Sidewalk Following Using Color Histograms

John S. Seng
Dept. of Computer Science
California Polytechnic State University
San Luis Obispo, CA 93407
jseng@csc.calpoly.edu

Abstract

In this paper, we present the development of an algorithm for sidewalk following for a mobile robot. The algorithm classifies pixels of a video frame as either sidewalk or background. The classification of pixels is performed using color histograms of the hue and saturation of various areas of an image taken by an on-board camera. This color histogram approach allows the algorithm to detect and follow concrete paved areas of arbitrary shape and color. The development of the algorithm is performed using a robotics research platform developed by undergraduate students at Cal Poly State University, San Luis Obispo.

1. Introduction

Road following is an important task in the field of mobile autonomous robotics. The ability to detect the road area in front of a robot via computer vision is a necessary step in achieving autonomous navigation in outdoor environments. The task of road following is to detect which areas in front of a robot are paved road areas (e.g. sidewalk and asphalt) and which areas are not (e.g. grass, bark, or buildings). This task is complicated by the fact that road areas can often appear as irregular shapes and varying colors. Once a robot can detect what area in front of it is road, it can plan which direction it should take to move forward.

In this paper, we focus on the task of sidewalk following, a specific form of road following. We study specifically the task of visually detecting and following concrete sidewalk areas. We will describe a sidewalk following algorithm that is under development for an autonomous robot platform. The sidewalk following algorithm processes monocular video to generate color histograms of the current road color. As the robot drives, the color histograms are used in classifying road pixels, and these pixels are used to update the sidewalk histograms in subsequent video frames. The algorithm is based on the principle that the color of sidewalk area in front of a robot can be represented with a color histogram and thereby can be distinguished from the surrounding background. The goal of this algorithm is to detect regions in front of the robot which are either sidewalk or unobstructed pavement areas.

There are several approaches to detecting sidewalk areas in video images. In this work, we have opted to utilize a pixel-based approach because a sidewalk area may not necessarily have a straight linear shape, and we desired an algorithm which can handle regions of arbitrary shape. Our algorithm does not look at particular features in an image, such as edges.

A sidewalk detection algorithm provides guidance to the navigation system by providing information on how the sidewalk surface changes ahead. Such an algorithm may be used to detect obstacles, but in this work we assume that close range obstacle detection will be performed by ranging sensors (either laser or sonar). Thus, we assume that the area directly in front of the robot is free from obstacles.

This work is part of a larger project which is the creation of a robotics platform developed by Cal Poly undergraduate students [5]. The students have developed this platform over the course of a 1 year period as part of a senior design project in the computer engineering program. A primary goal for this platform is the autonomous navigation of the Cal Poly campus. The platform consists of a motorized wheelchair base controlled by an on-board dual CPU core motherboard running Linux. The robot has a single on-board NTSC camera as video input which provides color images at 15 frames per second. All of the test video for this work is taken from the robot camera as the robot is driven by an operator, and the video is processed offline.

This paper is organized as follows: Section 2 covers related prior work. Section 3 discusses and describes the sidewalk detection and following algorithm. Section 4 describes future work. Section 5 concludes.

2. Related Work

Our algorithm is primarily based on the algorithm presented by Tan et al. [7]. The authors of that work describe a color histogram algorithm to detect and follow road areas in video taken from a moving vehicle. Our algorithm adopts several of the features and characteristics of the algorithm described by Tan et al. The algorithm presented in that work is based on color histograms of normalized red and green color values. In contrast, our work uses hue and saturation values, uses additional contours to determine the road and background regions, and uses a dynamic secondary histogram area for training.

There are several varying approaches used to detect sidewalk and road paths from monocular video. These techniques include: modeling the road areas as Gaussian distributions of pixel color, supervised training of a classifier, and edge-based techniques.

A road following algorithm developed by Dahlkamp et al. [3] uses Gaussian color models to model the road color. That work describes an algorithm for building multiple Gaussian road models. These color models are updated from a given frame to the next using a training area that is free from obstacles. The particular algorithm studied in that paper does not maintain a color model of the visible area which is not road.

