Proceedings of the 1991 IEEE
International Conference on Robotics and Automation
Sacramento, California - April 1991

Autonomous Landmark Selection for Route Recognition by A Mobile Robot

Jiang Yu Zheng†, Matthew Barth and Saburo Tsuji
Department of Control Engineering, Osaka University
Toyonaka, Osaka 560, Japan

Abstract

This work introduces an approach to build a qualitative description of scenes along a route, which is used in route recognition by a mobile robot. The description consists of a series of *Landmarks* autonomously selected by the robot from a *Generalized Panoramic View*, which has been generated as a visual memory of scenes along routes. The basic idea to bridge the quantitative panoramic view to qualitative landmarks is to examine the 'distinctiveness' of patterns in the image and select landmarks from unique patterns that are remarkable by which to navigate.

1. Introduction

1.1 Background

Long distance mobile robot navigation is faced with the two major problems of how to memorize scenes along a route and how to recognize the route when referring to these scenes. In order to investigate these topics, we have considered the following strategy. First, a robot constructs an internal representation of a route from specific scenes during a trial move guided by a human. When the same route is subsequently pursued again by the robot itself, it locates and orients itself based on the memorized scenes[1]. A *Panoramic Representation* used for presenting and retrieving scenes has been previously proposed[2,3]. Briefly, this representation is generated from slices of continuous images taken by a camera on a mobile robot. The resulting *Generalized Panoramic View* (PV)[4,5] contains the major visual information in the continuous images with a small amount of data, and depth information is estimated using multiple PVs from the same image frames. Figure 1 gives an example of a PV along a route in a miniature model.

Although the panoramic representation consolidates the local visual information into a global form, it still requires much memory space if the environment has large scale. A large majority of scenes taken along the route may change very little over long distances, and thus provides less meaningful information when composing a global representation of the route. Therefore, for long routes, a *qualitative representation* is better suited for route memorization. Also for route planning, a map on the qualitative level tends to be necessary.

This paper describes a method of constructing such a qualitative representation from the panoramic representation. Specific scenes, denoted as 'landmarks', are extracted from the PV. The advantages of a qualitative description lie in two major aspects: (1) It is small and compact, and its access is straightforward. (2) It is robust to the small changes in parameter values and view points and is thus reliable during recognition.

1.2. General Strategy

How do we proceed towards a qualitative description of scenes from quantitative image data? Let us first consider how a human memorizes a route. According to cognitive science, a key function of the visual memory of humans is *selectivity*[6]. A human only remembers the most distinct scenes when subjected to a large amount of visual information. When a human traverses a new route, what he most likely memorizes are the most 'striking' scenes in the range of observation, and he will probably describe the route by these scenes. Usually, we refer to these distinctive scenes as *Landmarks*.

One of the more challenging tasks for a mobile robot is the autonomous extraction of landmarks along a route. Up to now, such tasks have been done primarily with the assistance of human operators identifying objects in specific environments[7,8]. When such *a priori* knowledge is used for landmark selection, route recognition is a goal-driven event. However, when no *a priori* knowledge can be given about the landmarks, route recognition algorithms must rely on data-driven methods.

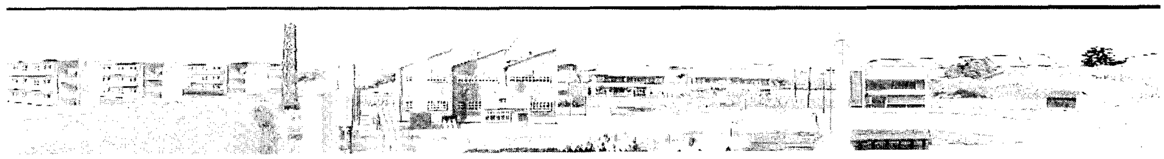

Fig.1 A segment of generalized panoramic view of scenes along a route.

† *Current Address: ATR Communication Systems Research Laboratories, Seika, Soraku, Kyoto 619-02, Japan*

Ours is primarily a data-driven method. Since a mobile robot should be able to move in various environments, our landmarks are not defined as any specific type of object. In our method, bottom-up algorithms operate to autonomously extract landmarks from the scenes during robot's learning phase along the route. The criteria for landmark selection are based on the following ideas:

1. The scenes extracted as landmarks should be remarkable in size and explicit in contrast or color so that they will not be missed in recognition.
2. The scenes extracted as landmarks should be unique and distinctive either in properties or in structure, compared with other scenes in certain ranges.

