MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Digital Lane Marking for Geo-Information Dissemination

Raymond Yim, Masashi Saito, Jinyun Zhang

TR2010-113 December 2010

Abstract

Roadway lane marking detection is essential for lane departure warning and enabling autonomous vehicles. Instead of using lane markings only as visual aids to indicate the boundary of two adjacent lanes, it is possible to also include digital information in the lane to provide geographically specific information. This enables in-tunnel navigation data, location-specific service advertisement, construction zone and lane merger warning. Technologically, such digital information should be decodeable by inexpensive equipment, even while vehicle is traveling at highway speed. We propose a Differential Orthogonal Encoding scheme that allows information to be decoded even when the captured image of lane marking suffers from a high level of motion blur, and we show the superiority of the encoding scheme through physical experiment and theoretical analysis. A vehicle equipped with IMP/cm**2 camera traveling at 30m/s (108km/h) can decode lane marking with data density of 77kb/m**2.

ITS World Congress 2010

Digital Lane Marking for Geo-Information Dissemination

Raymond Yim, Masashi Saito, Jinyun Zhang
Mitsubishi Electric Research Labs
201 Broadway, Cambridge MA 02139, USA
+1-617-621-7597, +1-617-621-7543, +1-617-621-7595
{yim,saito,jzhang}@merl.com

Naoki Shimizu
Information Technology R&D Center, Mitsubishi Electric Corporation
5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan
+81-(0267)41-2115
Shimizu.Naoki@ab.MitsubishiElectric.co.jp

ABSTRACT

Roadway lane marking detection is essential for lane departure warning and enabling autonomous vehicles. Instead of using lane markings only as visual aids to indicate the boundary of two adjacent lanes, it is possible to also include digital information in the lane to provide geographically specific information. This enables in-tunnel navigation data, location-specific service advertisement, construction zone and lane merger warning. Technologically, such digital information should be decodable by inexpensive equipment, even while vehicle is travelling at highway speed. We propose a Differential Orthogonal Encoding scheme that allows information to be decoded even when the captured image of lane marking suffers from a high level of motion blur, and we show the superiority of the encoding scheme through physical experiment and theoretical analysis. A vehicle equipped with $1MP/cm^2$ camera traveling at 30m/s (108km/h) can decode lane marking with data density of $77kb/m^2$.

INTRODUCTION

Lane Departure Warning (LDW) systems were first introduced to commercial trucks at the turn of the millennium, and it is currently widely available in most trucks sold in North America and Europe. The technology has subsequently been introduced to passenger vehicles by many makers including Nissan, Toyolta, Honda, Citroëon, General Motors, Mercedes-Benz and Fiat. It has been reported that single vehicle roadway departure crashes account for 1.2 million of the 6.3 million crashes annually in the United States [1], and more importantly it accounts for 43 percent of fatal crashes. It has also been estimated that approximately 10 percent of such crashes can be prevented with LDW [2]. Multiple government agencies, such as National Highway Traffic Safety Administration (NHTSA), are evaluating its cost-benefit to decide whether to mandate LDW to new vehicles.

Parallel to the development of LDW systems, Automated Highway Systems (AHS) has recently attracted much attention from academia and automotive industry since the demonstrations by the California Partners for Advanced Transit and Highways (PATH) program [3], and more recently the 40-million Euros cityMobil project since 2006 [4] and the 28-million Euros Highly Automated Vehicles for Intelligent Transport (HAVEit) project since 2008 [5]. These programs envision AHS to be initially deployed in over 19 European cities including London, Rome and Vienna. Ambient intelligence from road lane detection sensor and cooperative intelligence from C2X communications are some of the key enabling technologies for AHS.

Most lane detection systems for LDW and AHS use video capturing cameras to detect the presence of the lane markings. Global Position System (GPS) and highly detailed map may be used, although its accuracy is not sufficient when vehicle is in urban canyon settings. Other technology choice includes infrared sensor and Light Detection and Ranging (LIDAR), even though these systems are more costly in general. As a matter of fact, to keep the production cost low, most implementation of LDW systems obtains sensory input from cameras equipped with VGA-format CMOS image sensors with logarithmic sensing response [6]. While CMOS sensors are generally cheaper and consume less power, they also typically have less light sensitivity compared to CCD-based sensors. In other words, CMOS sensors are more susceptible to noise. To improve picture quality, it is possible to increase the exposure time of each frame so that noise would be averaged out; nonetheless, for fast moving vehicles, the resulting frame would suffer from motion blur.

