

Exploiting Composite Features in Robot Navigation

Jennifer Davison, Kelly Hasler

Faculty Sponsor: Karen T. Sutherland,
Department of Computer Science

ABSTRACT

Humans are very good at picking out characteristics to use as landmarks in environments which have very few structured features such as points, lines, or corners. They use those landmarks to navigate or identify objects around themselves.

Our work focuses on the problem of deciding what features in a typical unstructured environment an automated navigation system can use as landmarks. Our first step was to create a library of digital images. Digital pictures were taken from different view points of the bluffs surrounding La Crosse. We then generated precise definitions for the landmark features, such as ridge lines, saddles, and valleys, found in such environments. We have isolated features within digital images and we have written algorithms that color the pixel values in the image to accent these land features. These will be used to formulate a decision making strategy to locate these features in other digital images. We have obtained electronic digital elevation maps of the La Crosse area from the U.S. Geographical Survey to use in matching these features to the map data.

INTRODUCTION

The purpose of the research was to identify in a computer program those composite land features that humans use to navigate. The major questions in robot navigation are “Where am I?” and “Where am I going?” Landmarks or land features identified in the robot’s view are matched to a map of the area and measurements to those landmarks are used to answer these questions. This process is referred to as *localization*. Most of the previous research using landmarks in outdoor environments has used point features, such as mountain peaks, as opposed to composite features [Sutherland 1994, Sutherland and Thompson, 1994, Murphy et al. 1997]. Point features are easier to identify but produce more error when used for localization than do composite features. Thus, this work addresses the question of “What is a composite feature and how can it best be used to localize?” To identify a composite land feature, a precise definition must be known. Since no clear, formal, specified definitions existed for any of these features, they had to be formulated.

By interpreting previous work in the use of composite landmarks in unstructured environments [Fuke and Krotkov, 1996, Thompson et al., 1996] and consulting with researchers in that area, the terms for such features and their definitions were developed. The features were separated into two categories: *Primitives* and *Formations*. The Primitives are themselves

composite while the Formations are groupings of Primitive composite features. Figure 1 summarizes the Primitives and Figure 2 summarizes the Formations. This terminology will not only be used in this work, but has been shared with the navigation research community at large, which was in need of a standard for describing natural outdoor land features. It was interesting to discover that researchers who referenced these features in their work were not clear as to exactly how the terms should be defined. Two of the Primitives, Bowl and Cirque, are shown in Figure 3.

Humans can visually pick out land features in the environment as well as in digital images. Since the differences in color or grey scale in the images represent a land feature, it follows that distinguishable patterns or groupings of pixel values in the images should represent these features. Due to the fact that the images to be analyzed were taken in the La Crosse area, and that the predominant feature in this area is ridge lines, we focused on isolating and identifying ridge lines. To find these ridge lines, algorithms were written to pick out the patterns of pixel values in a digital image which represented ridge lines.

Primitives

Bench: a terrace along the bank of a body of water, often marking a former shore line

Depression: an area of land that is lower than the surrounding area

Reentrant: an indentation between two salients (promontories) in a horizontal plane

Valley: a stretch of lowland lying between hills and mountains sometimes containing a stream

Basin: a large or small depression in the land with sides sloping inward to a central point

Bowl: a basin that's horizontal cross-section forms a two-dimensional closed figure

Cirque: a spoon shaped, steep-walled basin on the high slopes of mountain ranges that when clustered together form sharp jagged ridges and isolated peaks

Gully: a channel or hollow that water may drain through

Pass: a narrow opening or gap between peaks

Flat area: an area of land that is mostly level

Protrusions: land that juts up beyond the surrounding land

Mesa: a small high plateau or flat area with steep sides

Peak primitives: a crest or summit of a hill or mountain ending with a point

Ridge: the horizontal line formed by the meeting of two sloping surfaces

Water area: an area of water surrounded by land

Wall: an area of land that forms a barrier

Figure 1: Definitions of Primitive land features

METHOD

The first step was to take digital images of the bluffs surrounding La Crosse. An image library was created consisting of images taken at multiple locations and from different view-points at each location in order to clearly capture distinguishable land features within the bluffs. USGS (United States Geological Survey) elevation data for the La Crosse area (Quad Boundaries: SW 640893.50, 4845594.50; NW 640599.81, 4859478.00; NE 650642.81, 4859698.00; SE 650957.44, 4845815.00) was obtained and used to render a terrain map of the area, shown in Figure 4. The data used was 7.5 minute DEM (Digital Elevation Model) with readings taken every 30 meters. This allowed for simulated terrains to be rendered covering the same areas as the images in the library and for the elevation at any given point to be electronically available. Figure 5b shows a rendering of the elevation data for the same area in which the image in Figure 5a was taken.