We measure how unique a scene is from its *Distinctive Range* in which no other similar scene appears. We describe a scene not only on its own attributes (i.e. color, shape, or structure) but also on its spatial relationship with other scenes. Our landmarks are selected as sections of a PV which contain distinct and unique scenes.

Since the panoramic representation incorporates the majority of information in the scenes visible along the route, it is well suited for comparisons between extracted patterns when determining distinctiveness. Furthermore, as a goal, the landmark-based qualitative representation should nearly maintain the equivalent information necessary for route recognition as the PV.

2. Basis for Landmark Detection

2.1 Panoramic Representation

We briefly describe how a panoramic view is formed. A camera on a mobile robot moves along a smooth curve on a horizontal plane, with its optical axis directed sideways from the direction of motion. The scenes along the route are observed through a vertical slit. This is realized by capturing data at the slit from each image taken continuously, and pasting them consecutively at the position along the trajectory as figure 2 depicts[3]. The scenes along the route are therefore projected towards a vertical surface determined by the trajectory of the camera. The PV has a wide field so that it is useful in presenting a global relationship of the scenes. The amount of data in a PV is equivalent to only a slice in the spatiotemporal volume constructed from continuous images[5].

If we set two vertical slits on the image frame and generate two PVs simultaneously, we can measure 2(1/2)-D scene information, given by the time delay of features passing the slit lines. The acquired 2(1/2)-D information is path-oriented and 3D information can be computed from the motion parameters given by the robot. The final panoramic representation is composed of the PV, 2(1/2)-D information, and a description of the camera motion.

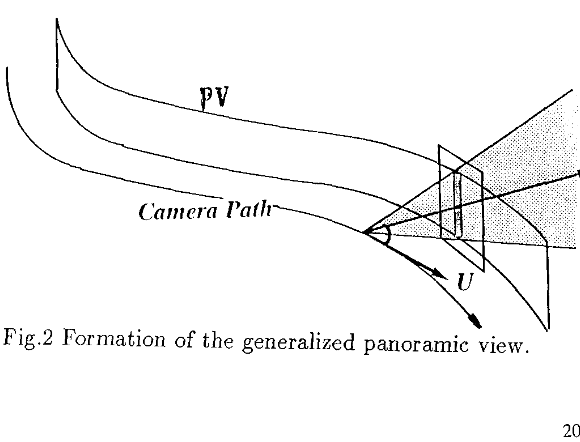

Fig.2 Formation of the generalized panoramic view.

We have realized matching of the generalized panoramic views on an iconic level. This is important for acquiring 2(1/2)-D information, and referencing memorized scenes during route recognition. The matching proceeds in a coarse to fine fashion, first matching profiles of the PVs along the route using *Dynamic Programming*, and then matching structural lines by comparing their attributes.

2.2. Features for Determining Landmarks

In determining landmarks, we use several inherent qualities provided by the PV. These qualities fall into a natural hierarchy, as illustrated in figure 3. We can group these qualities or features roughly into three groups:

1. *Low-level features*: Statistical features intrinsic to the scenes. They require little extraction processing, and it is usually upon these low-level features that segmentation takes place. Some examples are brightness, hue, and texture.
2. *Mid-level features*: Iconic features calculated from the low-level features, after segmentation takes place. Some of these derived features are area, perimeter, shape descriptors, etc.
3. *High-level features*: Features based on structural or syntactical properties in the scenes. In this case, grouping of mid-level features usually takes place.

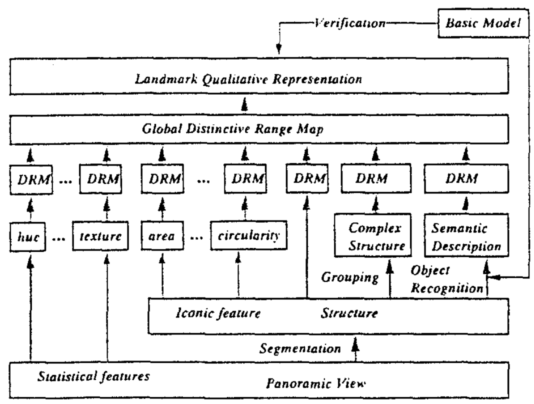

Fig.3 Block diagram of extracting landmarks.

