Learning an approximate map of the environment by unsupervised bimodal landmark exploration

Lambert Schomaker and Rudolf Fehrmann
AI Institute, Groningen University, The Netherlands
`schomaker@ai.rug.nl`

1. Problem

There exist several good technical solutions for robotic navigation. However, biological navigation methods are often very robust and do not require auxiliary hardware such as GPS or radiographic beacons. Therefore, it remains a challenge to understand and possibly exploit mechanisms known from biological navigation.

Biological navigation

- gradient based (light, polarization, salt, temperature)
- beacon based (homing, goal-directed)
- dead reckoning (time, energy spent, integral of optic flow)
- **landmark based** (salient patterns, orientation flight, Figure 2)
- **cognitive maps** (rats, Tolman [1])
- Cartesian grids (humans & gofai robotics)

2. Research questions

In the long term, we want to develop biologically plausible "cognitive" maps. In this study, we start by asking the following two questions:

1. How can a perceptual map of (a) **salient** and (b) **relevant** landmarks be constructed?
2. If landmarks can be detected during exploration, is there a sufficient amount of information to construct a vectorial map of the environment by a simple mechanism?

3. Design considerations

What is a landmark?

- a **salient** and
- **relevant** perceptual pattern

→ **Salient:** The robocup field is mainly color coded. Search for an image representation which captures color shade while remaining a sufficient degree of invariance to lighting (weather) conditions. Solution: Hue-Satiation images, setting luminance to a constant value.

→ **Relevant:** Define a **proximity event** as a perceptual event that may signal FFF (friend/foe or food). In monkeys, neurons in the medial-intraparietal area will respond to stimuli that invade 'egocentric space' [2]. In our case, we will use multimodality, i.e., vision and sonar, to take snapshots at occurrences of proximity detection by the sonar during field exploration.

4. Experimental setup

- Pioneer IIDX robot; 16 sonar sensors, color CCD camera (Fig. 1)
- odometry on linear trajectory segments
- XSAM agent-based software platform (cf. [3])
- 4x4.60m 'robocup' style field with yellow and blue goal, green floor, a round window to the outside world, fluorescent lighting

Figure 1. Pioneer robot.

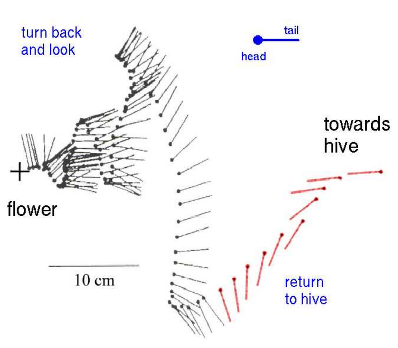

Figure 2. Orientation flight in bees [4]

5. Constructing a map of perceptual landmarks

Perception:

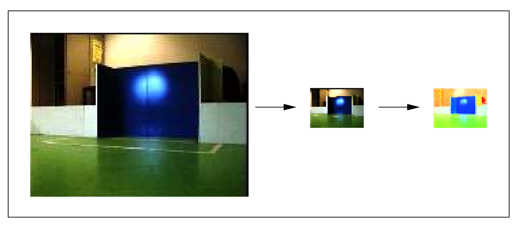

Figure 3. Resolution reduction and HS transform

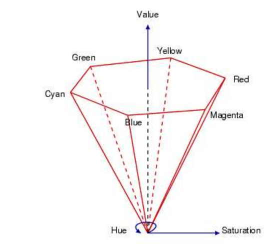

Figure 4. HSV space

- Image: 40x30 (WxH) pixels using Hue and Satiation (HS) only: $N_{dim} = 2400$ (Fig. 3 & 4)
- Proximity: 16 sonar-sensor readings
- Total bimodal perceptual frame: $N_{dim} = 2416$

Exploration

- Collision avoidance behavior...
- ...while sampling perceptual frames at 2 Hz

Procedure

- Run 1: Train a Kohonen self-organized map: Kohonen Landmark Map (KLM) on the bimodal perceptual frames
- Run 2: Collect & average the dead-reckoned landmark-to-landmark distances
- Save the (sparsely filled) 25x25-dim. average-distance matrix Δ for later analyses

