IAV2004 - PREPRINTS
5th IFAC/EURON Symposium on Intelligent Autonomous Vehicles
Instituto Superior Técnico, Lisboa, Portugal
July 5-7, 2004

APPEARANCE BASED LANDMARK SELECTION AND RELIABILITY EVALUATION FOR TOPOLOGICAL NAVIGATION

Lorenz Gerstmayr [∗,∗∗] **Alexandre Bernardino** [∗]
José Santos–Victor [∗]

[∗] *(lgerstmayr, alex, jasv)@isr.ist.utl.pt*
Vislab, Instituto de Sistemas e Robótica,
Instituto Superior Técnico, Avenida Rovisco Pais 1,
1049–001 Lisboa, Portugal
http://www.isr.ist.utl.pt/vislab/
[∗∗] *Department of Cognitive Neuroscience,*
University of Tübingen, Auf der Morgenstelle 28,
72076 Tübingen, Germany
http://www.uni-tuebingen.de/cog

Abstract: In this paper we propose two algorithms for selecting landmarks, and methods to measure the reliability of the selected landmarks, that will be used in future work to optimize topological navigation of a blimp. The first algorithm evaluates image similarities between possible landmarks. The second one, which can also be used for online landmark selection, uses Incremental Principal Component Analysis and a measure of how good a new landmark can be expressed in the existing eigenspace.

Keywords: Aerial vehicle, landmark selection, visual navigation, topological navigation

1. INTRODUCTION

As (Thrun, 1998) states, the concept of landmarks is easy to understand but difficult to define. He defines landmarks as elements, i.e. objects or image features that can serve as reference. This definition can be a superordinate definition for the one used here and in related work like (Knapek *et al.*, 2000; Ohba and Ikeuchi, 1997). We define a landmark as a subimage of a known image that is as dissimilar to all other considered subimages as possible.

⋆ This work was partially funded by the project RESCUE - Cooperative Navigation for Rescue Robots FCT SRI/32546/99-00 (http://rescue.isr.ist.utl.pt/).

The future aim of the described work is the topological navigation of an autonomous blimp, i.e. a zeppelin without a skeleton, integrated in the RESCUE project, whose purpose is to navigate and map urban areas in search and rescue scenarios. A practical setup (see fig. 1) was developed at Vislab and consists of an indoor blimp that flies over a huge poster of an aerial image of Lisbon, Portugal, shown in fig. 2. The blimp is equipped with a camera looking downwards to the poster.

For topological navigation it is necessary that the agent memorizes several distinct places and the relation between these places in a graph like structure, the topological map (Franz and Mallot, 2000). The agent is able to navigate from one known place to another via other memorized places. Therefore it is necessary to localize itself at each known place. In order to maximize the robustness of localization, these known places or landmarks have to be selected carefully. In this paper we focus on the selection of good landmarks. (Ohba and Ikeuchi, 1997) state that good landmarks are supposed to be detectable, unique and reliable. Landmark selection is closely related to object recognition methods (Ohba and Ikeuchi, 1997) or image retrieval (Schmid and Mohr, 1997) where discriminative parts of the image are used to characterize the object or image respectively.

Fig. 1. The indoor blimp flying over the poster of the aerial image

All the landmark selection methods try to reduce uncertainty. Some approaches like (Thrun, 1998; Olson, 2002) use a probabilistic approach reducing localization uncertainty. Others try to use landmarks in areas where it is known that landmarks lead to small localization errors (Sutherland and Thompson, 1994; Burschka *et al.*, 2003).

We consider approaches that address the landmark selection problem in two steps: first they pre–select a set of interesting image points or landmarks by using attention or interest operators like corners (Knapek *et al.*, 2000; Jugessur and Dudek, 2000; Little *et al.*, 1998; Ohba and Ikeuchi, 1997; Schmid and Mohr, 1997) or edge density ((Bourque *et al.*, 1998)). In a second step they test the pre–selected points for reliability or uniqueness, and reject the candidates not satisfying that criteria. Then, for localization or recognition, subimages around the selected points are used.