We apply different algorithms within this hierarchy in order to extract a measure of distinctiveness among the scenes. From each attribute, we derive a *Distinctive Range Map* (DRM) and subsequently merge all of the DRMs into a global distinctive range map. This is then used for landmark extraction. In this paper, we will focus on analysis using lower level features.

3. Finding Scenes Unique in Attributes

3.1. Constructing Parameter-distance Space

Since the attributes of scenes yield a symbolic description more easily than their structure, we first investigate the distinctiveness of the attributes of scenes. Let $P(i)(i = 1, 2, ...I)$ denote patterns in a PV, and $A(i) = [a_1(i), ..., a_K(i)]$ denote an m-dimensional vector whose components are the pattern attributes. The $a_k(i)(k = 1, 2, ...K)$ can be any type of low- and mid-level attributes.

For simplicity of computation, we consider each attribute separately. A *Parameter-Distance Space* (PDS) is defined so that the values of the pattern's attributes are given along the coordinate axis A, and the corresponding location of the patterns in the PV are given along the horizontal axis S. Such a space is depicted in figure 4. The parameter values on a vertical line in the PV are mapped to the PDS at its corresponding location to form a histogram. A large homogeneous pattern in the PV will form a long horizontal cluster in the space[1].

It is easiest to construct a PDS for low-level intrinsic statistical features that do not require segmentation of the PV. As an example, we concentrate on color, since it is view point invariant. After performing a color constancy algorithm[11,2], hue and saturation are extracted from the digitized R, G, B data. The PDS is then thresholded to eliminate the low clusters from small regions. Figure 5 shows the hue and saturation PDS from a segment of the PV in figure 1. Another attribute which can be directly computed from the PV is texture. We examine the horizontal and vertical spatial frequencies which reflect the complexity of artificial objects in the PV.

In order to extract the group of mid-level features, specifically the iconic attributes that are calculated from the intrinsic low-level features, segmentation takes place on color, using an iterative region expansion technique. After analyzing connected regions, it is possible to extract a large number of iconic features. Some of the simpler features are area, perimeter, and low-order moments. As an example of an iconic feature used for landmark selection, we consider the circularity and orientation of a region. Another useful iconic feature that is used in landmark selection is a measure of parallelism in the sides of the regions.

3.2. Locating Distinctive Range

In the identification of distinctive or salient patterns, conventional methods only classify clusters in parameter space mapped from an image, and then isolate those clusters separated from the majority[9,10]. However, the mapped clusters in parameter space no longer contain any spatial information of patterns in this method. It neglects an important factor that the spatial relationship between patterns also influences the distinctiveness. In landmark selection, landmarks need not have unique attributes if they do not appear elsewhere in a large context; e.g., a pattern can still be very unique if it is spatially far from a lot of patterns with the same color. We can describe the *Distinctiveness* of a pattern in a PDS by a non-expansible rectangle in which no other clusters exist. Thus, not only is the saliency in parameters considered, but also the spatial relationship with the nearest similar patterns is examined.

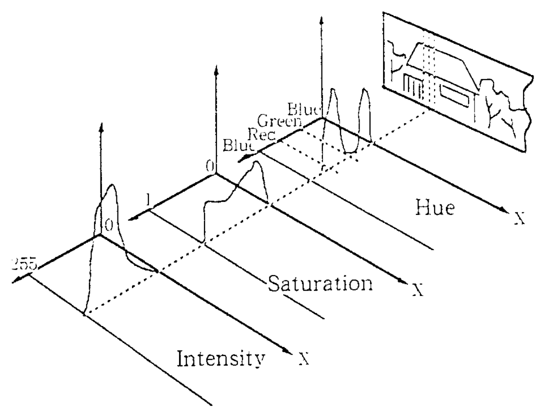

Fig.4 Constructing Parameter-Distance Space of at tributes.

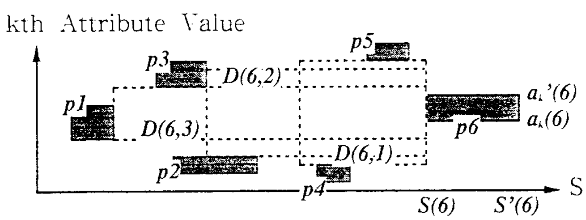

Fig.6 Computing Distinctive Range of a pattern in one attribute.

