Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Speeded-up Algorithm for Human/Vehicle Classification using Hilbert Scanning Distance

Young Min Baek*. Jeong Hwan Choi*. Jin Young Choi*

** Electrical Engineering Department, Automation and System Research Institute(ASRI)*
Seoul National University, Seoul, Korea
(e-mail: ymbaek, jhchoi@neuro.snu.ac.kr, jychoi@snu.ac.kr)

Abstract: In this paper, we propose speeded-up algorithm for human/vehicle classification in real-time surveillance system. We have approached this problem using silhouette-based template matching. The silhouette of the object is extracted, and then it is compared with representative template models. Template models are previously stored in the database. Our algorithm is similar to previous pixel-based template matching scheme like Hausdorff Distance, but we use 1D image array rather than 2D regions inspired by Hilbert Curve. Transformation of images could reduce computational burden to compute similarity between the detected image and the template images. Experimental results show robustness and real-time performance in human/vehicle classification.

1. INTRODUCTION AND RELATED WORKS

Human/Vehicle Classification is one of the most interesting fields in surveillance systems, and many investigators have researched in this problem. Three methods are mainly used; 1) Silhouette-based, 2) Feature-based, and 3) Learning-based Classifier. Xiaoxu Ma et al. (2005) proposed feature-based algorithm for vehicle classification. Feature-based algorithms can classify the vehicle category in more detail. Also this is very effective to occlusion problems. However, this approach requires a lot of computation power, and some constraints are expected that many local features with respect to view points and categories should be trained in advance. Learning strategy using neural network shows real-time performance, but various patterns of target objects should be trained before classification.

Silhouette-based scheme by Yiğithan Dedeoğlu et al. (2006) is simpler and faster than other ones; in addition, training process generally is not essential. Some representative models are chosen to be compared with target objects. Various kinds of similarity measures are proposed for this purpose. Hausdorff Distance which is proposed by Gunte et al. (1991) is a popular and well known algorithm to find similarity. This can give us similarity which is very useful to classify in Hausdorff dimension. However, in order to figure out Hausdorff Distance, all of the distances between pair-wise pixels should be computed. Being caused computational load by large amount of pixel process, this algorithm is not suitable for surveillance system which is processed in real-time manner. Reducing computation time is a very important issue for real-time systems. Dimension reduction scheme can be helpful for this purpose. If the two dimensional space image can be transformed into one dimensional space, computational time could be shortened.

To solve this problem, Li TIAN et al. (2006) proposed new distance measure named as Hilbert Scanning Distance that is inspired by Hilbert Curve. Hilbert Curve is a path that contains all regions of square in high-dimension space. Following this curve, we can get one-to-one correspondence 1D space image array. Furthermore, Hilbert Curve has more possibility to preserve neighbourhood pixels than a sequential curve or a zigzag path. Using Hilbert Curve, we can just solve the 1D template matching ratter than 2D image processing. Li TIAN used Hilbert Scanning Distance only for pixel-point template matching.

We approached classification problem using silhouette-based template matching. Template matching is a method to determine similarity between two images. However, previous matching algorithms (e.g. Hausdorff distance) are generally computationally heavy, and they are not suitable for real-time surveillance. In most cases, 2D image processing needs more computation times than 1D's case. So, we use 1-dimensional similarity measurement named Hilbert Scanning Distance(HSD) by Li TIAN. To find HSD, we transform 2D image into 1D image array by Hilbert Curve. Transformation of images could reduce computational burden to compute similarity between two comparing images. Proposed algorithm includes image preprocessing and transformation using Hilbert Curve. After transformation, we find the image similarity measure more quickly than other template matching algorithms. Our algorithm shows the robustness and effectiveness in human/vehicle classification in real-time manner.

The remainder of this paper is organized as follows. In section 2, we show Hilbert Scanning Distance for template matching. In section 3, we indicate our method to classify human/vehicle and demonstrate the classification results in section 4. Then, we conclude this paper in Section 5.

