MULTIPLE VEHICLE TRACKING USING GABOR FILTER BANK PREDICTOR

James Graham[1], Mehmet Celenk[1], John Willis[1], Tom Conley[1] and Haluk Eren[2•]
[1]School of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701, U.S.A.
[2]TBMYO, Computer Technology Department, Firat University, Elazig, Turkey

Keywords: Multiple vehicle tracking, Change detection, Gabor filter bank predictor, Motion estimation.

Abstract: This paper presents a time-varying Gabor filter bank predictor for use with vehicle tracking via surveillance video. A frame-based 2D Gabor-filter bank is selected as a primary detector for any changes in a given video frame sequence. Detected changes are localized in each frame by fitting a bounding box on the silhouette of the vehicle in the region of interest (ROI). Arbitrary motion of each vehicle is fed to a non-linear directional predictor in the time axis for estimating the location of the tracked vehicle in the next frame of the video sequence. Real-time traffic-video experimentation dictates that the cone Gabor filter structure is able to tune itself into a selected target and trace it accordingly. This property is highly desirable in the fast and accurate moving vehicle or target tracking purposes in range and intensity driven sensing.

1 INTRODUCTION

This work focuses on the tracking of multiple vehicles by fitting them with image oriented bounding boxes as the ROIs and predicting their positions in a given traffic image sequence using a Gabor filter bank. Detecting the ROIs is carried out by using a bank of Gabor filters to determine the relevant ROI, paired with change detection and analysis to track objects from frame to frame. Many other traffic surveillance and prediction methods have been proposed and implemented for the same purposes. Such schemes include the studies Wang, et. al, 2006, Maire and Kamath, 2005, Celenk and Graham, 2008, Zhi-fang and Zhisheng, 2007, Wang and Lien, 2008, Atev and Papanikolopoulos, 2008, Taj, Maggio, and Cavallaro, 2008, Qian Yu, and G´erard Medioni, 2008. Our approach uses the Gabor filtering as an estimator for robust ROI detection.

Gabor filters are able to discriminate the minimum bound via concurrent utilization of the spatial and the spectral domain information. Moreover, an extension application of the Gabor filter is the Gabor filter bank, which is comprised of a set of Gabor filters of different frequencies and orientations that provide a complete cover of the spatial frequency domain. An example is Macenko et al.'s (2007) work, which provides both a good explanation of the approach to using Gabor filtering and a highly relevant practical application in lesion detection within the brain. The following sections describe the overall approach and experimental results. Conclusions are given at the end.

2 TRACKING BY GABOR FILTER BANK

In this paper, the vehicle tracking problem is tackled by selecting the Gabor filter bank responses as the error invalidation criterion for estimation based tracking and surveillance. The basic premise behind the underlying objective is that prediction of an entire image is not necessarily useful, desired, or even practical. In lieu of this reasoning, the Gabor filter is chosen to help determine an ROI in the next frame as the most likely location of the moving target. More specifically, the current frame's results are projected "forward" for processing in the next frame as illustrated in Figure 1. The approach has a flaw in its operation due to the fact that only a portion of the each frame is processed. This is alleviated by interspacing a full frame analysis every few frames. Assuming piece-wise linearity, prediction of a vehicle V in the next frame at time

Graham J., Celenk M., Willis J., Conley T. and Eren H. (2009).
MULTIPLE VEHICLE TRACKING USING GABOR FILTER BANK PREDICTOR.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 632-635
DOI: 10.5220/0001806006320635

$t+1$ from its identified location in the present frame t is defined by

$$V\left(x,y;t+1\right)=V\left(x-d_x,y-d_y;t\right) \quad (1)$$

where d_x and d_y are the horizontal and vertical displacements of the vehicle in the next frame, respectively. Notice that these two displacements are determined by the speed of the motion of a vehicle. Piece-wise linearity assumption will not be a limiting factor for the generality of the method being presented here since any motion can be represented as a superposition of piece-wise linear motion. Keeping this in mind, frame rate needs to be sufficiently high to maintain a steady sample flow of moving target or vehicle.