The digital images were then converted into a form that allowed for easy access to the actual pixel values. They were cropped so that extraneous features in the image were eliminated and ridge lines were featured. The cropped images were then converted from color to grey scale, then into a matrix of numerical values with each value representing a grey scale pixel value, with values ranging from 0, representing black, to 255, representing white. These pixel values were then written into a text file. Visual patterns of pixel values were obvious in this text file. To more clearly see these patterns, an algorithm was developed to color different ranges of pixel values different colors. It was found that in each picture there was a unique number pattern that represented the ridgeline. The unique pattern consisted of a contrasting pattern of lighter number values located in close proximity to darker number values. Although ridge lines along the sky line were easy to find, these patterns were also apparent along ridge lines that did not border the sky.

Finally, pattern recognition algorithms were developed in an attempt to identify this unique pattern that represented the ridge line. The goal was to produce from the pixel intensity values a single entity which could be labeled "Ridge line". Although ridge lines on the map as well as on the scene rendered from the map data are shown as lines, they appear as blobs in the actual images. Once the blobs are identified, they must be compared in some way to the linear map data in order to match the view to the map.

Formations

Peak formation: an area of land consisting of more than one peak

Ridge line: an area of land consisting of more than one ridge and one or more peaks

Y-valley: an area of land consisting of more than one valley combined together

Figure 2: Definitions of Formation land features

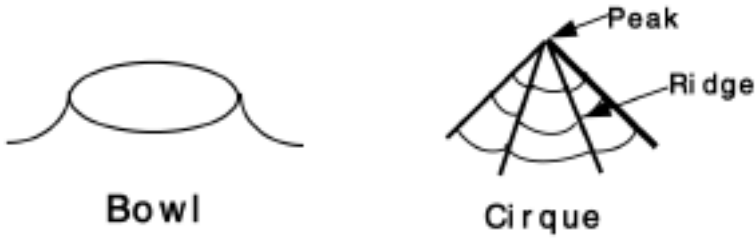


Figure 3: Two Primitives: Bowl and Cirque

The first algorithm pulled out pixels with values exceeding a given threshold. There were two problems with this approach:

1. When images are taken in an outdoor setting, there is no control over the light intensity. Images existed in the library with ridge lines visible to the human eye which had no pixel values over the threshold.
2. It was then necessary to add another level of processing to the algorithm in order to decide which of these thresholded pixels were part of the ridge line and which were outliers and to connect the legitimate pixels into a single entity.

At this point, it was decided that, to deal with the light intensity problem, instead of an absolute threshold, a difference measure should be used, pulling out pixels which differed from their surroundings by a percentage of the total range of pixel values. The second problem was dealt with by grouping pixels into a blob, using a standard region growing algorithm as described in [Faugeras, 1993]. A "Ridge line signature" was then developed for each connected blob for matching with the map ridge lines.

RESULTS

It was found that there is a unique pixel value pattern when a ridgeline appears in a digital picture and this pattern can be extracted using a pattern recognition algorithm.

After examining the U.S. Geological Survey data available, it was clear that the data is smoothed to the point that many land features commonly recognized by a human observer, such as cliffs and outcroppings of rock are lost in the smoothing. This smoothing is quite apparent in Figure 5b. The major ridge lines are clearly visible, but the rendering shows a much smoother terrain than does the image. Since the readings are taken at 30 meter intervals, much of the detail in the land features is lost. Although the U.S. Geological Survey data might be of more value if it included more data using smaller intervals than 30 meters, the additional data would then add to any required processing time. It should therefore be noted that using the USGS data for localization on a map will require that large scale features be utilized rather than some of the smaller distinguishable features so often used by humans. This is significant in that it has shown that preliminary work which depends on these features cannot assume that they will scale up.