Result 1: Kohonen Landmark Map

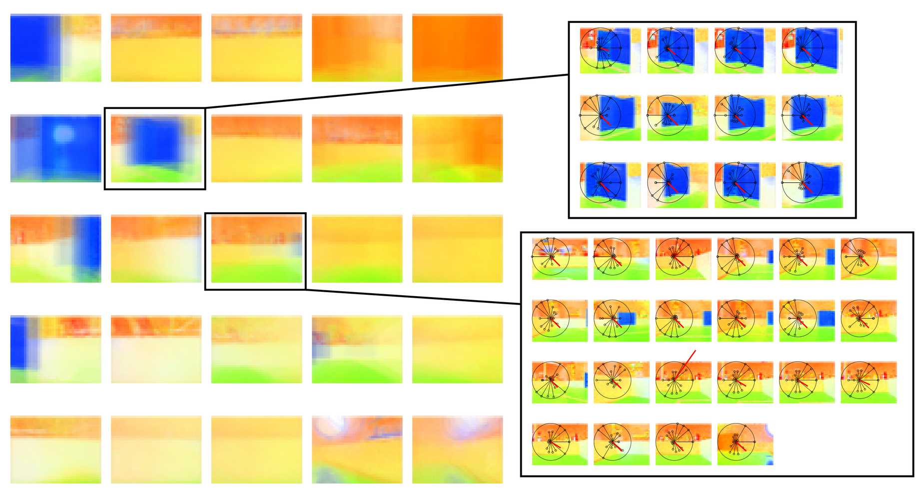

Figure 5. KLM of images and sonar readings. Image information overrules sonar influence. The blue and yellow goal are found, as well as the window in the room. Raw data (inserts, right) show an intuitive collection of nearest neighbours.

A 5x5 Kohonen self-organizing map was considered large enough. Training sets (half an hour of exploration) typically contained 500-700 proximity-event samples. Training parameters: 200 epochs, fast cooling, resulted in a KLM as depicted in Figure 5.

Result 2: Can the distance matrix be used for approximate map building?

Example: three given distances yield a triangle in the 2D plane. Four distances can be mapped to 2D or 3D. How to project an N-dimensional distance matrix to, say, two dimensions and keep the topology?

The position of a point in 2D is a linear transform on its distances Δ_{ij} to the other points, which can be solved in a supervised manner if enough points (x_i, y_i) are known:

$$\vec{O} + \beta\Delta_{i*} = (x_i, y_i)^t$$

Theoretically, due to the redundance in the distance data, an unsupervised estimate of the 2D structure can be obtained, by determining the Eigenvectors of Δ by principal components analysis (PCA): $M\vec{e}_k = \lambda_k\vec{e}_k$ where M is the covariance matrix and the two orthogonal axes $\vec{e}_k$ with largest λ_k are used. The resulting representation is an affine transform of the real 2D space. Problems: missing data (Figure 6) and landmark ambiguity.

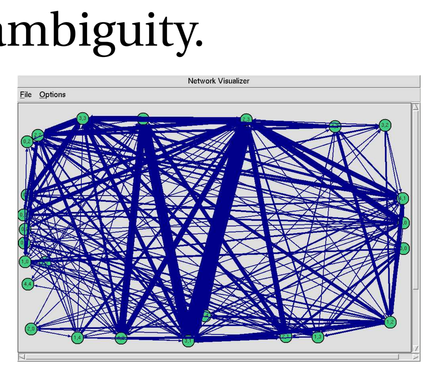

Figure 6. Landmark-to-landmark transition density. The distance matrix will not be complete: missing-value insertions are needed for PCA.

6. From distance matrix to 2D field topology (artificial data)

In order to test the reliability of PCA-based 2D maps, artificial data were generated on the basis of a square grid of 5x5 points, spaced $dx = dy = 1$ apart. Using Monte-Carlo simulations, distance matrices were computed between the points in the grid while varying

- (a) the probability of missing (inter-landmark) distances and
- (b) the amount of noise on the dead-reckoned distances, proper.

Computation

For each distance matrix:

1. fill in missing transitions with Gaussian noise (using distance mean and variance in total set)
2. apply PCA to the distance matrix
3. compute projection of each landmark distance vector on the first two PCA dimensions yielding $(\hat{x}_i, \hat{y}_i)$ estimates.
4. the resulting estimated space is an affine transform of the real space, plus estimation noise
5. correlate the estimated landmark positions $(\hat{x}_i, \hat{y}_i)$ with the the real point positions (x_i, y_i) for landmarks i in the square 5x5 grid.

Results (Figure 7) show robustness against noise on the distance measurements and an acceptable sensitivity for absent landmark transitions: Up to $p(miss) = 0.4$, the correlation between estimated and real positions yields values of 0.9 or higher.