A benchmark road follower system is the SCARF (Supervised Classification Applied to Road Following) system [2]. This system also uses Gaussian color modeling to distinguish road and background areas. This work assumes that the road ahead is linear in shape and that the width of the road is known. In our work, we do not place any assumptions on the road shape except for what is required to achieve proper training of our algorithm.

A road detection system described by Lombardi et al. [4] utilizes edge detection along with road shape model matching to detect road areas in an image. The approach in that work first uses an edge detection operator to group large image homogeneous regions. After the regions are detected, a Bayesian classification process is used to match the road to one of multiple road shape models. Their work has shown to detect road areas in the presence of vehicles and occluded regions. In our work we focus only on pixel-based detection and do not perform any road shape model matching.

Supervised classification of terrain is another technique which can be used to perform road detection. Such an algorithm is described by Angelova et al. [1] In that work, the authors capture texture, color histograms, and color information from a number of sample terrain image patches. These images patches are used to train a classifier. The classifier is then used to detect road areas in images. The classifier is also capable of classifying grass, gravel, and brush areas. In our work, we focus on unsupervised detection which is based only on the assumption that the area ahead of the robot is sidewalk.

3. Sidewalk Following Algorithm

The sidewalk following algorithm we study processes monocular video taken from the camera on the robot. Video frames are initially captured using an on-board NTSC camera with a resolution of 480 lines. In order to reduce computation, the images are then converted to a lower resolution of 320x240. The algorithm processes each frame individually using color histograms obtained from prior frame. 4 color histograms are maintained which model the current sidewalk and 1 color histogram is used to represent the surrounding background.

In this work, we refer to sidewalk as the areas of concrete that the robot is capable of driving on and background as obstacles that the robot cannot drive on.

3.1 Sidewalk and Background Histograms

The first step in determining the sidewalk and background areas is to generate color histograms of both regions. These color histograms are used to model the general color distributions of the sidewalk area that the robot is currently driving on and also the distribution of color of the current background. The advantage of using color histograms is that they can be quickly obtained, provide a good model of a sidewalk surface, and can be quickly compared.

In order to enable the initial sidewalk detection process, our algorithm is based on the assumption that the area immediately in front of the robot is a sidewalk area. As the algorithm proceeds, it can detect obstacles that are at a distance from the robot, but not obstacles immediately in front of it. Ranging sensors can be used to detect nearby obstacles, so we assume that this areas is initially free of obstacles.

When the algorithm initially starts, it is not known which areas of the video image are sidewalk and which are not. Because of this, a rectangular region (200x80 pixels) directly in front of the robot is used as a training area and assumed to contain sidewalk and is free of obstacles. The pixels in this training area are used to build a training histogram of the sidewalk in the current video frame. A 2-dimensional color histogram is generated using the pixels within this training area. In our work, we have found that using a histogram in the HSV colorspace provided good results. The histograms we use contain 10 bins in each dimension, and we use the 8-bit hue and saturation for each dimension. These 2-dimensional histograms provide the basis for classifying the sidewalk and background areas.

The histogram obtained from the training area is used in later stages to update 1 of the 4 sidewalk histograms. These 4 sidewalk histograms are the actual histograms used to segment the image into sidewalk and background areas. The segmentation is not directly performed using the training histogram.

A separate histogram called the background histogram is generated by randomly sampling pixels from the areas of the image which are not classified as sidewalk in the prior video frame. We empirically chose to sample 3000 pixels as this provides good results while allowing the algorithm to run quickly. Initially, the background area will be all the pixels not in the training area, and this may include pixels which represent the current