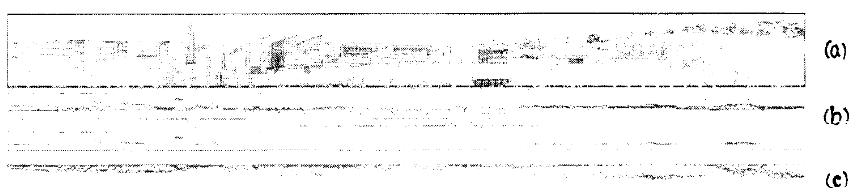

Fig.5 Parameter-Distance Spaces from the real panoramic view. (a) Original panoramic view. (b) PDS of hue. (c) PDS of saturation.

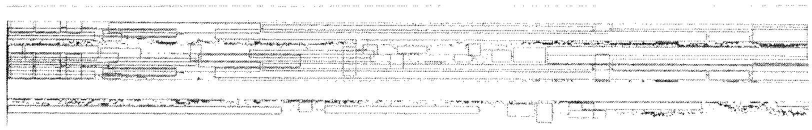

Fig.7 Distinctive range maps obtained from PDS in Fig.5.

Suppose $p(i)$ is the mapping of a pattern $P(i)$ from the PV to the PDS of the kth attribute with variation and location given by $[a_k(i), a'_k(i)]$ and $[S(i), S'(i)]$. If the pattern is very unique for a long distance, there should not exist any other patterns possessing the same value in the range. In other words, patterns near $P(i)$ in parameter values should be far from it in distance, and patterns near it in distance should be far in parameter values. We can locate several non-expansible regions centered at the value $(a_k(i) + a'_k(i))/2$ behind it in which no other patterns exist, as depicted in figure 6. Let $d(i,j)$ and $D(i,j)(j = 1, 2, ...J)$ denote the width and length of such rectangles of pattern i. We choose the rectangle with the largest area and denote its width and length by $d(i)$ and $D(i)$, which we call the *Discriminating Range* and *Distinctive Range*, respectively. The discriminating range $d(i)$ must exceed a certain value corresponding to the minimum discriminable threshold used for the attributes. If we can acquire a long distinctive region with a reasonable width, the pattern that follows it along the S axis will be unique over the long distance and will be selected as a candidate for landmark. Figure 7 shows the distinctive range maps of several features for a particular section of the PV.

4. Landmark Selection

4.1. Merging Distinctive Ranges to a Global Map

We now consider the approach of selecting landmarks for a given sub-route and how to connect them into a qualitative model. Once all of the distinctive range maps from various attributes or structure are determined, there is no insurance that we can find a sequence of distinctive ranges in a single parameter-distance space reaching an arbitrarily-given destination. Also, there is the possibility that we have multiple choices in selecting the distinctive ranges. In order to easily select landmarks, we combine the distinctive ranges into a *Global Distinctive Range Map*. The merged map still consists of many rectangles indicating distinctive ranges from different features.

4.2. Connecting and Updating Landmarks

After global distinctive range map is fused, we attempt to search for a connection of the distinctive ranges reaching the desired destination. Our method of connecting and updating landmarks is similar to what a human does while traversing a route. The patterns with long distinctive ranges are chosen as candidates of landmarks, and a sequence of candidates is incrementally strengthened as we move towards a destination. At the same time, if an incoming pattern has a longer distinctive range which is equal to the total length of the distinctive ranges of landmarks in a sub-sequence, the sub-sequence is updated with the incoming pattern. This is done so that the most unique patterns remain among the scenes along the route. Let $D(i)$:$[a(i), a'(i)][S(i), S'(i)]_k$ denote a distinctive range from the kth attribute in the global distinctive range map. The landmark candidate $L(i)$ follows it. DS denotes an arbitrary destination and C denotes the current viewed position. The landmark sequence that has generated is $[L(i_1)$,...,$L(i_l)$,..., $L(i_m)]$. The algorithm is given as follows:

1. Start from $C = 0$ and set the landmark sequence *null*.
2. Extend the current position $C = C + d$ until a new candidate of landmark $L(j)$ enters, i.e., $S'(j) < C$.
3. If the candidate $L(j)$ has a long distinctive range $D(j)$, e.g., there exists an $l < m$ so that either $S(j) < S(i_l)$, or $S(j) < S'(i_l)$ and $S(j) < S'(i_{l-1})$ are both satisfied, we substitute $L(j)$ for the subsequence $[L(i_l)$, $L(i_{l+1})$, ..., $L(i_m)]$, and the landmark sequence becomes $[L(i_1), ..., L(i_{l-1}), L(j)]$.
4. If $C > DS$ stop, else goto step 2.