Weather conditions can change the view significantly. For example, a snow covered area looks different than the same grass covered area. When the sky is overcast, the image is dark, leading to a smaller range of pixel values and less likelihood of picking up differences that signify ridge lines. A scaling of the pixel values can often widen the spread, but only if there are not too many outliers to skew the distribution. As an example, an image taken with an overcast sky might contain pixels in the range of 0 to 150. By scaling the pixel values, the range can be changed to 0 to 255, providing a larger difference in values along the ridge

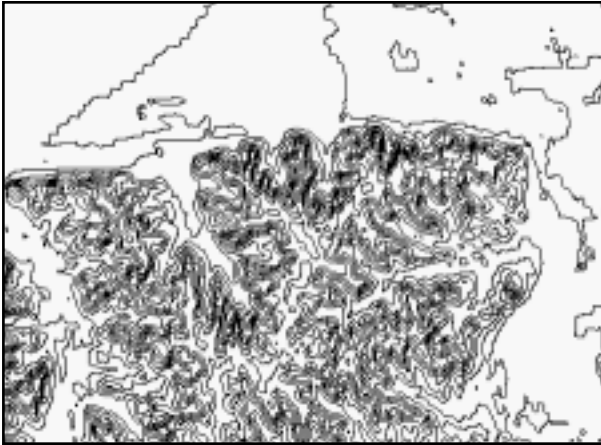


Figure 4: Rendered contour map of the La Crosse area.(⇒North)

lines. However, a patch of white in the image, such as a piece of pavement in the foreground, produces outliers and will skew the distribution and mitigate the effect of the scaling. These outliers can be identified and removed, but that takes valuable time. Vegetative growth can also change the look of the surroundings. A tree line may be mistaken as a ridge line when analyzing the digital image, which would throw off the pattern recognition algorithm. However, it should be noted that human navigators also occasionally mistake tree lines for ridge lines. It is questionable whether or not one should expect any better from an automated system.

Future work will include:

- Further testing of the ridge line signatures on additional data from the image library.
- Comparing localization using point features to that using ridge lines.
- The evaluation of additional composite land features and developing of a set of algorithms that can be used to incorporate their use into the localization process.

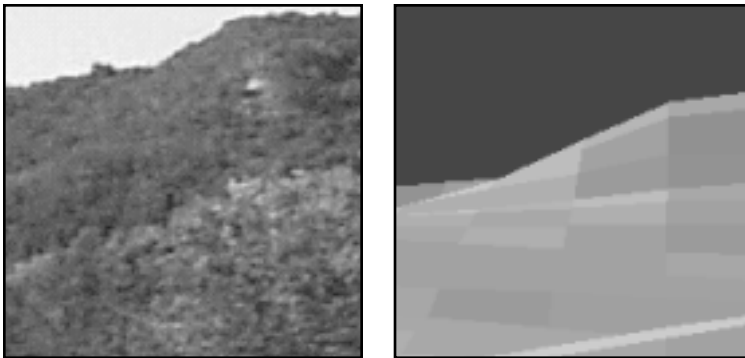


Figure 5: a) Image of local bluffs with b) matching USGS rendered data.

ACKNOWLEDGMENTS

We would like to thank The University of Wisconsin - La Crosse Undergraduate Research Program for funding this research. We would also like to thank the Computing Research Association's Collaborative Research Experience for Women (CREW) project for additional funding.

REFERENCES

- [Faugeras, 1993] Faugeras, O. (1993). *Three-Dimensional Computer Vision: A Geometric Viewpoint*.
- [Fuke and Krotkov, 1996] Fuke, Y. and Krotkov, E. (1996). Dead reckoning for a lunar rover on uneven terrain. In *Proceedings 1996 IEEE International Conference on Robotics and Automation*.
- [Murphy et al., 1997] Murphy, R., Hershberger, D., and Blauvelt, G. (1997). Learning landmark triples by experimentation. *Robotics and Autonomous Systems*, 22.
- [Sutherland and Thompson, 1994] Sutherland, K. and Thompson, W. (1994). Localizing in unstructured environments: Dealing with the errors. *IEEE Trans. on Robotics and Automation*, 10:740—754.
- [Sutherland, 1994] Sutherland, K. T. (1994). Ordering landmarks in a view. In *Proc. DARPA Image Understanding Workshop*.
- [Thompson et al., 1996] Thompson, W. B., Valiquette, C. M., Bennett, B. H., and Sutherland, K.T. (1996). Geometric reasoning for map-based localization. Technical report, University of Utah.