaternions in order to obtain the complete pose model of the human body.

The overall contribution of this thesis is the proposal of the previous method that, combining the previous methods, is able to improve the accuracy in the detection and the follow up of the human body pose in a video sequence, also reducing its computational cost .

This is possible due to the combination of the 4D-DPM method with the use of inverse kinematics techniques. The original DPM method should detect 14 point of interest on an RGB image to estimate the human pose. However, the proposed method, where a point of interest for each limb is removed, must detect 10 point of interest on an RGBD image. Subsequently, the eliminated 4 point of interest are calculated by using inverse kinematics methods from the calculated 10 point of interest.

To solve the problem of inverse kinematics a dual quaternions methods is proposed for each of the 4 kinematic chains that model the extremities of the skeleton of the human body.

The particle filter is applied over the time sequence of the 10 points of interest of the posture model detected through the 4D-DPM method. To design these particle filters it is proposed to add the following restrictions to weight the particles generated:

-
1. Restrictions on joint limits: The human pose is modeled with a set of open kinematic chains. In such a way that the points of interest are the joint articulation variables of each of the kinematic chains. Each of these variables has a restricted movement in a given range.
 2. Softness restrictions: it is proposed to weight the particles inversely proportional to the distance of the particle generated with the solution at the previous time instant.
 3. Collision detection: the geometric modeling used to model the skeleton of the human body is a set of poly-spheres because they allow us to perform collision detection between body elements very efficiently. It is proposed that the particle filter does not generate particles in which there are collisions impossible between these elements.
 4. Projection of poly-spheres: it is proposed to weight each particle generated directly proportional to the overlap of the projection of the poly-sphere model that defines this particle with some plane of the RGBD image.