A HYBRID HOUGH-HAUSDORFF METHOD FOR RECOGNIZING BICYCLES IN NATURAL SCENES

JOHN H. DUKESHERER[1] and CHRSTOPHER E. SMITH[2]

[1]Medtronic - Surgical Navigation Technologies, Broomfield, CO 80020
[2]Dept. Electrical and Computer Eng. University of New Mexico, Albuquerque, NM 87131

Abstract

Computer vision techniques have been applied to a variety of problems in Intelligent Transportation Systems (ITS), including vehicle following, vehicle guidance, license plate reading, pedestrian tracking, and vehicle counting. This paper presents research on a new computer vision approach to an ITS domain: bicycle counting. Traditionally, bicycles have not received the attention of transportation departments for a variety of reasons. As a result, effective and portable automatic methods for counting bicycles have not been developed. This paper presents a method for identifying bicycles in natural scenes taken from either time-lapse or full-motion video streams. The method is based upon a blend of two separate techniques, the Hough Transform and the directed Hausdorff distance. Experimental results demonstrate that neither technique individually achieves the performance of the hybrid technique.

1 Introduction

Transportation systems in the United States and worldwide are increasingly complex, require higher expenditures, and are more congested as modern society moves toward urbanization, expansion, and construction of high-technology highways and public transportation systems. Transportation agencies must address issues of cost and complexity by openly advocating the use of alternative transportation as a commuting and transportation choice. One popular choice is the increased use of bicycles. Although many communities have constructed and maintain bicycle paths and bicycle specific lanes on roadways, there are few traffic density studies for these paths and lanes.

This lack of comprehensive study is due to two basic reasons. The first is that bicycle paths are generally a lower priority item on transportation department agendas. They are relatively low in cost and provide recreational facilities, leading to greater acceptance by the public. The second reason is simply that an effective method for studying utilization and traffic flow of bicycle traffic is not available.

A fully automated and reliable bicycle counter is not currently commercially available. One possibility is the use of mechanically triggered mechanisms, much the same as those used to count automobiles (e.g. pressure hoses). Unfortunately bicycles are light weight and rider body weight can cause a high degree of variability in overall vehicle weight. Consequently, miscounting by existing, commercially available equipment is accentuated. An alternative would be to utilize manual counting by having transportation personnel stand at the side of a bicycle path and count each bicycle as it passes. This is a costly alternative and the amount of error introduced by human counters can be difficult to quantify.

What is required is an alternative to cost ineffective manual counting and to current error-prone automatic counters. We present research that is intended to aid in the production of a computer vision-based bicycle counting system. The goal is an offline technique that can process collected video and output the number of bicycles detected.

1.1 Problem definition and approaches

Automated detection and counting of bicycles using computer vision techniques generally involves imagery taken from a stationary camera with either full-motion or time-lapse video. In the case of time-lapse video, common techniques from computer vision such as frame differencing or background subtraction may be insufficient to correctly identify objects (e.g. bicycles) in motion. This is mainly due to changes in scene illumination, in shadowing, and in weather that, in time-lapse video, are far more pronounced due to the apparent motion of the sun in consecutive images of the video stream.

Computer vision techniques also have difficulties due to the variety of situations and backgrounds that this imagery might include. Bicycle paths fall into one of three categories: a) paths for the exclusive use of bicycles; b) paths for a mixed use of bicycles and pedestrians; and c) paths demarcated at the side of streets and highways. This results in a wide variety of objects in the imagery and in the background, including foliage, people, pets, automobiles, trucks, etc. To correctly identify and count bicycle traffic, distinguishing features of a bicycle and/or the bicycle's rider must be found that do not readily match features extracted from other possible objects in the imagery.

Two basic approaches appear to be the most logical for identifying bicycles and were initially investigated by Rogers and Papanikolopoulos [10]. The first approach is a geometric decomposition based upon (possibly) unique features of the bicycle such as the wheels, the frame, or the drive train. This approach was implemented using a Hough Transform in [10] but suffers from false negatives. The second approach is to use a template-matching method to match a template of a bicycle or a rider (or both) to objects in the gathered imagery. This method was minimally covered in [10] (using the Hausdorff distance), but the results were derived from only a single image and did not consider the variety of natural scenes possible along bicycle paths.