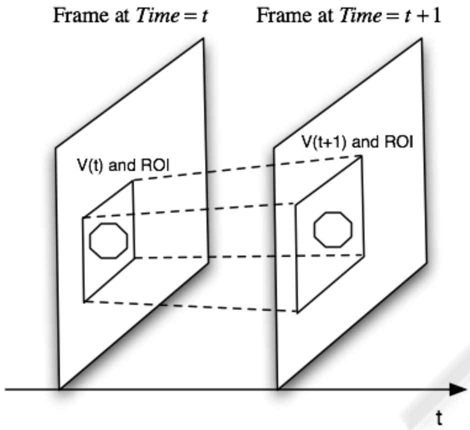

Figure 1: Gabor filter projective estimation in time.

Initially, the first two frames are used to detect any changes due to the vehicle motion. Based on the bounding box ROI results of this initialization step, a Gabor filter bank is generated as one per ROI using

$$h\left(x,y\right)=g\left(x,y\right)\cdot e^{j\left(\omega_x x+\omega_y y\right)} \quad (2)$$

$$g(x,y)=\frac{1}{2\pi S_x S_y}e^{-\frac{\left(\frac{x\cdot\cos\theta+y\cdot\sin\theta}{S_x}\right)^2+\left(\frac{-x\cdot\sin\theta+y\cdot\cos\theta}{S_y}\right)^2}{2}} \quad (3)$$

That is, $g\left(x,y\right)$ is a Gaussian that is spatially scaled and rotated by θ. S_x and S_y are the variances along the x and y axes and are equal to $\lambda\sigma$ and σ, respectively. The parameter σ is the spatial scaling, which controls the width of the filter impulse response. The value λ defines the aspect ratio of the filter, which determines the directionality of the filter. The orientation angle θ is usually chosen to be in the direction of the filter's center circular frequency ($\theta=\tan^{-1}\omega_x/\omega_y$). In the proposed system, the current frame is fed to a Gabor filter bank which calculates the output images for a series of Gabor bases with varying orientations. Here, the filter bank, which has a set of frequencies and orientations, will cover the spatial-frequency space and capture as much shape information as possible. A combined Gabor saliency map is produced from these resultant images. A set of bounding boxes is created with the saliency image and previous error results from the two corresponding Gabor filter. The prediction error (E) is calculated as

$$E=\iint_R\left[c\left(x,y;t+1\right)-c\left(x,y;t\right)\right]^2 dxdy \quad (4)$$

$$c\left(x,y;t\right)=V\left(x,y;t\right)*h\left(x,y;t\right) \quad (5)$$

$$c\left(x,y;t+1\right)=V\left(x,y;t+1\right)*h\left(x,y;t+1\right) \quad (6)$$

Here, $c\left(x,y;t\right)$ represents the convolution of the vehicle image value $V\left(x,y;t\right)$ with the Gabor filter impulse response $h\left(x,y;t\right)$ in the video frame at time t, $c\left(x,y;t+1\right)$ denotes the convolution of the tracked vehicle image value $V\left(x,y;t+1\right)$ with the Gabor filter's impulse response function $h\left(x,y;t+1\right)$ in the image taken at time t+1, and R is the region of support, or local spatial region used to estimate d_x and d_y, respectively. The solution to equation (4) requires the solution of

$$\frac{\partial E}{\partial R}=0 \quad (7)$$

Since the analytical solution of equation (7) is a nonlinear problem due to the randomly varying nature of *R*, the solution space is searched iteratively using the initial value and an incremental adjustment term. For computational efficiency, the estimated ROIs are searched in nine different positions for vehicles' images which match to those found in previous frames. The search strategy selected in this work is similar to the Matching Pursuit transform and the corresponding Gabor wavelets (Servais, 2006). In our case, however, the Gabor wavelet basis functions are adaptively formed in accordance with the shape and size of each vehicle being tracked.

3 EXPERIMENTAL RESULTS

A pair of data sets is used in experimentation from the Institut für Algorithmen und Kognitive Systeme of Karlsruhe University's traffic image sequence database; e.g., the Taxi and Rheinhafen sequences.

Images provided in the databases are in 2-D intensity format. The 2-D scene data used for this experiment is from a static surveillance camera. In the collected images, only the scene contents move while the camera remains stationary. The Taxi and Rheinhafen video frames have been converted into JPEG images for the sake of commonality with resolutions of 256x191 and 688x565, respectively. Figure 2 shows two example images from the selected databases depicting the scenes from which they were acquired.

Figure 2: Samples from Taxi and Rheinhafen databases.