The work presented in this paper only addresses the second part of the problem. We propose two algorithms based on Principal Component Analysis (PCA) techniques to select unique landmarks and evaluate their reliability. Here we do not make any pre–selection of features due to the rather uniform nature of the used data set. However, nothing prevents the proposed algorithms from having a prior attentional module that pre-selects some promising candidates to good landmarks.

With regard to the second step related work like (Knapek *et al.*, 2000; Schmid and Mohr, 1997; Ohba and Ikeuchi, 1997) compare image similarities and select landmarks that are as dissimilar as possible. In (Ohba and Ikeuchi, 1997) the number of landmarks is further reduced by discarding landmarks that are not stable to small changes of the viewpoint. (Little *et al.*, 1998) use stereo information to detect and discard corners that resulted from overlapping objects, keeping corners on planar surfaces. (Jugessur and Dudek, 2000) compute the standard deviation of the pixel values in the subwindows and only keep landmarks where it is above a threshold. (Johnson, 2000) proposes a method for terrain matching. He discards landmarks that are located in terrains with high curvature or in planes, because small changes in the sensor measurement result in great localization errors, or the landmarks are too similar, respectively. In general the advantages of selecting landmarks are better localization and a speed up of the computations because less landmarks are used.

The reminder of the paper is structured as follows: in the next section we will introduce the proposed algorithms. The third section shows the experiments. Results and conclusions will be given in the last two sections. Since PCA got an established method in the field of computer vision, the nonfamiliar readers are referred to (Murase and Nayar, 1995) from which we kept the mathematical notation.

2. OUR APPROACH

The long term goal of our work is aerial navigation on large urban areas. The data set used in this work is shown in fig. 2. As it can be seen, corners, edge density, symmetry and other features commonly used as attentional operators have a rather uniform distribution, thus limiting their discriminative power in this environment. In this work, instead of preselecting detectable landmarks, we divide the aerial image into a grid of subimages that are used as possible landmarks. For these subimages we derive features in the form of coefficients of a Principal Component Analysis (PCA). PCA tends to obtain compact and efficient representations of the global environment. Although PCA got an established method for recognition and robot localization, we do not know any related work on landmark selection that uses pure PCA features to select good landmarks.

The idea of both algorithms is to select landmarks that are as dissimilar as possible. The profile–based algorithm selects dissimilar images by comparing image distances. The algorithm based on Incremental Principal Component Analysis (IPCA) tries to select images that can not be expressed accurately in the existing eigenspace, and updates the eigenspace with these images.

Fig. 2. The aerial (gray scale) image used for the experiments

2.1 Profile–based landmark selection

This algorithm takes a set of images $\boldsymbol{g}_i;\ i = 1...n$ in the computed eigenspace and the number of landmarks l that should be selected as input. In the next step it computes a distance matrix $\boldsymbol{D}$ with the pairwise image distance as entries

$$D_{i,j} = \text{dist}\left(\boldsymbol{g}_i, \boldsymbol{g}_j\right). \tag{1}$$

As distance function, we use the Sum of Squared Differences (SSD) normalized by the variance of the vectors. Then a distance profile vector $\boldsymbol{p}$ with

$$p_j = \frac{1}{n-1} \sum_{i=1, i \neq j}^{n} D_{i,j}. \tag{2}$$

is computed by averaging over column j of the distance matrix. Since $D_{i,j}$ is close to unity if the subimages $\boldsymbol{g}_i$ and $\boldsymbol{g}_j$ are very dissimilar, a good landmark subimage that is very dissimilar to all other subimages will have a profile value close to unity. The selection is done by sorting $\boldsymbol{p}$ in an ascending order and selecting the landmarks corresponding to the greatest l values.