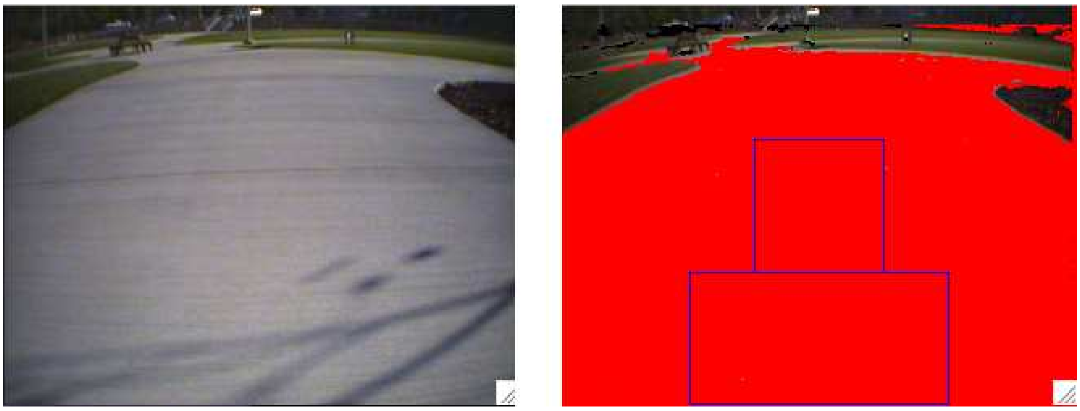

Figure 1. The image on the left shows a sample image taken from the on-board camera. The image on the right shows the classified sidewalk pixels in red.

sidewalk. This assumption is sufficient as the areas classified as sidewalk and background areas will shift as the algorithm proceeds. As more and more of the sidewalk pixels are actually classified as sidewalk, fewer of the misclassified background pixels will be included in the background histogram.

Once both the training histogram and background histogram are collected, they are normalized (all bins values are between 0 and 1). At the end of this step, the algorithm has a histogram of the sidewalk area immediately in front of the robot as well as a histogram of the pixels that are not sidewalk.

3.2 Sidewalk Classification

Once the training and background histograms have been acquired, the next step in the algorithm is to classify each pixel in the current frame as either sidewalk or background. This process is performed by comparing the frequency that a particular HS (hue-saturation) value appears in a sidewalk histogram as opposed to the background histogram. Colors that have a high occurrence in the sidewalk color histogram and have a low occurrence in the background color histogram are likely to be those pixels which actually represent the sidewalk. Conversely, colors that have a low occurrence in the sidewalk color histogram and have a high occurrence in the background histogram are likely to be those of pixels which do not represent sidewalk or shadowed sidewalk. Those pixels should accordingly be classified as background.

Individual pixels are classified as sidewalk and non-sidewalk areas using a probability metric similar to the technique described by Tan et al. [7] The probability that a pixel (given its hue and saturation) is a sidewalk pixel is given by: $P_{sidewalk} = S(h, s)/(S(h, s) + B(h, s))$.

$S(h, s)$ is the normalized frequency of a hue and saturation value from a particular sidewalk histogram. $B(h, s)$ is the normalized frequency of a hue and saturation value from the background histogram. If a particular color appears frequently in a sidewalk area but appears rarely in the background, then $P_{sidewalk}$ for that pixel will have a value close to 1. If a color appears equally in the sidewalk and background histograms, then its $P_{sidewalk}$ value will be .5. In order to effectively utilize all 4 stored sidewalk histograms, $P_{sidewalk}$ is computed for all 4 sidewalk histograms for each pixel and highest value of the 4 probabilities is used as the sidewalk probability for that pixel.

If the sidewalk probability for a pixel exceeds a threshold, then the pixel is classified as sidewalk. Those pixels which have a probabilities that are lower than the threshold are classified as background and are not considered a surface which the robot can drive on.

At the end of this step, the algorithm produces a binary image containing pixels classified as sidewalk and background.

3.3 Histogram Updating

As the algorithm progresses, the sidewalk and background color histograms are updated for each frame. Once the histogram of the training area has been collected, this training histogram is compared against each of the 4 current sidewalk histograms. If the difference between the training histogram and 1 of the 4 sidewalk histograms is within a threshold (computed as a sum of squared differences between the 2 histogram bin counts) and is the smallest difference out of the 4 histograms, then the matching histogram is updated with the training histogram. An update in our algorithm is a simple replacement of the matching histogram with the training histogram.