2. HILBERT CURVE & SCANNING DISTANCE

A Hilbert Curve is a continuous fractal space-filling curve first described by David Hilbert. This Curve can be used for a one-to-one mapping function to transform 2D image into 1D image. Also, this curve has more possibility to preserve neighbour pixels as much as possible than any other curves. For this characteristic, this curve has been studied for many applications like image compression(Sambhunath, 2000), 3D graphics, etc.

Hilbert Curve has one parameter which determines curve size, Hilbert Order. According to order, the curve size can be varied. When Hilbert Order is k , the size of Hilbert Curve will be $2^k \times 2^k$ in 2D space. Fig.1 shows several example of Hilbert Curve in proportion as order is increased from 1 to 5.

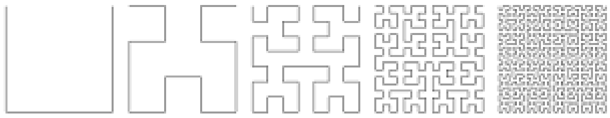


Fig. 1. Example of Hilbert Curves (Order : from 1 to 5)

2.1 Generation of Hilbert Curve

There are several implementations for generation of Hilbert Curve. In our implementation, Hilbert Curve is generated by an iterative function call. Basic Hilbert Curve begins when order is 1. The shape of curve is expanded from basic curves. An iterative rule is defined as follows:

Step.1 Initialize variables (Orientation=1, Angle=0)

Step.2 Call Curve generation function with initial angle and Hilbert Order

Step.3 In Curve generation function,

Step.3-1 Order decrease by 1, if order is less than 0, return

Step.3-2 Angle and Orientation changed sequentially (Table. 1)

Table 1. Curve Generation Change rule

	Phase 1	Phase 2	Phase 3	Phase 4
Angle	+90	-90	-90	+90
Orientation	-1	1	1	-1

Step.3-3 Call Curve generation function recursively with respect to 4 directions with new Orient and Angle value

Step.4 Curve generation end

Where Angle means heading for next step with respect to Orientation which indicates a criterion axis. A mapping rule of Hilbert Curve is generated only once, before classification is started.

2.2 Hilbert Curve Feature

As we mentioned before, Hilbert Curve has several features when we apply this Curve to template matching.

1) Hilbert Curve gives a one-to-one mapping function to transform 2D image to 1D image. Generally, computation time depends on complexity of the data. By reducing the image dimension, we can expect to speed up template matching scheme.

2) Hilbert Curve preserves coherences between pixels. We can estimate image resemblance calculating the nearest pixel distance. Hilbert Curve maintains neighbourhood pixel data well. This is the reason that Hilbert Curve is used for image compression as well.

3) Hilbert Curve should be a regular square. Actually, this property is a constraint to express 2D image. Our method is performed after object shape is detected, and then detected object is resized into square box with same aspect ratio.

2.3 Definition of Hilbert Scanning Distance

Assume that two binary 2D images are given. ($A = \{a_1,...,a_I\}$, $B = \{b_1,...,b_I\}$) These images are converted to 1D images using Hilbert Curve. ($S = \{s_1,...,s_I\}$, $T = \{t_1,...,t_I\}$) Fig.2 shows an example of transformation using Hilbert Curve.

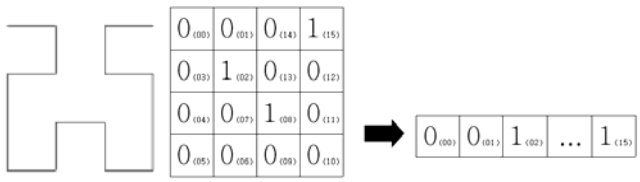


Fig. 2. Transformation of 2D image into 1D array using Hilbert Curve

Then, the directed HSD from A to B is computed by

$$h_{hsd}(A,B) = \frac{1}{N_s} \sum_{i=1}^{I} \min_j \|s_i - t_j\| \qquad (1)$$

Where ||.|| is the Euclidean norm distance in the 1D space. Parameter N_s is the number of pixels whose value is 1 in s_i . This normalization term is essential for pixel matching classification because total distance is dependent on the number of ‘1’ pixels. For example, if there is no ‘1’ pixel, the distance will be zero which is not true without N_s . In case of existing few ‘1’ pixels in the image, the distance will be very low. N_s is used to find average distance value for each pixel.