2.2 IPCA–based landmark selection

The second algorithm uses IPCA to compute the image ranking. IPCA was developed to overcome several drawbacks of the usual batch method to compute the PCA, related with the update of an existing eigenspace. It iteratively updates an eigenspace by adding a new vector that is orthogonal to the existing basis vectors. When a new landmark is selected, a new vector is added to the eigenspace. The new vector is given by the residue of the current image projection in the existing eigenspace. The norm of the residue vector gives a measure of how good the image can be expressed in the existing eigenspace. A complete derivation of the used algorithm to update the eigenspace can be found in (Freitas *et al.*, 2003; Hall *et al.*, 1998).

The algorithm we propose for landmark selection requires a set of images S, the number of landmarks l and the set of remaining subimages $X' := X \backslash S$. It then computes an eigenspace model Ω for the start images S that are used as the the first landmarks.

Then every image $\boldsymbol{x}_i \in X'$ is transformed to the eigenspace and the resulting residue vector $\boldsymbol{r}_i$ is computed. Since the norm of the residue vector is a measure how good the considered image can be expressed in the current basis, the maximal norm of all the residue vectors is computed and the corresponding image is selected as new landmark.

Then the picked image has to be removed from X' and the eigenspace model is updated. The algorithm terminates if the number of selected landmarks is greater then l.

Since the images contained in the start set S are used as landmarks, the further ranking is dependent on the selected images. For this work we fixed the number of start images and considered two methods for selecting start views:

(1) Choosing the start set according to the profile ranking.
(2) Choosing images randomly. This method is very fast but it is not predictable how the start images influence the further ranking.

Although not explicitly used in the work presented here, the IPCA–based landmark selection algorithm is capable of selecting landmarks in unvisited areas and can therefore be used for Simultaneous Localization and Mapping (SLAM). Though most works in the area of landmark selection assume the existence of a prior visit to the environment in an offline phase, to select possible landmark locations, the proposed algorithm can also be used for on–line landmark selection.

2.3 Reliability Evaluation

When the blimp has to localize itself in the topological map a nearest neighbor search has to be done. For each time step, the current camera image has to be compared to the known landmarks. The landmark to which the current image is as similar as possible is selected as the blimp's closest landmark. To measure a landmark's reliability we assume the blimp is at the landmark's position, except for little errors in the orientation, the position or the altitude. These little changes result in views that are either rotated, translated or have a different view window respectively. Additionally we consider changes in the image brightness. For the moment we do not consider image deviations that are due to pitch and roll movements of the blimp.

The current view is transformed to the eigenspace resulting in a point with coordinates differing from those of the landmark. The task is now to find out how much the view can be changed so that the landmark is still selected as nearest neighbor. Therefore we propose to compute for a set $L := \{\boldsymbol{g}_{i'} | i' = 1...l\}$ of selected landmarks, the maximal image dissimilarity ϵ so that the correct landmark is still selected. We define this limit ϵ as half the minimal dissimilarity between the correct landmark $\boldsymbol{g}_j$ and the other landmarks:

$$\epsilon := \frac{1}{2} \min_{\forall i', i' \neq j'} \left(\text{dist} \left(\boldsymbol{g}_{i'}, \boldsymbol{g}_{j'} \right) \right) \tag{3}$$

For a landmark $\boldsymbol{g}$ and a view $\tilde{\boldsymbol{g}}$ the equation

$$\text{dist} \left(\boldsymbol{g}, \tilde{\boldsymbol{g}} \right) < \epsilon \tag{4}$$

gives us a sufficient criterion to choose the closest landmark to the current view.

A similar method was used by (Ohba and Ikeuchi, 1997) to select the most stable and therefore the most reliable landmarks. The authors have chosen an absolute limit and only consider little changes in the rotation. Most other related works only show experimental data to proof the reliability of the selected landmarks.

3. EXPERIMENTS

For the experiments the aerial image of Lisbon, Portugal, shown in fig. 2 was used. The size of the image is 1600×1200 pixels, and image brightness is discretized in 256 gray values. For the experiments the whole image was divided into a grid of 48 squared subimages sized 200×200 pixels (referred to as views) and another 35 views were positioned at the adjoining points of the other views resulting in a total amount of 83 views with at most 25% overlap. All views (possible landmarks) are used to compute the eigenspace. All the experiments were done for several dimensions of the eigenspace and for several levels of scaling. For the downscaled images the number of landmarks was kept constant.