In the case that there is no matching histogram, this means that all of the sidewalk histograms are not similar to the training histogram. When this occurs, the sidewalk surface in front of the moving robot is beginning to change in color distribution. The algorithm policy is to then replace the histogram that had the the smallest number of pixel matches. The number of pixel matches for a histogram is the number of sidewalk pixels in a frame that matched best with that particular histogram. The histogram is replaced with the training histogram. This

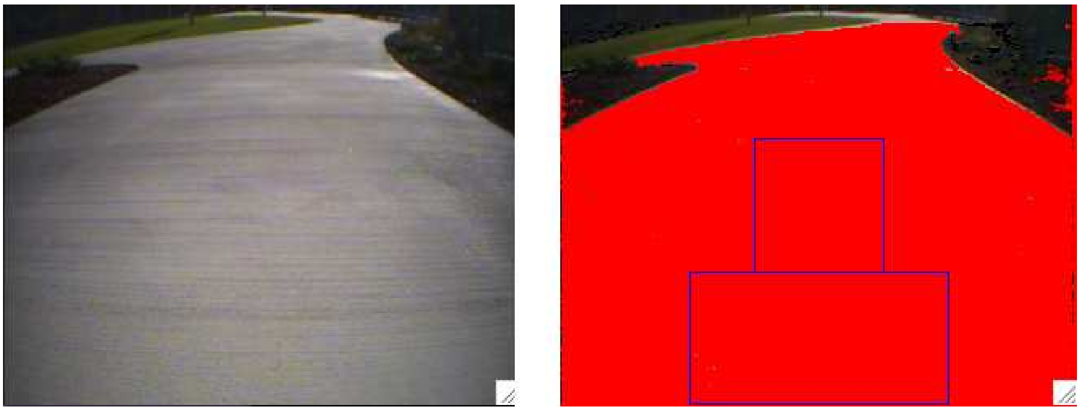

Figure 2. This is another example image taken the from test video sequence. The image on the left shows a sample image taken from the on-board camera. The image on the right shows the classified sidewalk pixels (in red).

replacement policy eliminates those histograms which contribute little to identifying sidewalk pixels.

The background histogram is updated by using a random sample of the pixels that are not classified as sidewalk. In addition to using only non-sidewalk pixels, we only select pixel which are not in a closed contour defined by the currently detected sidewalk.

In order to incorporate a larger area for training, we introduce a second training box when there is a high number of pixels classified as sidewalk in this new training box. This secondary training box is used for training if over 50% of the box is classified as sidewalk from the previous frame. Adding this second training box allows a larger training area when the area inside the second training box is similar to the initial training area. When the areas in the boxes are similar, then most of the area in the second box will be classified in a similar manner to the area in the first box. Including the pixels in the second training box is dynamic and is enabled based on the results of the prior frame.

3.4 Algorithm Performance and Analysis

In order to test our algorithm, we capture a video sequence using the robot's on-board camera and store the video to a digital camcorder. This provides an accurate representation of the colors that the robot would perceive. The robot is driven by a human operator on a sidewalk surface which contains irregular sidewalk intersections (not 90 degrees). The video is taken on a sunny day and contains shadows from trees, buildings, and the robot itself.

The algorithm provides relatively high performance averaging approximately 20 frames per second on a 2.16GHz Intel Core 2 Duo processor.

In this section, we provide a qualitative discussion of algorithm performance when running through the test video recorded while the robot was driven on campus. Figure 1 demonstrates the algorithm classifying a sidewalk area on the Cal Poly campus. The image on the left shows a sidewalk area in front of the robot with grass on the left and bark chips on the right. The sidewalk is not well structured and contains an intersection connecting another sidewalk path on the right. The image on the right shows the region of the image which is classified as sidewalk (draw in red). The lower blue box shows the primary training area while the upper blue box shows the smaller secondary training area. The upper blue box is visible which shows that over 50% of the pixels were classified as sidewalk in the previous frame. Figure 2 shows a similar video frame where the algorithm is able to detect a significant number of sidewalk pixels.

Figure 3 shows a video frame where the algorithm performs poorly. In cases where the robot's shadow is visible, the algorithm may or may not classify the shadowed area as sidewalk. In Figure 3, the fraction of shadowed pixels in the background exceeds the fraction of shadowed pixels in the sidewalk area.