We propose a third possible approach that utilizes the Hough Transform and the Hausdorff distance in a hybrid algorithm. The Hough Transform is initially used to find possible bicycle wheels in the imagery. The locations near these potential wheels are then analyzed using the Hausdorff distance to determine if a strong match exists with the bicycle model. This method is effective since the Hough enforces a strong geometric constraint that excludes image artifacts that cause the Hausdorff to report false positives. In turn, the Hausdorff can verify the existence of a bicycle even when the Hough finds only one of the two wheels (a typical result).

2 Prior Research

Prior research in the field of computer vision for object detection has mainly centered on the decomposition of image elements into known geometric shapes that can be readily identified. The Hough Transform is a well-known algorithm for ascertaining the presence of geometric shapes in edge-detected images. Others have used template-based methods to gather information from edge-detected images regarding the presence of more complex shapes produced by objects that are an amalgam of several geometric shapes or objects that do not decompose well into sub-shapes. The Hausdorff distance is one such solution where a template is compared to an image to find matches for that template in the image.

2.1 Bicycle detection

Bicycle detection had not been investigated prior the work presented in [10]. This research looked at two competing ideas, the Hough Transform and the Hausdorff distance. They reported the problems we discussed earlier and concluded that a Hough-exclusive method was the best alternative.

The Hough Transform [2, 5, 4] is used to find edge elements that conform to known geometric shapes such as lines, circles, ellipses, and other simple curves, with detection performed through use of the parametric equations that describe the particular shape. Because bicycles contain derivations of simple geometric shapes such as somewhat predictable line configurations (the frame) and circles (i.e., the tires and wheels when viewed from the side) this approach has merit for the problem of detecting bicycles.

The principle idea is as follows: assume we are given a set of points and also assume that we know the parametric form of the curve to which these points belong. We may then determine the parameters of the curve underlying the points in two steps. First, each edgel in the edge-detected image is used to solve the parametric equation(s) of the curve of interest. Since only one point on the curve is used in the solution, the system of equations is under-constrained and many solutions will exist. For instance, if the Hough is used to detect circles, a single point could belong to an infinite number of circles in the image. Each of these possible circles will have a center and a radius derived from the parametric equations.

In the second step, for each possible solution (at a given level of granularity), a vote is registered for that particular curve interpretation. In the circle example, a three dimensional voting matrix is used to collect the votes for each possible circle in the image. The dimensions of the voting matrix are the center of the circle in X and Y and the radius of the circle. Cells of the voting matrix that receive a high number of votes relative to the radius of the potential circle supports the existence of that particular circle [3].

2.2 The Hausdorff distance

Central to the problem of pattern recognition in computer vision is the ability to determine the extent to which one shape is similar to another. The geometric comparison of shapes is a fundamental tool for model-based object recognition, where many of the methods used refer to a similarity measure between the model features and features of the image.

The Hausdorff distance, derived from set theory, measures the divergence of a set of features with respect to a reference set of features [1, 8]. Our initial investigation indicated that using the Hausdorff distance to measure similarity between data from an image and the data contained in a template would provide reasonable results for bicycle detection.

Much of the published material available outlining use of the Hausdorff distance concentrates heavily on the optimization of transformations of model sets used in object detection, focusing on instances where the models are allowed to translate and scale with respect to the images [6, 7, 9, 11, 12, 13].

Our approach differs from the presented research in two important ways. First, due to the predictable nature and size of bicycles in imagery and due to the defined placement of devices for image acquisition, we will perform only fixed point calculations using the Hausdorff distance. Second, our approach makes use of the forward directed Hausdorff distance alone. Calculation of the reverse Hausdorff distance will not provide reliable indications of matching, a topic we discuss in more detail later.

3 The Physical Properties of Bicycles

Bicycles are made up of three distinctly identifiable objects that could be used separately or in combination to identify an object in an image as a bicycle.

The rider: Much work has been done to identify humans in images. We do not use the rider since variations in the human form and position are extremely diverse, especially in cycling domains;

The frame: Because of the nearly endless number of distinct frame configurations available, we choose not to focus on the frame. Included are several photographs of mass produced, commercially available bicycles to illustrate this point (see Figure 1).

The wheels: The wheel set is composed of two rims of equal size and tires encircling the rims at the outside edge. Bicycles generally have one of two possible standardized wheel sizes. Mountain bicycles are typically built with a standard 26 inch diameter rim size and road bicycles are built with a 700c rim size (approximately 29 inches).