The first part of the experiments consisted in computing the best landmarks with both algorithms, for each possible combination of scalings and dimensions of the eigenspace. For all methods a number of $l = 8$ landmarks was selected to allow a better comparison of the results. The start list for the IPCA–based ranking contained 3 landmarks.

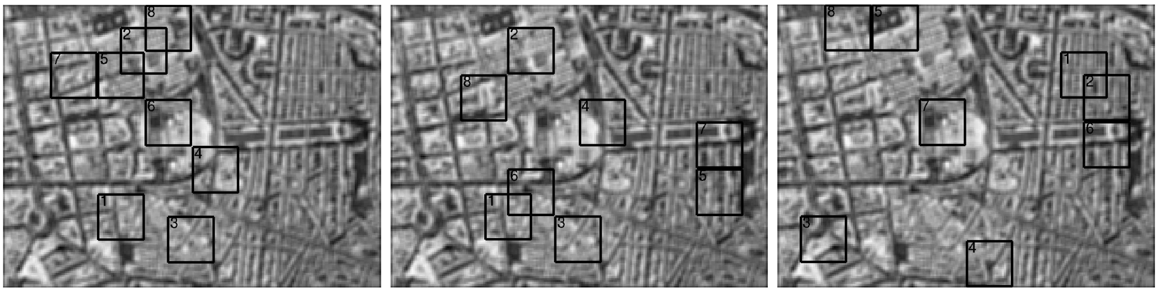
The second part of the experiments consisted in measuring the reliability of the selected landmarks. This is done with respect to deviations in orientation, altitude and image brightness, starting with $\tilde{\boldsymbol{g}}$ as an identical view and making it more and more dissimilar. For each parameter (orientation, scale and brightness), we computed both the upper and lower limits of the range where (4) still holds. Values between the upper and lower limit curves in fig. 5 do not produce erroneous localizations. For evaluating deviations in the position we spiraled around the exact position and counted the number of views, for which the image similarity value is below the threshold ϵ and for which another view in a 3×3 neighborhood is similar enough. So we computed the size of a catchment area, for which exact localization to the considered landmark is possible.

4. RESULTS

Since none of the proposed methods is clearly superior to others, we only present some representative results. For a complete overview of the results, interested readers are referred to (Gerstmayr *et al.*, 2004). Since for the range of parameters we tested we did not find a clear dependence between image downscaling and landmark reliability, all the presented results were downscaled with a scaling factor of 8 and we focus on the influence of the number of dimensions of the eigenspace.

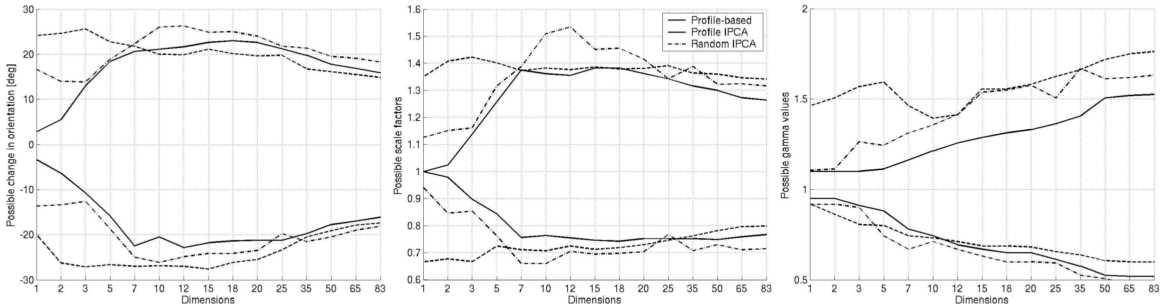
Fig. 3 shows the results of the landmark selection for the methods proposed using a 12–dimensional eigenspace. All methods have in common that they select landmarks that all look different. Often a landmark contains a unique pattern formed by streets or buildings. The profile based algorithm shown in fig. 3(a) selects landmarks only in areas that are not grid like or repetitive because these areas contain the most similar views of the image. Fig. 4 shows the computed profile vectors for all considered dimensions of the eigenspace. The abscissa represents all the possible landmarks or views and the ordinate represents the dimension of the eigenspace. So each "row" of the plot contains one profile vector computed with (2). The plot shows that for more than 10 dimensions the selection gets stable because the same landmarks show out as very distinctive (coded by white) or very similar (coded by black). It also shows a drawback of the aerial image or the arrangement of landmarks used: there are only little views that show out as very distinctive but plenty of views have profile values coded by an average gray scale.