The digital lane marking technology that is proposed by this paper is founded on the assumption that as LDW-enabled and autonomous vehicles become more widely available, most future vehicles will readily have the functionality to track lane markings on road surfaces. Hence, a simple software upgrade would allow these vehicles to also decode digital information that may be embedded in these lane markings. Specifically, such digital information can be stored on the lane markings by introducing small variations that can be decoded from the captured image frame. Simplistically, one may think of these markings as bar codes that store digital information; nonetheless, existing bar codes are not designed for system undergoing high mobility, and bar code readers tend to operate in rather controlled environments with consistent ambient lighting. The challenge in such digital lane marking technology lies in the development of a coding scheme that would maximize the amount of information that can be stored in a given lane marking, and allow approaching vehicles to decode such information even when the captured frame suffers from a large amount of motion blur, while doing so in a range of ambient lighting environments.

Embedding digital information into lane marking enables a wide range of applications. GPS has successfully enabled positioning technology for vehicle navigation systems. However, due to clock errors, ionospheric effects and multipath distortion, the position accuracy of civilian GPS under clear sky is still about 5 meters [7]. The accuracy drops further when the line of sight component of the arriving signal is blocked when vehicles pass through buildings, bridges and forest, and positioning is impossible when vehicles go inside tunnel. Information embedded in lane marking can include supplemental position information that augments existing navigation systems, leading to faster satellite fix time and improving the accuracy of the vehicle position awareness. Such positioning accuracy will become more important as driving features become increasingly automated in the future. Besides position information, vehicles can also download local maps on the fly without wasting radio resources. This is particularly useful for roads under constructions. Local authorities need not go through the trouble of updating a centralized database in order to communicate with vehicles in the sector. Furthermore, information on lane markings provides local merchants a mean to advertise their services to approaching vehicles. On-board navigation unit can be configured to display such service announcements to the drivers if preferred, even when drivers are not subscribed to any data communication services. Finally, security codes that may be encoded in lane markings can provide a means to provide location authentication for purposes such as accident scenario reconstruction.

In this paper, we compare the performance of four encoding schemes: Direct Encoding, Longitudinal Bipolar Encoding, Latitudinal Bipolar Encoding, and Differential Orthogonal Encoding. Through theoretical and experiment studies, we show that, even when the decoder software is equipped with a rather complex maximum likelihood decoding algorithm, Direct Encoding and Longitudinal Bipolar Encoding degrades rapidly as the vehicle speed, and thus the resulting level of motion blur, increases. Latitudinal Bipolar Encoding has better performance compared to Direct Encoding and Longitudinal Bipolar Encoding; however, the resulting complexity remains high. Finally, we introduce Differential Orthogonal Encoding, which is derived from Hadamard codes, and its respective differential receiver, and we show that the code outperforms all other codes in comparison. We also implemented a test platform with simpler decoding algorithms to show real-time decoding capabilities for the various codes.

ENCODING SCHEMES

Figure 1 depicts some basic definitions that will be used throughout this paper, and illustrates the Direct Encoding scheme. The longitudinal direction aligns with the direction to which the vehicle travels on the road, and the latitudinal direction is perpendicular to the direction of travel. Each bit of information is represented by a module in the lane marking. Each module has the same size, L meters in the longitudinal direction, and H meters in the latitudinal direction. Each different encoding scheme represents a bit in a different manner. Each lane marking is composed of multiple modules. In the figure, two lane markings, each with 6 modules, are represented using Direct Encoding.

