IMPROVING THE RECOGNITION OF GEOMETRICAL SHAPES IN ROAD SIGNS BY AUGMENTING THE DATABASE

Padmaja Dukkipati
The University of Texas at Tyler
Department of Computer Science
Tyler, TX, 75799
sdukkipati@yahoo.com

Leonard Brown
The University of Texas at Tyler
Department of Computer Science
Tyler, TX, 75799
Leonard_Brown@uttyler.edu

ABSTRACT: *Improving safety is a key goal in autonomous road vehicles. Driver support systems that help drivers react to changing road conditions can potentially improve safety. As with any vehicle, autonomous vehicle driving on public roads must obey the rules of the road. Many of these rules are conveyed through the use of road signs, so an autonomous vehicle must be able to detect and recognize the road signs and change its behavior accordingly. This implies that the system must be able to detect a real world road sign and match its image to images that are already present in its underlying database. In order to be effective it is critical that the system is able to perform this matching accurately. The ability to match a picture of a real world image with images already in the database based on visual characteristics is called Content-Based Image Retrieval (CBIR). This paper proposes a method for improving the accuracy of CBIR systems by augmenting their underlying databases.*

Keywords: *Content-Based Image Retrieval; Database Management Systems; Similarity Search*

1. INTRODUCTION

Fatal car crashes occur every day around the world. An estimated 30% of the fatal car crashes can be attributed to the driver inattention and fatigue [1]. This can be reduced by autonomous vehicles which work in part by automatically recognizing road signs placed near streets and highways. The roadway is well structured, and the appearance of the road signs is highly restricted. Each type of sign must be of a particular size, color and shape with few exceptions. Thus, the ability to correctly identify the size, color, and shape of an object is useful when attempting to automatically recognize an image of a given road sign.

Usually signs are upright along side the road. Unless tampered by an accident or by other environmental conditions, signs will appear approximately orthogonally to the road direction. Still, even with the well structured road system, there are problems associated with automatically identifying road signs that fall into two types, detecting the sign and recognizing the sign. The detection of road signs can be a problem in the areas of temporary work and uneven ground surface. The recognition can be problem when signs are faded because of rain, sun etc. or may be covered with tree leaves so that the sign is not visible. In addition, with minor accidents, signs can appear tilted.

Despite the numerous problems that can hamper the recognition of road signs, it is imperative that an autonomous vehicle system does not fail to detect the presence of a sign. Even failing to recognize one sign can be very dangerous. In order to detect the signs correctly to avoid accidents, the recognition system should retrieve them accurately. In this paper, we present our approach for improving the retrieval accuracy of road sign images for such a system.

The remainder of our paper has the following structure. Section 2 describes different image retrieval systems that match images based on their visual characteristics. Section 3 gives our proposed approach for improving the ability to recognize shapes in such systems. Section 4 presents the results of a performance evaluation measuring the effectiveness of our proposed approach. Finally, Section 5 summarizes our work and presents directions for future research.

2. RELATED LITERATURE

A system which assists a driver by alerting or displaying the road signs on the dash board is known as a Driver Assistance System (DAS). To illustrate how a DAS may work, consider the DAS developed by the National Information Communications Technology

Australia (NICTA) [2]. The DAS has three cameras. One camera scans the road ahead, and the other two cameras, called the gaze monitoring pair, is set on both sides of the instrumental panel on the dashboard to monitor where driver is looking. The images from the cameras are fed to a computer system fitted behind the dash. The Face lab [3] package is used to analyze images from the stereoscopic cameras and determine where the driver is looking. The software scans the video images and detects road signs by recognizing their symmetrical shapes. Once the sign is detected, the image is compared to a list of signs stored in the computer's database. If it recognizes a stop sign, the computer checks if the car is slowing down, if not, then it alerts the driver.

The above discussion indicates that the DAS operates by repeatedly obtaining an image and forming a query that requests all images stored in the database that are similar to the input image, which is called a similarity search. The reason is that is expected that the query image will not be an exact match of one of the database images. Instead, the system needs to evaluate the similarity between the query image and the images in the database.