4 A Directed Hausdorff Distance Approach

We include a Hausdorff-exclusive approach since the results of this method influence the choice of a hybrid approach. Additionally, the work reported in [10] used both the forward and reverse distances. As discussed later in this section, our investigation demonstrates that the reverse distance is not appropriate for this application.

The Hausdorff distance is derived from set theory and is a similarity measure for two sets. It can distance can be described as the maximum distance of a set of points to the nearest point in the other set. Given two finite point sets A = $\{a_1\ldots, a_p\}$ and B = $\{b_1\ldots,b_q\}$, the Hausdorff distance is given by:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (4.1)$$

where $h(A, B) = \max_{a \in A}\min_{b \in B}\|a - b\|$ and $\|.\|$ is some underlying norm on the points of A and B (e.g., the L2 or Euclidean norm).

Figure 1 Bicycle styles

The function $h(A, B)$ is the directed Hausdorff distance from *A* to *B*. This distance identifies the point $a \in A$ that is farthest from any point of *B*, and measures the distance from a to its nearest neighbor in *B*, using the norm $\|.\|$. If $h(A, B) = d$, then each point of *A* must be within distance *d* of some point in *B*, and there exists a point of *A* that is exactly distance *d* from the nearest point in *B*.

The Hausdorff distance, $H(A, B)$, is the maximum of the two directed distances $h(A, B)$ and $h(B, A)$. It differs from other approaches to shape comparison since there is no explicit pairing (correspondance) of points of *A* with points of *B*.

For image comparison, the directed Hausdorff distance $h(B, A)$ uses the template A and only those edgels in image B that lie under the extent of the template at any given offset of the template on the image.

The Hausdorff distance is asymmetrical. The forward distance, $h(A, B)$ and the reverse distance, $h(B, A)$ are rarely equal, stemming from the asymmetry of max-min functions. This asymmetry is detrimental to our particular problem. We only use the forward Hausdorff distance, $h(A, B)$, where set A is the template and set B is the image. A bicycle in a natural scene typically has background edgels showing through the wheels and frame (see Figure 2). When operated on by the reverse Hausdorff distance, $h(B, A)$, such images can produce faulty results since dense edgel regions produce small distance measures.

4.1 Experimental results

We found that the directed Hausdorff distance provides accurate results only for identifying bicycles in "clean" imagery such as the bicycles shown in Figure 1. The method consistently returns small, predictable distances when applied to these types of images. Figure 3 shows a Hausdorff-detected bicycle in one such image.

Unfortunately, this performance does not extend to natural images. Images with dense edgel regions have a high rate of both false positives and false negatives. Two such instances are shown in Figure 4 and in Figure 5. Figure 4 shows the results from a natural scene using a template that contains partial wheels (the same template as was used for Figure 3). Figure 5 shows a single false positive on the same image when