In the implementation, we follow the same discrete formulation of the Gabor filter as in Macenko et. al (2007), which specifies the Gabor filter variables to be $S_x = 1$, $S_y = 1$, and $\theta = \{0, \pi/4, ..., \pi, ..., 7\pi/4\}$. Eight different orientations for the Gabor filter bank are adapted since more would not provide any significant improvement and fewer would begin to lose significant information about the image content. The aforementioned eight orientations have been realized via the nine predicate templates of size given by the bounding boxes obtained at time *t*. Below, the nine templates are shown in the order in which the spatial coordinates x and y are offset from the predicted location at position zero.

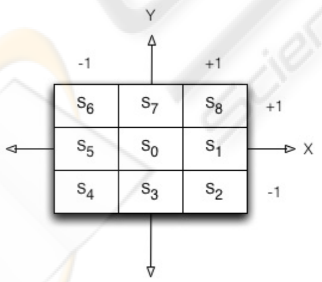

Figure 3: Spatial offset templates.

Upon passing an image through the filter bank consisting of nine Gabor wavelets arranged in the above shown structure, we have created a combined saliency image as described in (Celenk, Zhou, Vetnes, and Godavari, 2005). The saliency image has the background saliency map subtracted to leave only the correct region of interest (ROI). The resulting ROI is then passed through a noise reduction and blocking filter to remove specks which result from small background changes and to square the edges of the ROI to improve reliability.

Figure 4 shows the results of the algorithm's tracking features. Over the course of the image progression, the algorithm tracks the movement and contents of the ROI boxes and attempts to isolate and track relevant objects. The red lines in the figure represent the tracked path of the objects, while the numbers act as bounding box labels. Close observation of the images shows that there are 2 kinds of tracked bounding boxes. One type is the single blue box as seen with regards to boxes 2 and 3, and the double lines green and blue boxes. Regions such as those of 2 and 3 exist because their associated feature is no longer being tracked, at least in the current frame. This can happen for one of three reasons; the object has moved off the screen, tracking has been temporarily lost on the object, or the object is an erroneously tracked region as in the case of region 3. Region 3 was an artifact associated with region 1 that was tracked across several frames.

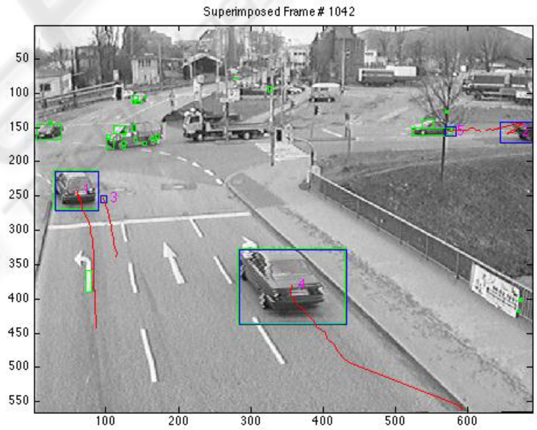

Figure 4: Tracking of various moving vehicles.

Quantitative measurements of tracking error have been made in the least mean square sense (LMSE) yielding an acceptably low tracking error as in (Celenk and Graham, 2008). Here, the main goal is to capture and track the moving objects in various road conditions. Figure 5 illustrates the progression of tracking over time, while Figure 6 simply shows an overlay of both Figures 4 and 5 at a slightly later time. Comparing the images of Figures 4, 5, and 6 with the Rheinhafen data as seen from Figure 2 gives a better idea of exactly how the objects have progressed.

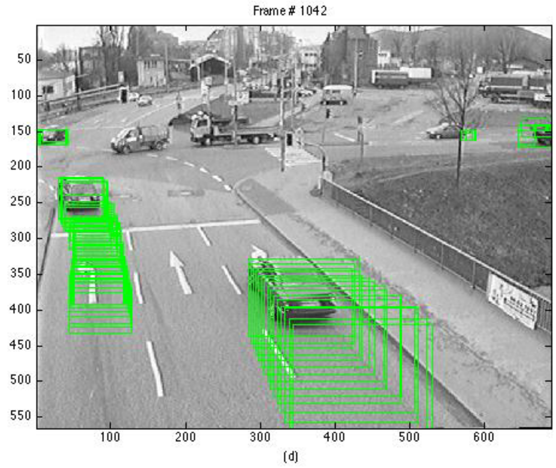

Figure 5: Progression of tracked boxes.