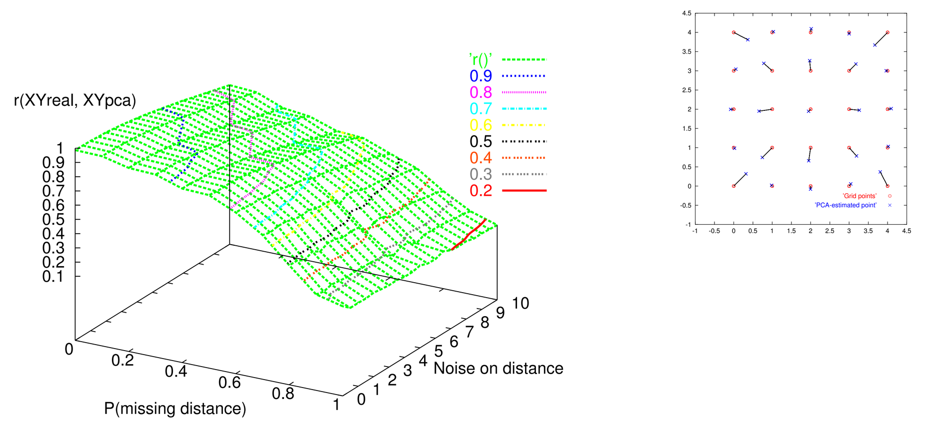

Figure 7. Correlation between real and PCA-based 2D grid points as a function of the probability of missing landmark distances and noise on the distance measurements. Simulation results from an artificial field of 5x5 points, regularly spaced $dx = dy = 1$ apart. Very similar results were obtained for irregular 25-landmark maps. Inset (right): example of grid solution

7. Reconstruction of robocup field topology (real data)

Figure 8 shows the results of an independent exploration run by the robot in the real field.

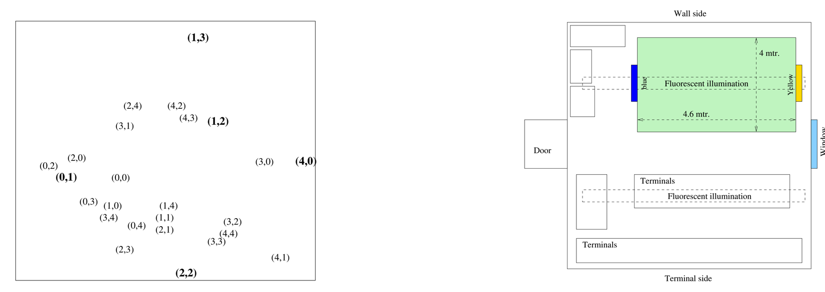

Figure 8. Topological landmark representation based on PCA of the distance matrix (left). Without human intervention, the system has correctly placed the blue(0,1) and yellow(4,0) goals in opposing locations. The position of the views on the window in the room, i.e. nodes [(1,2),(1,3)] at the top, indicate that the y-axis should be flipped. The map (on the right) represents the whole room, of which the actually explored field is the large green rectangle, upper right.

8. Conclusion

- Collecting visual landmarks at proximity events facilitates the training of a self-organized, Kohonen Landmark Map
- The distance matrix which is based on average dead-reckoned distances between such landmarks can be used to find an approximate vectorial map of the environment, using PCA
- several neural-network variants for PCA exist (e.g., diabolo MLPs)
- the location information is there, but rather crude (Figure 8)
- behaviors which exploit heading and angular information would be conducive to calibrate the necessary affine 2D transform
- cf., turn back and look behavior in bees (Figure 2)

References

[1] E.C. Tolman, "Cognitive maps in rats and man," *Psychological Review*, vol. 55, pp. 189–208, 1948.

[2] Colby C.L. and M.E. Goldberg, "Space and attention in parietal cortex," *Annual Review of Neuroscience*, vol. 22, pp. 319–349, 1999.

[3] L.G. Vuurpijl and L.R.B. Schomaker, "Multiple-agent architectures for the classification of handwritten text," *Proc. of IWFHR'98*, pp. 335–346, 1998.

[4] B.A. Cartwright and T.S. Collett, "Landmark learning in bees," *Journal of Comparative Physiology*, vol. 151, pp. 521–543, 1983.

Proc. of the 15th Belgian-Dutch Artificial Intelligence Conference (BNAIC'03), Nijmegen, The Netherlands, Oct. 2003, pp. 275-282.