Visual SLAM for humanoid robot using intrinsic scale

Positionnement interne

Ce sujet de thèse s'insère dans le cadre de l'ARC INRIA “H-Vis : Vision, navigation et action intentionnelles des robots humanoïdes” dans laquelle le projet e-Motion s'est engagé, en collaboration avec le groupe Gepetto (LAAS-CNRS) et les projets INRIA Bipop, Bunraku et Lagadic.

Environment :

- **Project Team :** e-Motion (project leader Christian Laugier)
- **Research axes :**

The goal of e-Motion is to develop geometric and Bayesian models for autonomous navigation in open and dynamic environment, i.e. in an environment observed by different sensors, in perpetual evolution. The team deal with the multimodal perception of dynamic world, reconstruction, localization and motion planning in such an environment.

- **Phd direction :** Chritian Laugier (DR) and Agostino Martinelli (CR)
- **Place :** INRIA Rhônes-Alpes
- **Contact :**
  - e-mail : christian.laugier@inrialpes.fr
  - phone : +33 4 7661 5222
  - fax : +33 4 7661 5477

Subject description :

This subject takes place in the context of Humanoid Robot navigation in an open and dynamic environment. To operate in such an environment, the robot needs to localize correctly in this world and to modelize the world in order to avoid obstacles.

The currents approaches consist in simultaneously resolving the localization problem and the mapping problem (SLAM), the “eyes” position and the position of 3D objects are jointly estimated from visual features extracted in the images. In [1][2], the camera position is tracked using a map build with natural visual landmarks. These landmarks are generally detected with a corner detector, like the Shi and Tomasi detector [3], and represented as a 3D vector point in the 3D world. As camera sensor only measure the object's bearing, a landmark has to be seen at least two times from different points to be correctly estimated. To deal with this problem, [4] propose to use an inverse depth representation in order to increase the linearity of the system and make it possible to initialized the landmark position at the first observation. However, even with this representation, the landmark depths can't be accurately estimated if its parallax angle stay small (for example if the landmark is in front of the camera and the camera moves forward or backward. This point is really annoying if we want to avoid obstacles in front of the robot. To resolve this problem, we believe that a scale information associate to each landmark can improve the depth estimation. Indeed, as shown in [5] the observe object size in a camera is related to the object depth and the object size in the real world. If we use a SIFT detector [6] or a Laplacian scale-space based detector, it is possible to obtain an intrinsic scale for each landmark detection.

The work of this thesis will so consist in introducing a scale information in a visual SLAM

framework in order to improve landmarks and camera localization. Resulting algorithms should be also usable by humanoid robot in open environment, that's mean that dynamical models would be necessary.

Required skills and knowledges

This proposal is limited to candidates having a master's in computer science. A minimum knowledge in the following points is required :

- Bayesian and probabilistic methods used in robotics
- Computer vision and image processing
- C++ programming

Bibliography

[1]: A. Davison, I.Reid, N. Molton, and O. Stasse, "Monoslam : Real-time single camera slam",IEEE Transaction on Pattern Analysis and Machine, 2007
[2]: E. Eade and T. Drummond, "Scalable Monocular SLAM",IEEE Conf. Computer Vision and Pattern Recognition, 2006
[3]: J. Shi and C. Tomasi, "Good features to track",IEEE Conf. Computer Vision and Pattern Recognition, 1994
[4]: J.M.M Montiel, J. Civera and A. J. Davison, "Unified Inverse Depth Parametrization for Monocular SLAM",Proc. Robotics Science and Systems, 2006
[5]: A. Nègre, C. Braillon, J. Crowley and C. Laugier, "Real-time Time-To-Collision from variation of Intrinsic Scale",Proc. of the Int. Symp. on Experimental Robotics, 2006
[6]: D. G. Lowe, "Object Recognition from Scale-Invariant Features",Proc. Int. Conf. Computer Vision, 1999