The histogram replacement algorithm works well in that the histogram with the lowest number of matches is replaced when the sidewalk surface is rapidly changing. Histogram replacement occurs primarily when driving over a large sidewalk crack or when there are diffuse shadows from trees and bushes.

We found that 4 sidewalk histograms were sufficient to capture the color variation of the sidewalk. Under bright lighting conditions, 2 histograms were sufficient to classify the majority of sidewalk pixels. The average fractions of sidewalk pixels classified by the 2 best matching histograms were 50.04% and 46.38%. In shadowed areas, there was less color variation and most of the sidewalk could be represented with just 1 histogram (86.71% on the average). In transitional areas, the distribution was relative equal across the 4 histograms.

4. Future Work

One primary limitation of the current algorithm is that it requires the training areas to be fixed in location. It may be advantageous to provide training areas which can dynamically move in position as lighting conditions

Figure 3. In this image, the algorithm incorrectly ignores the shadowed area. The number of shadow pixels in the background exceeds the number of pixels in the training area. This leads to incorrect classification.

change. This may aid in the detection of sidewalk areas that are covered with sharp shadows.

The current algorithm implementation does not contain any aspects of supervised training. Given that a goal of the robot is to traverse a university campus, the sidewalk areas it will cover will be relatively the same color given a particular campus location. This information can be used to train a classifier and can be used in conjunction with the algorithm described here to achieve better sidewalk detection.

Additionally, the current implementation of the algorithm does not incorporate any edge or texture features into the classification of road areas. When viewed against neighboring surface regions, sidewalk areas typically exhibit a strong edge. Incorporating edge information may provide more robust sidewalk detection. In addition, sidewalk area have a distinct texture. In our work, we did not consider texture information because of the additional computation required, but using texture cues across small image patches may increase detection accuracy. Other research has shown how terrain can be successfully classified using texture alone on grayscale images [6].

5. Conclusion

This paper describes an algorithm used to detect and follow sidewalk areas in a monocular video sequence. The algorithm uses 2-dimensional histograms of the hue and saturation values of pixels to model sidewalk and non-sidewalk areas in an image.

A training area directly in front of the robot is used to build the initial color histogram of the sidewalk. This area is assumed to be free from obstacles and contains pixels which are representative of the current sidewalk. This training area is used to generate a 2-dimensional histogram of the hue and saturation values of the sidewalk pixels. The training histogram is used to update 1 of the 4 histograms which are used to model the sidewalk in front of the robot.

One additional histogram is used to represent the background area in a video frame. This background histogram is acquired by randomly sampling pixels which were not classified as sidewalk in a previous frame. This histogram provides a model of the color distribution of the current surroundings.

Individual pixels are classified as sidewalk if a pixel's hue and saturation appear more frequently in the sidewalk histograms as opposed to the background histogram.

This algorithm has been developed in conjunction with an autonomous robot design project at Cal Poly State University.

References

[1] A. Angelova, L. Matthies, D. Helmick, and P. Perona. Variable-length texture representation. In *Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference*, June 2007.

[2] J. Crisman and C. Thorpe. Scarf: A color vision system that tracks roads and intersections. In *IEEE Transactions: Robotics and Automation*, Feb. 1993.

[3] H. Dahlkamp, A. Kaehler, D. Stavens, and S. Thrun. Self-supervised monocular road detection in desert terrain. In *Proceedings of the 2006 Robotics Science and Systems Conference*, Sept. 2006.

[4] P. Lombardi, M. Zanin, and S. Messelodi. Switching models for vision-based on-board road detection. In *Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference*, Feb. 2005.

[5] T. Norrie and J. Seng. Team-based project design of an autonomous robot. In *Proceedings of the 2007 International Conference on Frontiers in Education: Computer Science and Computer Engineering*, June 2007.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 2002.

[7] C. Tan, T. Hong, T. Chang, and M. Shneier. Color model-based real-time learning for road following. In *Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference*, Sept. 2006.