Part 8

Vision-Based Navigation
Session Summary

Yoshiaki Shirai

Department of Computer-Controlled Machine Systems, Osaka University

Vision based navigation can roughly be divided into the following two classes:

- Navigation without maps.
- Navigation with maps.

Navigation without map requires two subtasks: map building, and motion planning. The former has been intensively studied for a few decades. A typical way is so called SLAM (simultaneous localization and map building), which uses dead reckoning and visual observations (see papers in another session).

Motion planning is necessary to reach a goal. An important problem is how the goal is specified. The goal can be pointed if it is seen from the starting position. If not, the goal can be specified using a rough topological map or specified in terms of properties of the goal. Another way is to manually guide a robot to the goal.

Navigation with a map is dealt with in this session. A problem is how to find correspondence between a given map and the environment which is observed by vision. For this purpose, prominent features of the environment are observed. Those features are called landmarks.

The navigation with a map can further be divided into the following classes according to available landmarks:

(a) All landmarks are observed everywhere.
(b) Only a part of landmarks are observed at a position.
(c) No landmarks are given.

In case (a), the problem of finding landmarks is easy because they can be tracked. Such landmarks are equivalent to beacons emitted from known positions. Even if many landmarks are observable, three of them is enough to determine the position in the horizontal plane.

Peter Corke proposed a method similar to visual servoing because visual servoing is more efficient than solving geometrical problems.

The discussions on his presentation are:

Thorpe: A proof of convergence should not be too hard.
Dickmanns: The possibility of dynamic feature selection to optimize a Jacobian had been previously explored by Wunsch.
Nebot: When moving between local navigation regions some sort of landmark switching strategy would be required.

In case (b), visible landmarks changes as a robot moves. The important problems are what kind of landmarks should be used and how to find landmarks at each position.

Santos-Victor addresses those problems. He proposes the use of a topological map of touristic landmarks. Given a goal, the robot moves along a road or a corridor...
while looking for the predicted landmarks. This is more effective than making a geometric representation of the environment. He also proposes to take the advantage of biological vision systems which are proved to be successful in different navigation tasks.

The discussion on his presentation are:
Chatila: The approach of using image motion (optic flow) for the control of a mobile robot is interesting from the viewpoint of demonstrating the biological plausibility of such models. However it seems to require massive modification to the environment such that their practical use is limited.
Answer: One of the examples shown may have lead to that impression since the environment had been modified to have highly textured artificial patterns. However, the reason is that this particular example has about 8 years old when the available computational power would require strong textures. We can currently apply the same techniques with higher resolution images at video rate. Examples were shown for the underwater robot control or for the binocular head.
Shirai: How important is the log-polar transformation? For example, when navigating in a corridor, many distant landmarks are useful.
Answer: Indeed, many distant landmarks are useful in the context of navigation. The omni-directional camera still suffers from perspective effects, so that distant objects will always remain smaller in the image. The advantage of the panoramic image is the uniform angular resolution. Related to the question, it is interesting to note that the optimization of the sensor radial distribution versus the mirror profile has some redundancy. For instance, one could use a hyperbolic mirror and determine the sensor distribution or do a joint optimization of both the sensor layout and mirror profile.

Featherstone: In the human retina, the photoreceptor in the foveal region have different characteristics of those located in the periphery, in terms of sensitivity to light and color. Have these aspects been taken into account in the log-polar sensor?
Answer: Indeed, there are two types of photoreceptor: cones and rods. Cones provide color vision and rods provides gray vision under low illumination. In the human eye, there are very few cones in the periphery of the retina. The central foveal is densely populated with cones. These aspects are not covered in the log-polar sensor.
Zelinsky: There have been other ways of producing image geometries similar to the log-polar by using mirrors instead of a special sensor.
Answer: Other ways of producing space variant resolution images is to use lenses, although there are some difficulties with severe brightness changes across the entire image. There have also been other efforts of building mirrors such that the panoramic images have some desirable properties: constant resolution. However, only a polar sensor can realize the equal angular resolution at no additional computational cost.

In some biological systems such as that of bees, the goal position is specified by other bees. The task is to fly toward the goal direction and to land at the goal. Srinivasan already presented a theory that utilizing information on optic flow, a bee can fly in the center of a narrow channel.
This time he investigates how to control the flying velocity for a smooth landing. The theory is similar to the previous one: a smooth landing is realized by keeping optic flow constant while approaching to the goal. This is a simple and beautiful theory which is proved by experiments. Because of its simplicity, the landing method may be applied to controlling flying vehicles.

On the other hand, Hartley deals with a problem of measuring the angle between two lines in a 3-D space. He proposes a mathematical method to precisely measure the angle by utilizing a horizon or an estimated horizon. This is effective for finding landmarks of artificial objects because their image usually have straight lines which meet a fixed angle.

The discussion on his presentation are:

Shirai: Since you have the horizon, is it not possible simply to compute a rectified image of the plane by synthetically rotating the view direction?
Answer: Yes. There are other ways of carrying out this task. One of the ways is to compute 2-dimensional projective transformation (homography) that will rectify the plane to a geometrically-correct Euclidean view. In doing this the difficulty is that one either needs to know the camera calibration accurately, or else to carry out a full 8-degree of freedom homography computation. In our method, we do not need to know a full calibration of the camera. Computation of the conformal point is simpler and more stable than computing the principal point and focal length of the camera. Computing a full 8-degree of freedom homography is not an ideal solution, either, since the problem of over-fitting the transformation arises. Knowing the conformal point, the correct homography relies on only 3 parameters.

Jarvis: It seems that you have invented a new method of doing odometry. Could you indicate some of the advantages of this method with respect to other methods?
Answer: One of the major advantages is speed; another is that it does not need a full camera calibration. There are two factors that affect speed – the speed of eliminating false point tracks on the ground and the rapid method of computing the rotation angle of the vehicle. Errors in tracking and determination of which points are on the ground plane are problems with a feature-based approach. Our method can efficiently eliminate false tracks, and tracks corresponding to objects not on the ground. The motion model for points (used in the RANSAC procedure) needs only two points to compute. RANSAC complexity increases exponentially in the number of points needed to instantiate the model (4 points for a full homography). Further, the computation of the rotation angle is very simple and by eliminating points not on the ground plane, we avoid the need for ground segmentation.