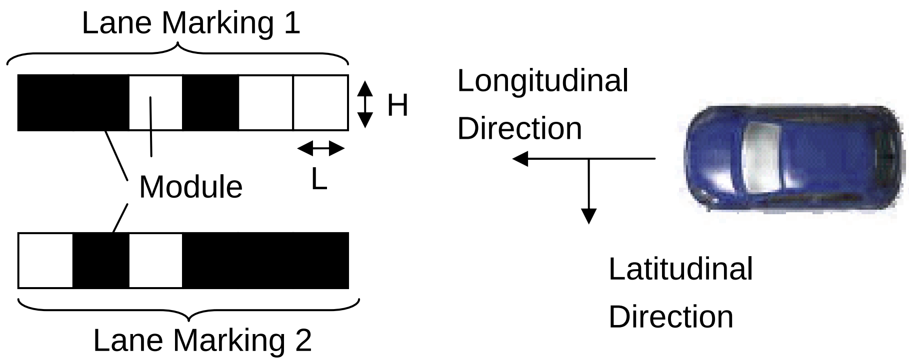

Figure 1 Direct Encoding and Definitions

Direct Encoding: a white colored module represents bit "1" and a black colored module represents bit "0". While such encoding scheme is straight forward, we immediate see a drawback of this scheme. Specifically, in Figure 1, Lane Marking 2 contains more black modules than white. This affects lane tracking capability for LDW systems. We will also see that the performance of Direction Encoding is not good.

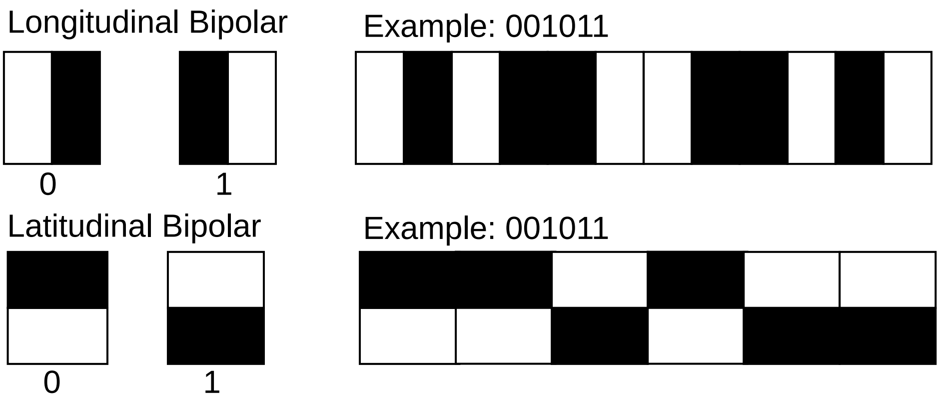

Figure 2 Longitudinal and Latitudinal Bipolar Encodings

Longitudinal Bipolar Encoding: As depicted in Figure 2, each module in Longitudinal Bipolar Encoding is divided into two sub-modules. A bit "0" is represented by white-black sub-modules in the longitudinal direction in the example, and a bit "1" is represented by black-white sub-modules. We can see a representation of a lane marking conveying bit sequence "001011" in the upper right of Figure 2.

Latitudinal Bipolar Encoding: Similar to the longitudinal counterpart, each module in Latitudinal Bipolar Encoding is divided into two sub-modules. A bit "0" is represented by black-white sub-modules in the latitudinal direction in the example, and a bit "1" is represented by white-black sub-modules. We can see a representation of a lane marking conveying bit sequence "001011" in the lower right of Figure 2.

An immediate advantage of bipolar encoding is that the resulting lane marking has a one-to-one ratio of black and white colors, since each module is consists of equal amount of black and color sub-modules. Due to motion blur, we will see that the performance of Latitudinal Bipolar Encoding is superior to that of Encoding in the Longitudinal Direction.

DIFFERENTIAL ORTHOGONAL ENCODING

Under motion blur, adjacent modules in the longitudinal direction will interfere with one another. Hence, it is desirable to encode adjacent modules such that they are orthogonal to each other. Consider the Hadamard matrix [8]; it is well known that each row of the matrix is orthogonal to any other row under an inner product operation. For example, using 4x4 Hadamard matrix, the following four orthogonal codes results: c_0 = [1 1 1 1], c_1 = [1 -1 1 -1], c_2 = [1 1 -1 -1] and c_3 = [1 -1 -1 1]. In systems such as Code Division Multiple Access (CDMA), such code can be directly used to create orthogonal channels in wireless communication systems. However, both CCD and CMOS sensors detect light intensity only. In other words, it is impossible to detect negative light intensity which results in a negative number at the receiver. Modification is needed to use such orthogonal codes for lane markings.