Retrieving images by evaluating the similarity of their visual characteristics is known as Content-Based Image Retrieval (CBIR). One of the best known and representative commercial CBIR systems is the QBIC system developed by IBM [4]. Other well known commercial CBIR systems include Virage [5] which is used by Alta Vista to facilitate image searching and Photobook [6] which is a representative research CBIR system developed at MIT. Similarity between images in these types of systems is often computed by comparing low level visual features that must be extracted from each image involved in the comparison. Shape is one of the key visual features used by distinguishing visual data along with other color and texture [7]. Of these features, shape is often the most important because in some applications only shape is needed. Although color is also important when recognizing road signs, recent research [2] considers shape to be the main feature in recognition. Generally shape is classified into two categories which are Contour-Based shape descriptors and Region-Based shape descriptions [7]. Contour-Based shape descriptors are shape features extracted using only the contour of a shape. These descriptors only employ shape boundary information and capture shape boundary features. Region-Based shape descriptors are shape features are extracted from the whole shape region. These descriptors make use of all the pixel information across the shape region.

When a CBIR system processes a query of the type "*Retrieve all images in the database that match this query image*", it extracts and attempts to match the visual features extracted from the query image to collections of visual features previously extracted from the database images. Each feature extracted from the query image is compared with the same type of feature obtained from the database images. The comparison between the query image and a database image can be performed using distance functions like the Euclidean distance. In such a function, the distance between two features is indirectly proportional to the similarity of the features. Thus, two input images with features that have a distance of 0 are considered to be identical. For a system to determine if two images are similar, then, it needs to utilize some threshold value that is compared with the output value of the distance function. If the output value is below the threshold, then the input images are considered to match.

The main problem with CBIR is that it solely depends on the extraction and comparing of primitive visual features which have no understanding of the image's semantic contents. So, even though the CBIR system is able to identify the low level feature similarities between it and images in the database, it may still fail to recognize similarities in semantic content between the database images and the query image.

In order to improve the accuracy of CBIR systems, several approaches have been developed. Often, these approaches involve changing the procedures used to extract low-level features from images or changing the function used to determine similarity between extracted feature sets. The problem with these improvements is that they cannot be applied to existing CBIR systems. Thus, the internal functions of a system must be rewritten in order to implement the changes proposed by these approaches. Another type of approach for improving the accuracy of a CBIR system involves processing multiple queries. Such multiple querying approaches including [8, 9] utilize several query images for a single query by submitting each one to the database separately. The responses obtained from processing each of the query images are combined to give a cumulative result. The problem with this approach is that the desired image features must be extracted from each query image, which increases the time needed to process the entire query.

As an alternative to the previous approaches, our proposed method can improve the accuracy of a CBIR query without requiring multiple query images to be submitted to the system. This approach, called database augmentation [10, 11] will be presented in the next section.

3. SHAPE BASED RETRIEVAL AND AUGMENTATION

Our approach for improving the accuracy of a CBIR system operates on the principles presented in [12]. Specifically, we consider that minimizing the number of occurrences of false negatives when retrieving images is often considered more important than reducing the number of false positives. The reason is that users can filter out any images retrieved by the database system that are unwanted, but they have no way of knowing the existence of a matching database image that the system failed to retrieve. So, our approach focuses on reducing the number of false negatives at the expense of possibly generating new false positives.

Our proposed approach for improving the accuracy of CBIR systems addresses the problems of feature matching by augmenting the underlying database with new images created by editing the original images already present. For each image object z in the database, the system will store z along with a set of images created by modifying z. Augmenting the database in this manner can improve the accuracy of CBIR systems in any situation when a particular database image, say x, is expected to be retrieved in response to a query image q, but the features extracted from q and x do not sufficiently match. The idea is that the features of q may sufficiently match x', where x' is a slightly modified version of x. By maintaining a connection between x and x' in the database, the system can use the fact that the features of q and x' match as the basis for returning x in response to the query, even though the respective features of q and x do not sufficiently match. Thus, the accuracy of the CBIR system can be improved without having to change how the system extracts features from the images and without having to submit multiple query images to the database.

The critical part of the above approach for database augmentation involves determining how to generate a new image x' from a database image x. We have developed a set of rules that govern how new images should be created for CBIR systems that retrieve road signs based on geometric shapes using Region-Based features. Specifically, our rules are designed for systems that extract a Freeman Code or Chain Code [7] from each image in order to represent the shape of its objects. The Freeman Code is a popular tool for representing the shapes of objects in images by describing each shape as a sequence of unit-size line segments with a given orientation.