The final sequence yields the most distinctive landmarks to pursue when the destination is given. We do not aim at finding a unique landmark sequence nor do we necessarily want the sequence to be similar to that which a human uses. Figure 8 shows sequences of distinctive ranges extending to several specified positions where the robot is supposed to stop or turn. The landmarks locate at the positions behind these distinctive ranges. In this result, only hue feature is used. Figure 9 shows another result of the landmark sequences where saturation is integrated into the Global Distinctive Range Map as well. Some parts in the obtained sequences are not connected. Although we can choose even short distinctive ranges to connect them, it is better to describe these gapes by fuzzy distance, say about 10 meters and etc..

5. Experiments

Experiments were carried out using a miniature model (figure 1), as well as in an urban area with real complex scenes[13]. In the real outdoor environment, a camera mounted on a mobile robot recorded the scenes along a three kilometer route. The robot speed varied slightly from 5 km to 6 km per hour, and continuous images were captured at the rate of 10 frames per second. Prior to any processing on the PVs, a color constancy algorithm is applied in order to equalize any temporal and spatial variations of the ambient light. In dealing with low level attributes, there is not much difference on the performance of the algorithm between indoor miniature model and outdoor scene. Labeling is performed with a multi-valued connectivity analysis routine[12] after region segmentation[2], followed by the extraction of features such as area, perimeter, moments, etc. Edges are obtained from filtering the PVs followed by edge finement. In the outdoor situation, we pay more attention to the vertical lines because the horizontal lines in 3D space usually are curves and are easily destroyed by the shaking of the camera when the road is uneven.

6. Discussion

The landmarks described in this paper have not included those based on semantic meaning that might be obtained from organizing regions to meaningful identities commonly appearing in the environment. We have not considered the functions of these identities, nor the possibility of guidance based on character interpretation, which often play an important role in human landmark selection. Neither we want to select a landmark sequence uniquely, nor we want them to be similar with the landmark we human use. We do intend to use increasingly complex features for landmark

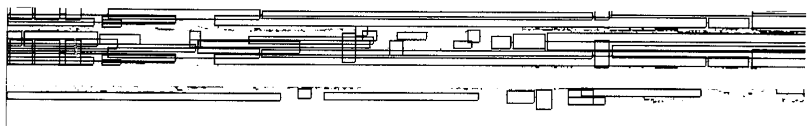

Fig.8 Selected landmark sequences displayed by distinctive ranges. The start point is at left and destinations are given arbitrarily.