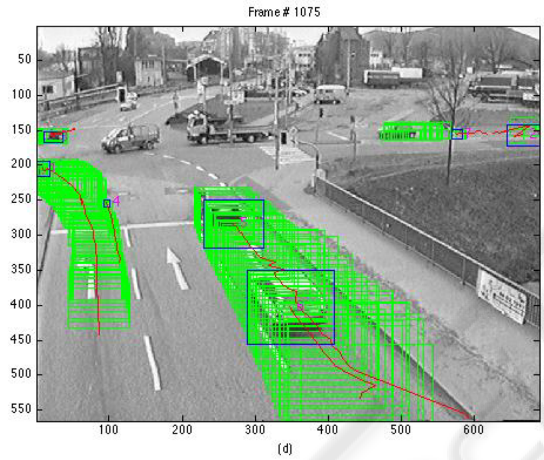

Figure 6: Overlap of tracking and progression data.

4 CONCLUSIONS

This paper presents a time-varying Gabor filter bank predictor for use with vehicle tracking via surveillance video. Detected changes are localized and motion is fed to a non-linear directional predictor in the time axis for estimating the location of the tracked vehicle in the next frame of the video sequence. Real-time experimentation has shown that the cone Gabor filter structure can adjust itself into a selected target and track its motion. This property is highly desirable for processing a fast moving vehicle or target tracking purposes. Future work involves extending the plane structured Gabor filter bank to a 3D spatio-temporal arrangement with feature selectivity. For high performance and/or real time implementation the Gabor filter bank lends itself to parallel (e.g., GPU or FPGA) implementation.

REFERENCES

Atev, S. and Papanikolopoulos, N., 2008, “Multi-view 3d vehicle tracking with a constrained filter,” 2008 IEEE ICRA, May 19-23, pp. 2277 - 2282.

Celenk, M. and Graham, J., 2008, “Traffic Surveillance Using a Gabor Filter Bank Based Kalman Motion Predictor,” Proc. VISAPP ‘08, Funchal, Maderia – Portugal, Jan 22-25.

Celenk, M., Zhou, Q., Vetnes, V., and Godavari, R., 2005, “Saliency field map construction for ROI-based color image querying,” Journal of Electronic Imaging, Vol. 4, No. 3, pp 033012-1 ~ 033012-9, July –September

Institut für Algorithmen und Kognitive Systeme, Image Sequence Server, Karlsruhe University. http://i21www.ira.uka.de/image_sequences/

Macenko, M., Luo, R., Celenk, M., Ma, L., and Zhou, Q., 2007, “Lesion detection using Gabor-based saliency field mapping,” Medical Imaging 2007, Proc. SPIE Vol. 6512, Feb. 17-24, San Diego, CA.

Maire, M. and Kamath, C., 2005, “Tracking Vehicles in Traffic Surveillance Video,” Lawrence Livermore National Lab., Technical Report UCRL-TR-214595.

Servais, M., 2006, “Content-based motion compensation and its application to video compression,” Ph.D. Thesis, Univ. of Surrey.

Taj, M., et. al, “Objective evaluation of pedestrian and vehicle tracking on the CLEAR surveillance dataset,” Multimodal Technologies for Perception of Humans, Vol. 4625/2008, June 27, 2008, pp. 160-173.

Wang, C.C.R. and Lien, J.J.J., 2008, “Automatic vehicle detection using local features—a statistical approach,” IEEET-ITS, Vol. 9, No. 1, March, pp. 83-96.

Wang, Y., et. al, 2006, “Renaissance: A real-time freeway network traffic surveillance tool,” Proc. IEEE ITSC, pp. 839-844.

Yu, Q. and Medioni, G., 2008, “Integrated detection and tracking for multiple moving objects using data-driven mcmc data association,” IEEE WMVC, pp. 1-8.

Zhi-fang, L. and Zhisheng, Y., 2007, “A Real-time Vision-based Vehicle Tracking and Traffic Surveillance,” 8th ACIS Int. Conf. Software Eng., AI, Networking, and Parallel/Distributed Computing, Vol. 1, pp. 174-179.

• Mr Haluk Eren’s contribution is supported by TUBITAK