Each of our rules for augmentation are in the form of *if (condition) then (operation).* Each condition is a simple test applied to each image, x, as it is inserted into the database. If the image meets the given condition, then a new image is formed by applying the associated editing operation to image x. Both x and the new image derived from x are then inserted into the database.

Our first rule is if the major axis of a shape is not parallel to the y-axis, then the image is rotated so that those axes are parallel. The purpose of this rule is that to ensure that the query image will be oriented in the same direction as the images in the database. Although chain codes can be normalized to account for variations in orientation, we have found that it is not always effective, especially when the difference between the orientations of two images is not a multiple of 90°. The second rule is if the image's shape has less than 5 different chain-code sequences or the sides are equal then the image should be merged with its copy. The purpose of this rule is that when the road is uneven, signs with simple shapes may give an impression of two images merged together. The third rule is if the intensity values of the interior of the shape are different from the intensity values of the shape's boundary, then the shape should be filled with the boundary color. The purpose of this rule is that we noticed that images containing shapes with thick boundaries did not always successfully match images containing similar shapes with thin boundaries or similar shapes that were filled. Finally, our fourth rule is that if the chain code has many irregular sequence numbers then the image should be stretched in order to eliminate the irregularities. The purpose of this rule is that images that are partially occluded or noisy may generate shapes of objects that are not smooth. These rules are summarized in the Table 1.

4. PERFORMANCE EVALUATION

This section describes the result of a preliminary performance evaluation testing the effectiveness of our proposed augmentation rules. To perform this evaluation, we utilized the data set of road signs obtained from Street Signs USA web site [13]. This data set of 125 images is divided into five categories, Diamond, Rectangular, Triangular, Circular, and Other. We loaded these images into an Oracle 9i database management system using their built-in functions to extract the low-level visual features from each image. These features were stored in the database along with the 125 images.

Rul e	Condition	Operation
1.	If the major-axis is not parallel to the y-axis,	Then the image should be rotated, so that the major axis is parallel to y-axis
2.	If the image's shape has less than 5 different chain-code sequences or the sides are equal ,	Then the image should be merged with its copy.
3.	If the intensity values of the interior of the shape are different from the intensity values of the shape's boundary,	Then the shape should be filled with the boundary color
4.	If the chain code has many irregular sequence numbers,	Then the image should be stretched

Table 1. Shape-Based Augmentation Rules

Our test involved submitting several shape-based similarity searches to the data set using Oracle 9i's built in image matching functions. We used the database images as query images as well during these tests. We varied the threshold value used to determine similarity by Oracle from 10 to 90 and executed similarity searches using all images from each category. For each query, we recorded the system's precision and recall. Afterwards, we manually applied our augmentation rules to the images in the data set and repeated the queries. The augmentation process added 110 new images to the database, almost doubling the total number of images.

The results of our tests are displayed in Table 2, and they are as follows. During the tests before augmentation, the query images usually returned only identical images for the smaller threshold values 10 and 20. Thus, those queries exhibited poor recall. Alternatively, for threshold values 70 and 90, many false positives were retrieved by the database. Thus, those queries exhibited poor precision.

Threshold Values	Average Recall Before Augmentation	Average Recall After Augmentation
10	0.08	0.08
20	0.1	0.12
30	0.19	0.22
50	0.46	0.49
70	0.72	0.74
90	0.89	0.96
Average	**0.41**	**0.43**

Table 2. Performance Evaluation

After augmentation, our tests indicated that our approach does reduce the number of false negatives generated from the image matching function of Oracle. The results comparing the average recall before and after augmentation are displayed in Figure 2. We computed the average recall gain of our approach as (Average Recall After Augmentation – Average Recall Before Augmentation)*100 / Average Recall Before Augmentation. This gave us a recall gain of 4.9%.

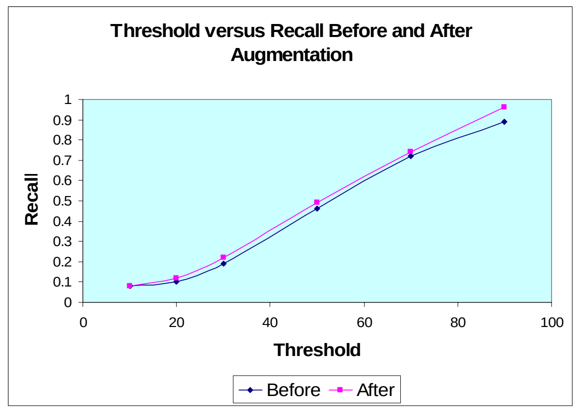

Figure 1. Threshold vs. Recall

The reduction of the number of false negatives as a result of augmentation came at the cost of an increase in the number of false positives. The comparison of the precision before and after augmentation is displayed in Figure 2. The loss in average precision was computed as (The Average Precision Before Augmentation – The Average Precision After Augmentation)*100 / Average Precision Before augmentation. This yielded a precision loss of 8.3%.