The results of the IPCA–based ranking are very dependent on the views contained in the start lists. For different startlists with very similar views the, results will be similar. If the views are distinctive, the results differ. It is far beyond the scope of this paper to deal with the influence of the start lists on the results.



(a) Profile based ranking (b) Profile based IPCA–ranking (c) Profile based ranking with random start list

Fig. 3. Landmarks selected by the proposed algorithms using a 12–dimensional eigenspace. The images were downscaled by a factor of 8 and smoothed with a Gaussian smoothing. The numbers are the rank of the landmark



(a) Orientation (b) Scaling (c) Image brightness

Fig. 5. Limits for the allowed deviations. Within the limits, landmark localization is reliable.

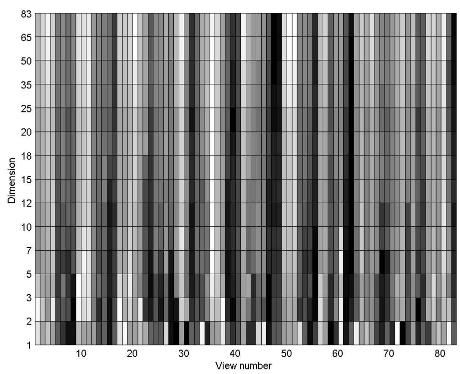


Fig. 4. Profile vectors for different dimensions of the eigenspace. White codes very dissimilar, the values were normalized to $[0, 1]$

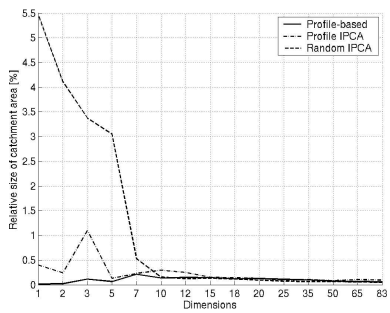


Fig. 6. Reliability according to size of catchment area

The results of the reliability evaluation are shown in figure 5. The abscissa of each plot represents the number of dimensions of the eigenspace. The lines in the plot correspond to the averaged lower and upper limit computed by simulating errors for the landmarks selected by the different selection methods. For a low dimensional eigenspace, the plots for the possible deviations in orientation and scaling show an increase in the reliability. For more then ten dimensions a slight decrease in reliability can be observed, which might be due to the fact that the images are then expressed more accurately resulting in greater dissimilarities for smaller deviations. The plot for the possible change in image brightness shows an increase of robustness with increasing number of the dimensions. Fig. 6 shows the relative size of the computed catchment areas compared to the total size of the image. The percentages seem to be rather small but one has to keep in mind that a view only covers approximately 2% of the total image and that landmarks are characterized by unique patterns. For localization these patterns have to be visible in the view and have to be at approximately the same position as they are at the landmark. Since the views are represented very coarsely in a low dimensional eigenspace, landmarks like landmark 1, 2 and 6 in the IPCA–based ranking with random start list are surrounded by great catchment areas and dominate the mean size shown in fig. 6. For more then 10 dimensions, the images are represented more accurately and the sizes of the catchment areas are comparable to those of the landmarks selected by the profile–based ranking. Since the profile–based ranking selects those landmarks that are most distinctive for each dimension of the eigenspace, no landmarks in repetitive areas are selected, which seem to result in greater catchment areas.