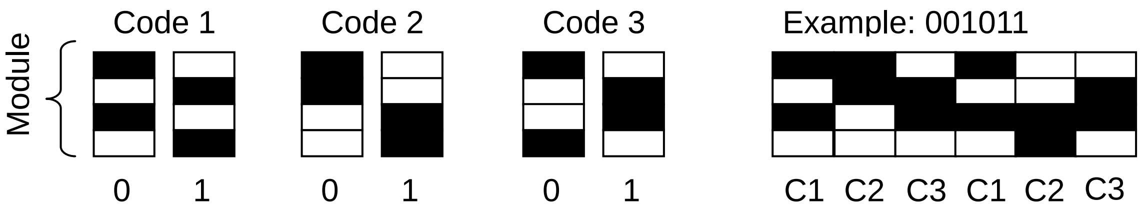

Figure 3 Differential Orthogonal Encoding

Figure 3 illustrates a Differential Orthogonal Encoding scheme when a module is divided into four sub-modules. Out of 4 orthogonal codes from Hadamard matrix, we eliminate the all-one code c_0 = [1 1 1 1], and change the remaining three codes to have positive values only. Specifically, we use c_1 = [1 -1 1 -1] as the basis for Code 1, and we use [1 0 1 0], or white-black-white-black, to represent bit "1" of Code 1, as shown in Figure 3. For the bit "0" of Code 1, we represent it using [0 1 0 1], or black-white-black-white. Similarly, with c_2 = [1 1 -1 -1], bit "1" of Code 2 is represented by [1 1 0 0], white-white-black-black, and bit "0" is represented by [0 0 1 1], black-black-white-white. Also, with c_3 = [1 -1 -1 1], bit "1" of Code 3 is represented by [1 0 0 1], white-black-black-white, and bit "0" is represented by [0 1 1 0], black-white-white-black.

On the lane marking, the three codes are used in a cyclical manner. On the right side of

Figure 3, we see an example of encoding the bit sequence "001011" using Differential Orthogonal Encoding. We see that the first two "0" bits are encoded differently since the first bit "0" is represented using Code 1, and the second bit "0" is represented using Code 2.

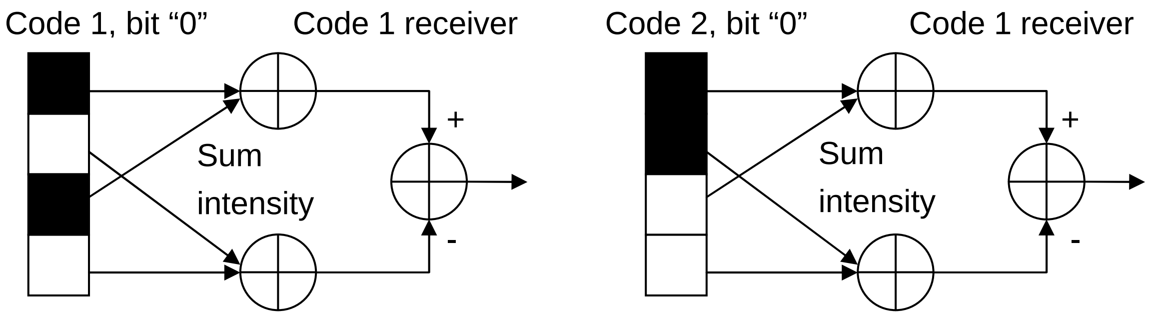

Figure 4 Differential Decoder

Figure 4 shows the receiver for Code 1 of the specific Differential Orthogonal Encoding scheme. Code 1 is generated based on c_1 = [1 -1 1 -1]. Hence, the receiver sums up the first and third sub-modules, and then subtracts the sum of the second and forth sum-modules. Assuming that white corresponds to value 1 and black corresponds to value 0, then the receiver has an output equal to -2. A negative number is interpreted as bit "0". On the left hand side, when bit "0"of Code 2 is decoded by the receiver for Code 1, the output value is 0. In other words, information encoded using Code 2 has no interference to receiver for Code 1.