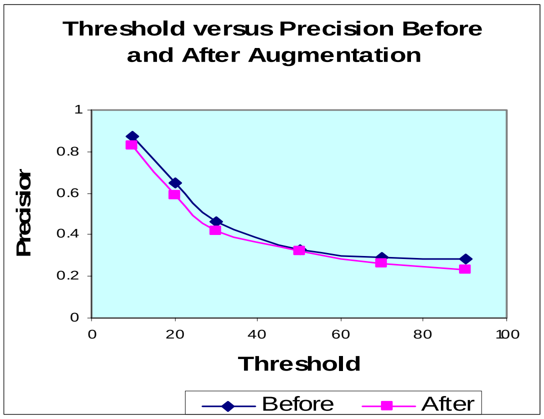

Figure 2. Threshold vs. Precision

5. CONCLUSION

Road signs are designed and positioned in the environment clearly in order to easily detect and recognize them. Still, there is a slight chance of failing to recognize a road sign which can be very dangerous. The augmentation procedure has shown that it can reduce the number of times such a system fails to correctly identify a road sign image.

Our future work involves implementing our rules in MATLAB so that they can be applied automatically to a given database image. This will allow us to individually test the effectiveness of each rule in order to determine if one rule outperforms another.

References

[1] T. Williamson and C. Thorpe, "*A trinocular stereo system for high-way obstacle detection*", 1999 International Conference on Robotics and Automation (ICRA '99), IEEE, 1999.

[2] URL www.cnn.com/2004/TECH/10/06/ roadsign.recognition, last accessed Oct 7, 2004.

[3] N. Apostoloff and A. Zelinsky, "*Vision In and Out of the Vehicles: Integrated Driver and Road Scene Monitoring*" The International Journal of Robotics Research, Vol. 23, No. 4-5, pp. 513-538, 2004.

[4] W. Niblack, et al., "*The QBIC Project: Querying Image by Content Using Color, Texture and Shape*", Proc. SPIE Storage and Retrieval for Image and Video Databases, Vol. 1908, pp. 173-187, 1993.

[5] J. Bach et al, "*The Virage Image Search Engine: An Open Framework for Image Management", *Proc. of SPIE, Storage and Retrieval for Still Image and Video Databases IV, pp. 76-87, 1996.

[6] A. Pentland et al., "*Tools for Content-Based Image Databases*", International Journal of Computer Vision, 18(3):233-254, 1996.

[7] R. Gonzalez and R. Woods, Digital Image Processing, Addison-Wesley, 1992.

[8] L. Tamine, C. Chrisment, M. Boughnem, "*Modelling Vagueness and Subjectivity in Information Access*", Information Processing and Management: an International Journal, Vol. 39, Issue 2, pp. 215-231, March 2003

[9] S. Tahaghoghi, J. Thorn, and H. Williams, "*Are Two Pictures Better Than One*", Proc. of the 12th Australasian Conference on Database Technologies, Queensland, Australia, pp. 138-144, 2001.

[10] L. Brown, "*Issues in Augmenting Image Databases to Improve Processing Content-Based Similarity Searches*", Proc. of the 20th Annual ACM Symposium on Applied Computing, Santa Fe, New Mexico, pp. 1254-1255, 2005.

[11] L. Brown and L. Gruenwald, "*Performing Color-based Similarity searches in Multimedia Database Management Systems Augmented with Derived Images*", Proc. of the 21st British National Conference on Databases, Lecture Notes in Computer Science, Vol. 3112, Springer, pp. 178-189, 2004.

[12] A. Gupta and R. Jain, "*Visual Information Retrieval*", Communications of the ACM, Vol. 40, No. 5, pp. 71-79, 1997.

[13] URL http://www.street-signs.usa.com, last accessed May 1, 2005.