5. CONCLUSIONS

In this work we present two algorithms for appearance–based landmark selection. The first algorithm evaluates image dissimilarities between possible landmarks. The second one enlarges an existing eigenspace stepwise by adding that view that can be represented worst in the existing eigenspace. The results show that both algorithms perform equally well. In case a more then 10–dimensional eigenspace is used, all methods select landmarks that are robust over a large range of image deviations and have a very similar averaged reliability. For our particular application using an eigenspace with 10 to 15 dimensions seems to be optimal, because there is a trade–off between generalization in low–dimensional eigenspace and exact representation in higher dimensional eigenspace. If many dimensions are used small image deviations lead to great differences between the images resulting in a loss of reliability.

Since the proposed methods will be applied in large scale environments, there will be the need for discarding possible landmarks in order to speed up the analysis. Therefore we want to apply image preprocessing like color segmentation or line enhancement to bring out the underlying geometrical pattern. We think that these overall geometrical aspects of roads and buildings are the most relevant features for human observers. Unlike low–level attention mechanisms, these methods should allow discrimination between different places. In order to enhance the pattern formed by the street contour grouping exploiting the color transitions from red roofs to black streets will be used. Other plans aim to run real world experiments with the blimp.

REFERENCES

Bourque, E., G. Doudek and P. Ciaravola (1998). Robotic sightseeing – a method for automatically creating virtual environments. In: *Proceedings of the ICRA 1998*. pp. 3186–3191.

Burschka, D., J. Geiman and G. Hager (2003). Optimal landmark configuration for vision–based control of mobile robots. In: *Proceedings of the ICRA 2003*. pp. 3917–3922.

Franz, M.O. and H.A. Mallot (2000). Biomimetic robot navigation. *Robotics and autonomous Systems* **30**, 133–153.

Freitas, R., J.A. Santos-Victor, M. Sarcinelli-Filho and T. Bastos-Filho (2003). Performance evaluation of incremental eigenspace models for mobile robot localization. In: *Proceedings of the ICAR 2003*. pp. 417–422.

Gerstmayr, L., A. Bernardino and J. Santos-Victor (2004). Appearance based landmark selection and reliability evaluation for topological navigation. Technical report. Vislab – ISR, Instituto Superior Técnico, Lisbon.

Hall, P.M., A.D. Marshall and R.R. Martin (1998). Incremental eigenanalysis for classification. In: *Proceedings of the BMCV 1998*. pp. 286–295.

Johnson, A.E. (2000). Surface landmark selection and matching in natural terrain. In: *Proceedings of the CVPR 2000*. pp. 413–420.

Jugessur, D. and G. Dudek (2000). Local appearance for robust object recognition. In: *Proceedings of the CVPR 2000*. pp. 834–839.

Knapek, M., R.S. Oropeza and D.J. Kriegmann (2000). Selecting promising landmarks. In: *Proceedings of the ICRA 2000*. pp. 3771–3777.

Little, J.J., J. Lu and D.R. Murray (1998). Selecting stable image features for robot localization using stereo. In: *Proceedings of the IROS 98*. pp. 1072–1077.

Murase, H. and S.K. Nayar (1995). Visual learning and recognition of 3–D objects from appearance. *International Journal on Computer Vision* **14**, 5–24.

Ohba, K. and K. Ikeuchi (1997). Detectability, uniqueness and reliability of eigen windows for stable verification of partially occluded objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**, 1043–1048.

Olson, C. (2002). Selecting landmarks for localization in natural terrain. *Autonomous Robots* **12**, 201–210.

Schmid, C. and R. Mohr (1997). Local grayvalue invariants for image retrival. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**, 530–535.

Sutherland, K.T. and W.B. Thompson (1994). Localizing in unstructured environments: Dealing with errors. *IEEE Transactions on Robotics and Automation* **10**, 740–755.

Thrun, S. (1998). Bayesian landmark learning for mobile robot localization. *Machine Learning* **33**, 41–76.