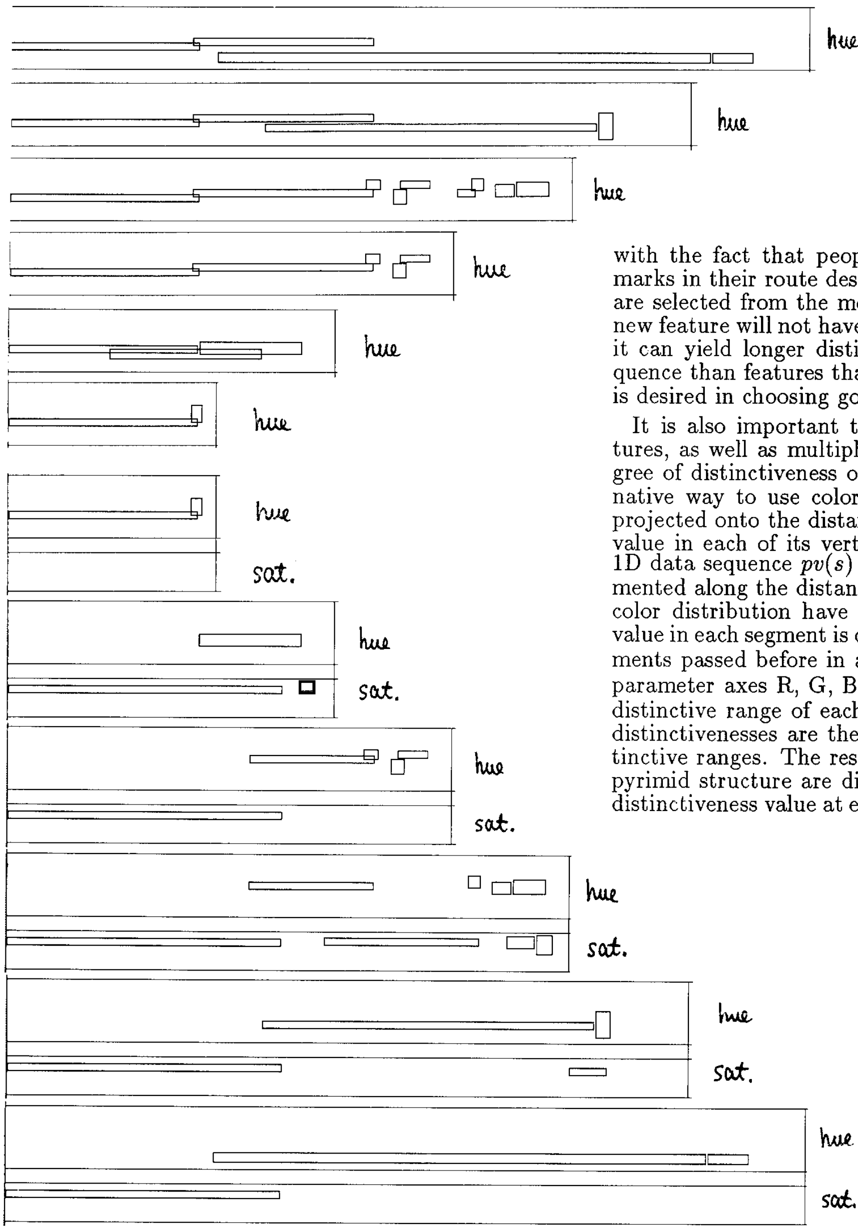

Fig.9 Landmark sequences selected from parameters of hue and saturation.

selection, working up the feature hierarchy described. This shall provide us with an even more robust representation.

The extracted final qualitative description depends highly on the features used. Some features are easier and more stable to extract than others, and thus play a larger role in landmark selection, which is consistent with the fact that people may choose different landmarks in their route descriptions. Since our landmarks are selected from the most unique scenes, increasing a new feature will not have influence on the result, except it can yield longer distinctive ranges in landmark sequence than features that have been considered, which is desired in choosing good feature.

It is also important to consider using multiple features, as well as multiple resolutions, to provide a degree of distinctiveness of landmarks. We try an alternative way to use color value of the PV. The PV is projected onto the distance axis by computing average value in each of its vertical line so that we obtain an 1D data sequence $pv(s) = [R(s), G(s), B(s)]$. It is segmented along the distance axis at the places where the color distribution have abrupt changes. The average value in each segment is compared with those in the segments passed before in a four dimensional PDS (three parameter axes R, G, B and one distance axis). Then distinctive range of each segment is computed. Their distinctivenesses are the volumes occupied by the distinctive ranges. The results on different resolutions as pyrimid structure are displayed in figure 10. It shows distinctiveness value at each position along the distance axis under the condition that the robot moves from left to right. The method is stable in yielding landmarks then the method described above, whose result is sensitive to the threshold when removing small clusters in PDS. On the other hand, the result of this method is not intuitive to varify and it is unable to find small regions distinct in attributes, because it averages the

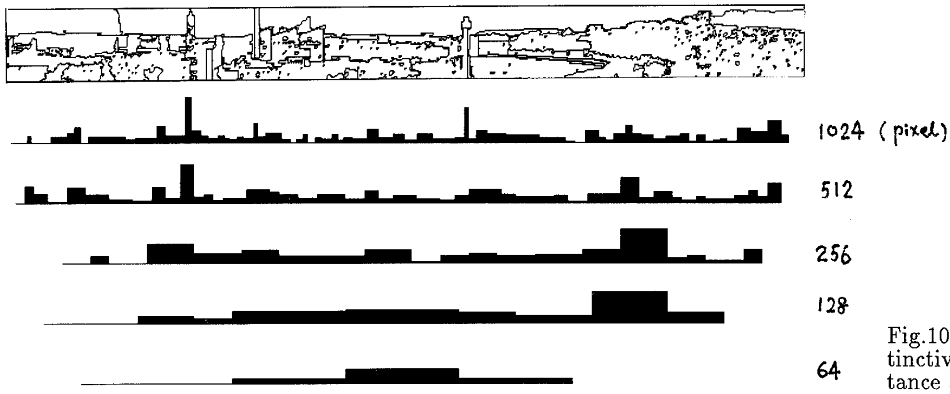

Fig.10 Distribution of distinctiveness along the distance axis.

value in each vertical line.

