Pose estimation camera localization for multi cameras

Technology

Pose estimation camera localization for multi cameras has been investigated for many years. Mainly, researches have concentrated on camera localization in multi camera systems, because of the constraint of cameras with low accuracy and limited field of view. Researchers are mainly concentrated on single camera localization to decrease inaccuracies and improve robustness of algorithms.

In previous work, Hanno Jaspers et al proposed a camera localization method for the multi camera systems to calibrate the intrinsic and extrinsic matrices of two data sets, Herz-Jesus-P8 and Fountain-P11. They achieved good result with 12cm error in a total distance of 15. 5m and the position error is a less than 6cm in overall distance of 11m. In their method, scale factor of camera is calculated using reference objects. Firstly, the urgent parameters, width and height of the object and the distance between camera and reference object, need be clear to compute scale factor and avoid scale ambiguity. The disadvantages of this method to calculate the important features of objects takes more time, human participates in this process and there must be an object with planar surface. Even, the accuracy is very low with low resolution cameras.

Yingen Xiong et al proposed a multiple camera calibration for analyzing human body poses, hand gestures and motion and head orientations in meeting rooms. In this approach, a box with dots and Vicon markers are exploited to calibrate the cameras. A good accuracy of a mere than 1 cm to camera locations of most cameras is attained.

Alvaro Collet et al proposed efficiently recognizing all objects in a scene and estimation full of cameras from multiple views based on the art single-view method. Initially, the single-view algorithm is executed for each images and RANSAC calculates the presence of thousands of corresponding points. Optimizing of the object pose is performed using a reduced generalized image comprising of points consistent across every images. Results of this method is that a modeling error of 1mm in the measurement of a coke can and takes increase to a depth estimation error of up to 3cm at a distance of 1m from camera. However, the distance is spacious enough to cause problems to a robotic manipulator attempting to grasp the object.

Aslan et al proposed a method to automatically calibrate the extrinsic parameters of multi cameras. They exploited the highest detected point of moving people to calibrate features. After the relative pose estimation for each camera pair, the resulting camera network is built up using a global error minimization technique. In different scenarios extrinsic parameters are performed, achieving a projection error of mere than 6px and a triangulation error of markers in the scene of about 5cm.

Sunglok Choi et al presented an asymmetric Gaussian kernel to estimate unknown scale factor accurately. The unknown scale factor is derived by comparing same point in the real world and the reconstructed world. In their method the asymmetric kernel almost achieved twice higher accuracy than the previous Gaussian kernel.