Li TIAN(2006) mentioned that threshold elimination function $\rho(x)$ will help to find a distance as well, but this function is not useful for our purpose. Being different from their application, we detect filled-in objects instead of pixels-based images. In addition, several image preprocessing are performed before calculating HSD. A flowchart of preprocessing will be discussed in more detail in section 3. Finally, HSD is defined by

$$H_{hsd}(A,B) = \max(h_{hsd}(A,B), h_{hsd}(B,A)) \quad (2)$$

This equation is similar to Hausdorff Distance. Speeded-up algorithm will be discussed in Section 3.4.

3. PROPOSED METHOD

Fig.3 shows a flowchart of finding similarity between two images. Through object detection, we can get the binary object image that is filled. Then image is resized from detected image to square size image with same aspect ratio. Because of normalization of image size, we can classify objects in a scale-invariant manner. Experimental results show that even if detected object is very small or large, classification result is almost correct. After that, the silhouette information is extracted using edge detection.

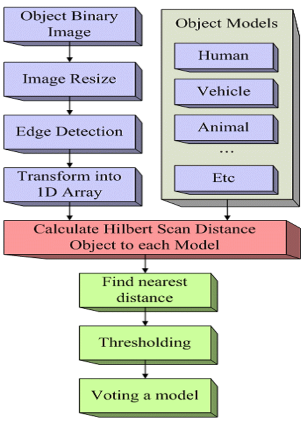


Fig. 3. Flowchart for Human/Vehicle Classification

A silhouette of 2D image is converted to 1D sequence through Hilbert Curve, and then 1D sequence is compared with representative model which is previously stored in the database. After each HSD is computed, we find smallest distance value. For eliminating objects that are not human or vehicle, selected model is confirmed using a threshold value. The threshold of distance is varied depending on application and configuration parameters.

3.1 Selection of representative models for the database

Selection of template images for database is one of the most important issues. Even if there is a same car, the silhouette of the car could be different according to points of view. This means we need several templates for one objects when some variance in the templates are allowed. Whereas humans seldom change according to view points, so we choose one representative model for humans. In our experiment, we have constructed total five models. The system has one model for human, human group and three models for vehicle. Fig.4 shows template models for each target object in our system. These models are converted into 1D array in advance in order to speed up.

(a) Human (b) Human group

(c) Vehicle Models with respect to different points of views (30, 60, 80degree rotated)

Fig. 4. Representative Template Models

Even if we cannot say that these representative models fit detected objects, a voting system to most similar models could give us a good reference for finding an adequate category. That is the reason why the number of representative models is a weighty factor. As the number of models increases, the success rate will also increase. However, computational loads also become heavier. Therefore, appropriate number of models is necessary.

3.2 Image resize of detected object

We need to transform detected images into suitable size image for finding Hilbert Scanning Distance. Because of square property of Hilbert Curve, detected image should be resized into square. However, resizing method without maintaining aspect ratio is not sufficient to our application because it could generate transformation distortion into the image. Therefore, our algorithm resizes detected images to square image with same aspect ratio of original image.

3.3 Edge Detection Algorithm

After resizing detected objects, we perform edge detection to the image. As image is consisted of binary image, any edge detection algorithms could be used. We choose the canny edge detection to find the silhouette. Here, edge points have value 1 where blank points have value 0. This value is used for next step to calculate distance. Fig.5 shows an example of total preprocessing sequence.

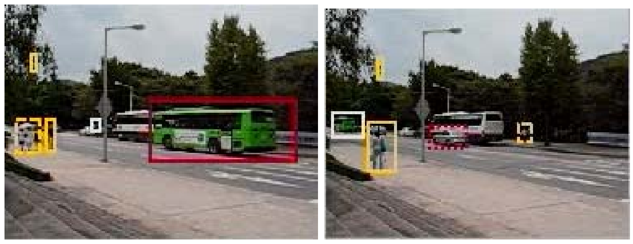
(a) Detected Object (b) Foreground Map

(c) Image resize (d) Edge detection

Fig. 5. Preprocessing for detected object

(a) PETS movie

(b) Seoul National University Outside Scene #1



(c) Seoul National University Outside Scene #2

Fig. 7. Classification results of detected images in various environments

3.4 Finding Hilbert Scanning Distance

To find Hilbert Scanning Distance, two images are transformed using Hilbert Curve. The basis sequence A has '1' value as edge line, while comparing sequence B has '2' substituted from '1'. Then two sequences are added up like Fig. 6.