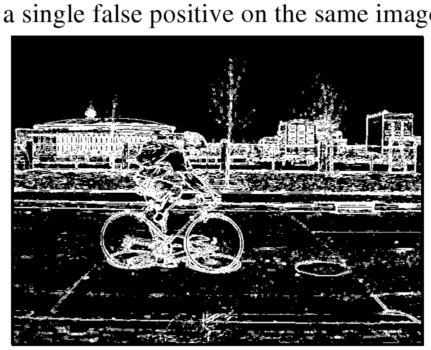

Figure 2 Edge-detected natural scene

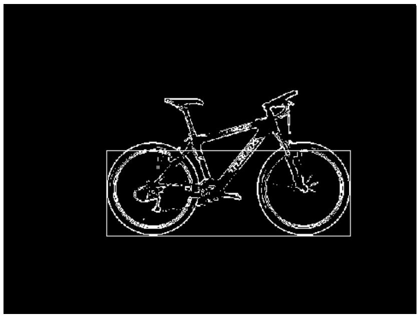

Figure 3 Clean image with Hausdorff result

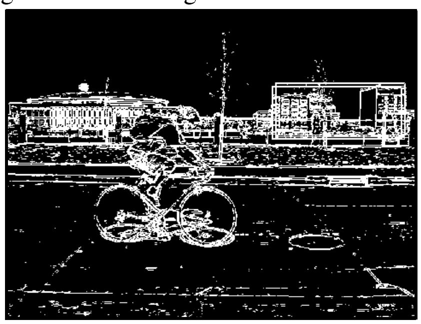

Figure 4 False positives in a natural scene

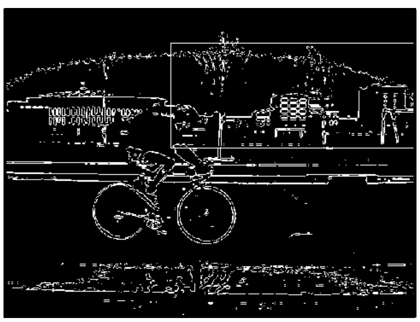

Figure 5 Result using an enhanced template

a higher edge-detection threshold was used and the template was padded with blank space around the perimeter. Both changes eliminated some of the false positives present in Figure 4; however there still remains a false positive and the method still does not identify the bicycle in the scene. Clearly an exclusive Hausdorff approach is not reliable for our problem.

A failure analysis indicates that the Hausdorff distance method gives false negatives only when the constraints in the template and image thresholds are optimized to eliminate the false positives. What is required is a method for eliminating false positives while not inducing false negatives.

We suggest that by using a circular Hough Transform, we can filter out the dense edgel areas that created false positives by identifying image regions that have geometric evidence supporting the presence of a bicycle. The Hough, unfortunately, also fails as an exclusive method since it rarely detects both wheels of a bicycle unless the vote threshold is arbitrarily low. It also detects vehicle tires, manhole covers, and other circular shapes in natural scenes, making it a good choice for determining candidate regions, but a poor choice for bicycle discrimination.

5 The Hybrid Hough-Hausdorff

The Hough Transform is based upon the parameterized equations for circles given by:

$$x_C = x - r\sin\theta \tag{5.1}$$

$$y_C = y - r\cos\theta . \tag{5.2}$$

In practice, the parameter d can be constrained by using camera placement, camera focal length, and knowledge of the diameters of bicycle wheels.

By using the Hough, we first limit the search area in an image to the probable locations of bicycle traffic. Eliminating searches in the extreme upper and lower bands of the image shortens the effective run-time. Within this reduced search area we apply the Hough transform. This has the advantage of acting as a first pass geometric filter. If no circle of the appropriate size is present, then the Hausdorff is not called.

In the event a circle of the correct proportions is detected, the directed Hausdorff distance is used to verify the existence of a possible bicycle. We use a minimal bicycle model that only incorporates wheel data since this is the most consistent feature across the variety of commercial bicycles. The Hausdorff distance calculation is performed for a small number of possible placements of the template taking into account that the circle located earlier by the Hough transform could possibly be either the front or rear wheel of a bicycle. This blended approach balances the geometric details identified by the Hough Transform with the template-related spatial relationships identified by the Hausdorff distance.

The basic algorithm of our approach is given by the following pseudo code:

1. For each $a_i \in A$ register a vote in the cell of the voting matrix for each parameter triple (x_C, y_C, r) that satisfies the parametric equation for a circle with a radius that falls between r_{min} and r_{max};
2. For each cell in the voting matrix, calculate the circle *circumference* based upon the radius from the voting matrix. If *votes_in_cell* > *z* * *circumference* then
   if $h(A, B)$ < *acceptable_distance*, tag the region as indicating a bicycle;

where a_i are the edge pixels (edgels) in A, the centers of the prospective circles are given by the ordered pairs (x_C, y_C), the radii of the prospective circles are given by r, the number of pixels expected in a circle of a specific radius is given by *circumference*, *z* is the percentage of pixels needed in *circumference* to indicate a valid circle, the range of radii for prospective circles is from r_{min} to r_{max}, the number of votes in a given cell of the voting matrix is *votes_in_cell*, and *acceptable_distance* gives the threshold for identifying a match for a section of image A to template B using the directed Hausdorff distance, $h(A, B)$.

5.1 Experimental results

Results using the hybrid algorithm demonstrate the effectiveness of this approach. The use of the Hough transform to prune the search space allows near optimal placement of the template used to calculate the forward Hausdorff distance. It also enforces geometric constraints that eliminate the false positives that the Hausdorff exhibited as a stand-alone method.

The Hough transform is used to locate possible sites for wheels in the image and each of the possible sites is recorded, illustrated by the placement of a set of cross hairs (see Figure 6). Note that the Hough has only detected the front wheel of the bicycle since the view of the rear wheel is not complete. This result is a common result for many natural scenes.