Because the robot can insure that there is no similar section in the distinctive range of a landmark selected in its trial move, the correspondence to a landmark found in a subsequent move is reliable and efficient. Since the landmarks have distinctive characteristics among the scenes along the route, it will be influenced less by certain tendencies in the changing environment, such as illumination.

Basically, the distinctive range expands from the current structure backwards to the first similar neighbouring structure. The expansion of rectangles for a less unique pattern will finish within a very local range. Therefore, the computation of distinctive characteristics is efficient because we compute relations of a pattern only with those near it both in distance and in parameter, instead of all of the patterns in the view space.

7. Conclusion

We have developed a method for a robot to autonomously create a qualitative landmark representation for use in long distance navigation. With this representation, a mobile robot can recognize particular locations along a route. The key task in this work is how to extract autonomously unique scenes, which are referred to as landmarks, from a panoramic representation. We have considered this task almost entirely from a bottom-up approach, since we want the robot to operate in many varied environments where a *priori* information can not be used.

Our method examines a PV and determines sections that are unique based on a number of features. For each feature under consideration, there is an associated distinctive range map. Since the distinctiveness of the landmarks are based primarily on relationship of patterns, the set of landmarks are influenced little by changes in the environment such as light intensity, as well as other parameters such as threshold values in segmentation and viewing distance. We have applied our algorithms to a three kilometer outdoor route, resulting in a robust qualitative landmark representation for our mobile robot.

Acknowledgement

The authors wish to thank Mr. Shigang Li for his contribution in preparing miniature model image.

References

[1] J.Y. Zheng, M. Asada, and S. Tsuji, "Color-based panoramic representation of outdoor environment for a mobile robot", Proc. 9th-ICPR, vol.2, pp.801-803, October 1988.

[2] J.Y. Zheng, Dynamic Projection, Panoramic Representation, and Route Recognition, Ph.D. thesis in Osaka University, Dec. 1989.

[3] J.Y. Zheng and S. Tsuji, "Panoramic representation of scenes for route understanding", Proc. 10th-ICPR, Vol.1, pp.161-167, June, 1990.

[4] J.Y. Zheng, and S. Tsuji, "Spatial representation and analysis of temporal visual events", Proc. IEEE Intl. Conf. Image Processing, Vol.2, pp.775-779. Sept. 1989.

[5] J.Y. Zheng, and S. Tsuji, "From Anorthoscope Perception to dynamic vision", Proc. IEEE Intl. Conf. Robotics and Automation, Vol.2, pp.1154-1160, May, 1990.

[6] G. Cohen, Memory in the Real World, Lawrence Erlbaum Associates, Hove. 1989.

[7] B. Bhanu, H. Nasr and S. Schaffer, "Guiding an autonomous land vehicle using knowledge-based landmark recognition", Proc. DARPA IUW, pp.432-439, 1987.

[8] D.T. Lawton, et. al., "Terrain models for an autonomous land vehicle", Proc. IEEE Conf. Robotics and Automation, pp.2043-2051, 1986.

[9] M. Barth, S. Hackwood, G. Beni, "An attentive sensing strategy for a multi-window vision architecture", in Proc. of SPIE/SPSE Symposium on Electronic Imaging: Science and Technology, Santa Clara, CA, Feb. 1990.

[10] M. Barth, Robotic Attentive Sensing and its Application to Integrated Circuit Wafer Inspection, Ph.D. thesis, University of California, Santa Barbara, August, 1989.

[11] M. D'Zmura, P. Lennie, "Mechanisms of color constancy" JOSA-A, Vol.3, No.10. pp.1662-1672, October 1986.

[12] J. Wang and G. Beni, "Connectivity analysis on multi-dimensional and multi-valued images", Proc. IEEE Intl. Conf. on Robotics and Automation, Raleigh, 1987.

[13] J.Y. Zheng, M. Barth and S. Tsuji, "Qualitative route scene description using autonomous landmark detection", Proc. 3rd-ICCV, pp.558-562, 1990.