0	0	0	1
0	0	0	1
1	0	0	0
1	0	0	1

Image A

1	0	0	0
1	0	0	1
1	1	0	0
0	0	0	0

Image B

1D sequence of A	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0	1
1D sequence of B	2	0	0	2	2	0	0	2	0	0	0	0	2	0	0	0
1D sequence of A+B	2	0	0	2	3	1	0	2	0	0	1	0	3	1	0	1

Fig. 6. Addition operation of two images

Now A+B has values of '1', '2' and '3'. '1' denotes only belong to A sequence, in the other hand, '2' denotes only belong to B sequence. '3' means that point belong to both A and B. For convenience, we call the point of each value as p_{num}. Now, we want to get the sum of distances from all of p_1 to nearest p_2 of it. There is no need to calculate p_3. In those points, p_1 and p_2 are overlapped, and the distance d_i should be 0.

As all points are scanned sequentially, the locations of p_1s are stored. Distance d_i of p_1 is selected comparing with forward p_2 and backward p_2. Then p_2 is met, this point is stored as a backward p_2. Then, we can calculate all distances of p_1 with that point. After that, backward p_2 become forward p_2. This process is done by only one scanning, so we can reduce a lot of computational time effectively.

4. EXPERIMENTAL RESULTS

4.1 Application for Human/Vehicle Classification

We can estimate similarity by proposed algorithms. Experiment were performed in a condition that there are only three object categories; human, human group, and vehicle. But it could be extended to other moving objects like animal, motorcycle and so forth. In the result of the experiment using three models, success ratio shows above 95% with the exception of occlusion's cases. Fig.7 shows the results of entire system which is integrated with object detection. We tested the movie files that are taken by our university and those that are from PETS to verify our algorithm. Some frames of these movies contain humans, human groups and vehicles. Yellow line indicates Human, white line indicates human group, and red line is for vehicles. These results show correctness for classification. Table.2 represents several Hilbert Scanning Distance of detected object when Hilbert Order is 6.

(a) Human (b) Human Group (c) Vehicle

Fig. 8. Voting results of detected objects in various environments

Table 2. Examples of Hilbert Scanning Distance

Model / Object					
	8	44	112	111	81
	19	35	102	98	96
	8	31	115	117	86
	9	36	108	105	86
	16	31	129	127	82
	35	27	113	105	56
	78	33	94	89	51
	33	22	128	120	58
	93	71	92	84	48
	108	65	117	106	49
	108	109	10	7	45
	109	123	11	11	47

As we can see in Table.2, Hilbert Scanning Distance would be less when detected object is similar to representative model. The size of objects for experiment has large variations; smallest size is roughly 15*15, and largest one is approximately 150*150, scale-invariant classification is achieved because image size is normalized to Hilbert Order. Fig.8 represents classification results of each detected objects in those movies.

4.2 Computational time and Success Rate according to Hilbert Order

The system specification for experiments was Pentium IV 2GHz with 1G main memory. The resolution of movie file is 320*240. The computation time for one object is about 20msec on average when Hilbert Order is 6. Frame rate is above 15frame/sec in this system, and the result is verified that proposed methods can be applied to real-time detection and classification.

The computation time and success rate of Hilbert Scanning Distance depends on one parameter; Hilbert Order. As order increases by 1, image size will be enlarged by 4. This causes image complexity. The computational complexity is represented by $O(n^2)$. In addition, the number of edge pixels slightly affects the complexity.