For each of the possible wheel sites, the directed Hausdorff distance is calculated for two placements of the template: one corresponding to a left wheel detection by the Hough and one corresponding to a right wheel detection. The Hausdorff template that we are currently using is left-right symmetric and does not require that we consider the mirror image of the template to detect bicycles travelling both left-to-right and right-to-left (the template is shown in Figure 7).

Using the origin of the circle detected by the Hough transform, we search from a calculated starting point, placing the template on the image in a 10×10 pixel region about the detected circle. If an acceptable (i.e. small) distance is returned by the forward Hausdorff distance, a bicycle detection is tallied at this position in the current image. Figure 8 illustrates the position of the template on the image where an acceptable forward Hausdorff distance was calculated.

We have applied the hybrid algorithm to a variety of images that include mixed traffic scenes and scenes

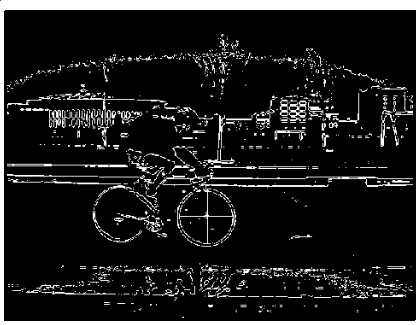

Figure 6 Hough selection of potential bicycle

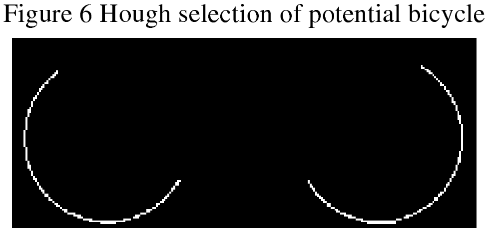

Figure 7 Hausdorff template

with no bicycles. Since the two observed problems with using either technique as an exclusive method involve a high rate of false negatives (for the Hough) and false positives (for the Hausdorff). we included imagery that might present problems for the hybrid method considering the known shortcomings of either the Hough Transform or the Hausdorff distance.

For instance, Figure 9 shows a scene from a bicycle lane that shares a surface street with normal vehicular traffic. A passenger bus presents the correct size and placement for the Hough to identify a potential bicycle, but our hybrid approach has rejected the hypothesis of a bicycle at this position. Figure 10 shows a similar example with an automobile. Again, the Hough filter has identified a region that the Hausdorff has rejected.

The only false positive that the hybrid approach has given during testing is shown in Figure 11. Here, in an image from the Criterium bicycle race, the bicycle is not orthogonal to the image plane on the camera. The result is technically a false positive, since the hybrid algorithm has counted two bicycles, where

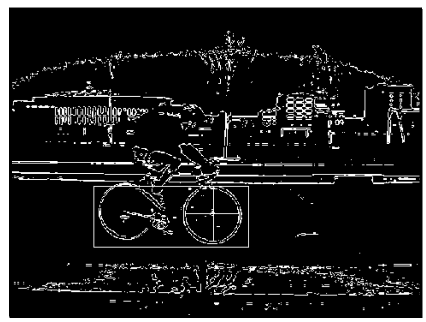

Figure 8 Hausdorff confirmation of bicycle

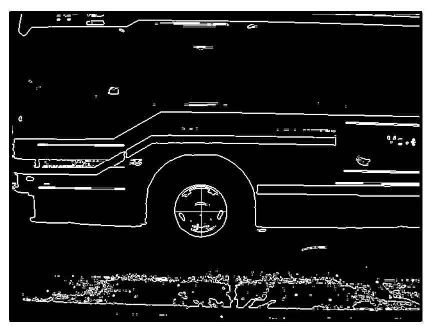

Figure 9 Rejection of a bus

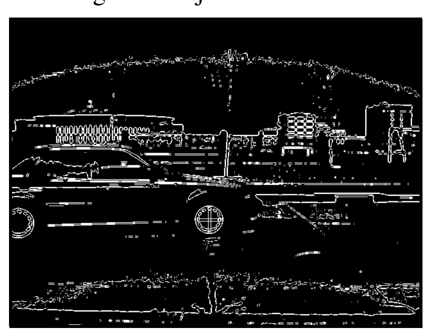

Figure 10 Rejection of an automobile

only a single bicycle exists. This could be eliminated by analyzing the amount of overlap between two identified bicycles; however, it is common for cyclists to ride side-by-side, producing scenes where a high degree of overlap would exist for two legitimate bicycles. We are currently investigating other methods for addressing the multiple counting of single bicycles.

5.2 Performance

Two performance factors are relevant in our work: efficiency and effectiveness. Efficiency of the Hausdorff distance was the major factor in the dismissal of this method by Rogers and Papanikolopoulos [10]. Their work showed the Hough Transform to be more suitable due to a lower runtime. Near-optimal placement of the template greatly reduces the runtime of the directed Hausdorff distance.

The effectiveness of the algorithm has been demonstrated during our preliminary testing. When applied to a variety of images (our preliminary data set included 25 images of bicycle paths that included vehicular traffic in both the foreground and background, we have found an accuracy over 95% in the detection of bicycles in natural scenes. For these same images, the Hough Transform rarely found both bicycle wheels, demonstrating that a Hough exclusive method would be successful in only 30% of the cases. During our final testing of the hybrid method, no false negatives were observed and only a single false positive, shown in Figure 11, was observed. The false positive is a minor error, since a bicycle exists at the location of both of the report detections; however, the method over-counted the number of actual bicycles.

6 Conclusions

In this paper, we have presented a hybrid approach to the problem of using model-based methods to locate bicycles in imagery. The method uses a Hough Transform for circles to identify potential bicycle wheel locations in images and then uses the forward Hausdorff distance to compare the image to a template for a bicycle. Our research indicates that the use of the directed Hausdorff distance in conjunction with the Hough transform performs satisfactorily for the task of bicycle detection in natural imagery. In preliminary testing with a variety of images, the hybrid method eliminates false positives and false negatives that plague methods based exclusively upon either the Hough or the Hausdorff.

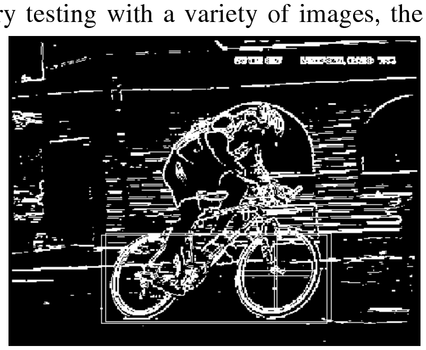

Figure 11 Non-orthogonal bicycle

Acknowledgements

This work has been supported by the Center for Transportation Studies at the University of Minnesota, the Minnesota Department of Transportation, the Department of Electrical and Computer Engineering at the University of New Mexico, and the University of Colorado at Denver.

References

[1] Atallah, "A linear time algorithm for the Hausdorff distance between convex polygons," *Information Processing Letters*, 17:207-209, 1983.

[2] Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, 13(2):111-122, 1981.

[3] Ballard, and, C. Brown, Computer Vision, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1982.

[4] Brown, C. M. "Inherent bias and noise in the Hough transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(5):493-505, 1983.

[5] R. Duda, and P. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, New York, 1973.

[6] W. Grimson and D. Huttenlocher, "Analyzing the probability of a false alarm for the Hausdorff distance under translation," *Proceedings of the ARPA Image Understanding Workshop*, II:1257-1262, 1994.

[7] D. Huttenlocher and K. Kedem, "Computing the minimum Hausdorff distance for point sets under translation," *Proceedings of the 6th annual ACM Symposium on Computational Geometry*, 340-349, 1990.

[8] D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850-863, 1993.

[9] D. Huttenlocher and W. Rucklidge, "A multi-resolution technique for comparing images using the Hausdorff distance," *Technical Report TR 92-1321,* Department of Computer Science, Cornell University, NY, 1992.

[10] Rogers, S. and Papanikolopoulos, N., "A robust video-based bicycle counting system," *Proceedings of the IEEE International Conference on ITS*, 2000.

[11] Rucklidge, W., "Lower bounds for the complexity of the Hausdorff distance," *Technical Report TR 94-1441*, Department of Computer Science, Cornell University, NY, 1995.

[12] W. Rucklidge, "Locating objects using the Hausdorff distance," *Proceedings of the 5th Annual ICCV Conference on Computer Vision*, 457-464, 1995.

[13] W. Rucklidge, "Efficiently locating objects using the Hausdorff distance," *International Journal of Computer Vision*, 24(3):251-270, 1997.