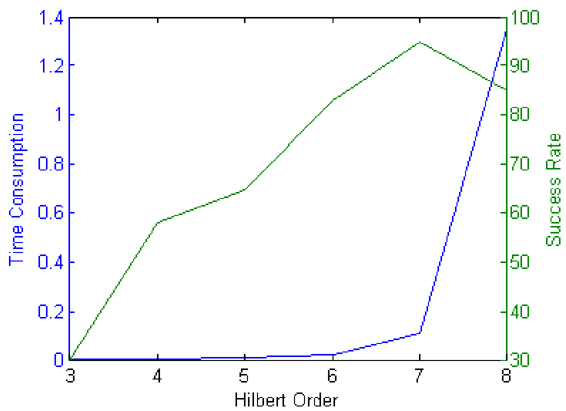


Fig. 9. Time Consumption and Success Rate for each Hilbert Order

As order increases, success rate increases as well. However, if order is higher than a fixed value, success ratio could be decreased. As image dimension become 1D space, some neighbourhood pixels are missing. Due to the higher order, distortion noises affect more easily to image. Fig. 9 demonstrates the computational complexity and success rate according to Hilbert Order. As the result of this graph shows, the result is most appropriate when Hilbert Order is 6 or 7.

5. CONCLUSIONS

We proposed a novel human/vehicle classification algorithm using Hilbert Scanning Distance for real-time surveillance system. We have also had experiments in various environments and view points. To overcome the time consumption of previous template matching algorithms in 2D image space, we transform the 2D image into 1D array, and find similarity measure quickly. This algorithm has benefits of follows;

1) Reducing data complexity

2) No need to learn the models

3) Robust to scale variation

The entire system that is integrated with moving object detection algorithm shows a distinguished result for our purpose in real-time manner. However, some false alarms were detected when occlusion happens. Future work is when occlusion happens, other measure will be added, something like trajectory or local features. Representative models also should be adapted to view point of the camera as we cannot define exact template models in every environment. Learning template models will be helpful to improve the performance.

ACKNOWLEDGEMENT

This research is supported by the Ministry of Commerce, Industry and Energy of Korea and Samsung Techwin Corp.

REFERENCES

Gunte Rote (1991). Computing The Minimum Hausdorff Distance Between Two Points Sets On a Line Under Translation. *Information Processing Letter,* page 123-127.

Haritaoglu. I., Harwood. D., Davis. L.S., (2000). W^4: Real-Time Surveillance of People and Their Activities. *PAMI,* **Volume 22**, page 809-830.

Lisa M.Brown (2004). View Independent Vehicle/Person Classification. *Workshop on Video Surveillance & Sensor Networks,* page 114-123.

Li TIAN, Sei-ichiro KAMATA, Kazuyuki TSUNEYOSHI, Haijiang TANG (2006). A Fast and Accurate Algorithm for Matching Images Using Hilbert Scanning Distance with Threshold Elimination Function. *IEICE TRANS. INF. & SYST. ,* **Volume 1**, page 290-297.

Yiğithan Dedeoğlu, B. Uğur Töreyin, Uğur Güdükbay, and A. Enis Çetin (2006). Silhouette-Based Method for Object Classification and Human Action Recognition in Video. *ECCV,* page 64-77.

R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, and O. Hasegawa (2000). A system for video surveillance and monitoring: VSAM final report. **Technical CMU-RI-TR-00-12**

Sambhunath Biswas (2000). Hilbert Scan and Image Compression. ICPR, **Volume 3**, page 3211.

Tatsuya YOSHIDA, Shirmila MOHOTTALA, Masataka KAGESAWA, Katsushi IKEUCHI (2002). Vehicle Classification System with Local-Feature Based Algorithm Using CG Model Images. *IEICE TRANS.* **Volume E85-D, No11**, page 1745-1752.

Xiaoxu Ma, Grimson. W.E.L. (2005). Edge-based rich representation for vehicle classification. *ICCV,* **Volume 2**, page 1185-1192.

Yang Ran Weiss, I. Qinfen Zheng Davis, L.S. (2004). An Efficient and Robost Human Classification Algorithm Using Finite Frequencies Probing. *CVPR Workshop,* page 132.