ANALYSES AND EXPERIMENTS

Figure 5 shows the model of the image receiver that we will use to analyze the various encoding schemes. We assume that each frame is captured with an exposure time *T*, and the camera amplifies the received signal by a factor *A*. The received signal to noise ratio of a pure white color signal is γ (typical value from 300-1000). The image detector has a maximum output value I_{max} (typical value 255), and it quantizes its output to integer value between 0 and I_{max}. Furthermore, we assume the marking moves in the longitudinal direction with velocity *v*. Each module contains *N* pixels. The number of received pixel depends on the size, *L* x *H*, of each module, the relative 3D distance, (latitude *X*, longitude *Y*, vertical *Z*, units in meter), from the camera to the marking, the resolution of the camera, R megapixels per centimeter square (MP/cm^2), and the focal length, *f* in meter, of the camera lens camera.

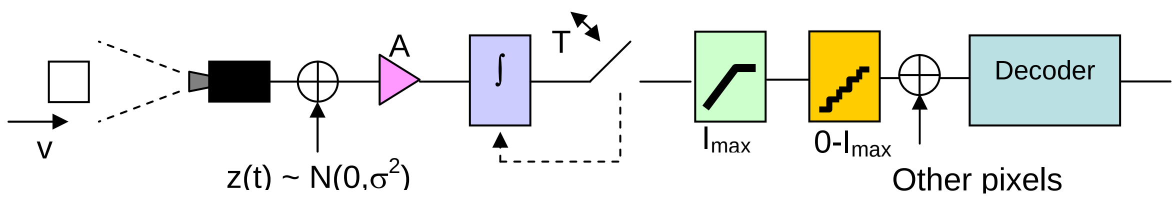

Figure 5 Model of Image Receiver

If we assume that the camera is aligned to the lane marking in the latitudinal dimension (*X* = 0), using orthographic transformation, we can find that $N = LH \frac{f^2}{R(Y^2 + Z^2)} 10^6$.

MOTION BLUR: DIRECT ENCODING

Under motion blur, the resulting frame is a convolution of the original lane marking image without motion and the point spread function. In signal processing, we can model the effect by a finite impulse response filter, as shown in Figure 6. In the figure, we assume that the vehicle speed is such that motor blur causes three adjacent modules to be summed up in the longitudinal direction. We also show two markings as example, and their respective observed outputs.

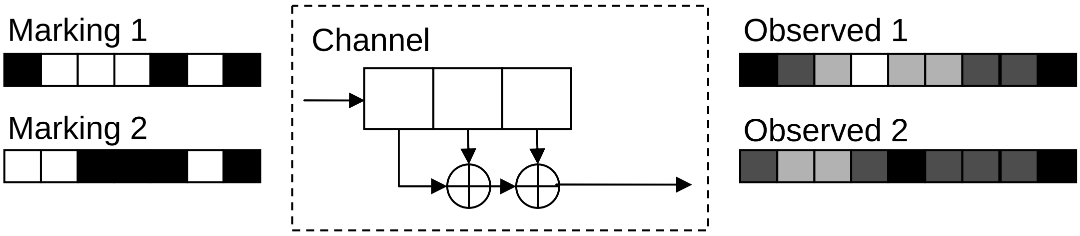

Figure 6 Effect of Motion Blur on Marking with Direct Encoding

With motion blur, the camera needs to adjust its amplification factor so that the maximum observed value does not go beyond I_{max}. Also, as the amount of blur increases, more levels of grayscale color would appear in the observed frame and it becomes harder to tell apart the various grayscale levels. By looking at the observed pattern, it is possible to find the most likely pattern that is present in the marking. Through analysis, we can upperbound the probability of error to $P_E \le (2^n - 1)[P_1^2 + P_1 (1 - P_1)]$, where

$$P_1 \approx Q\left(\frac{1}{\sqrt{2vT/L+2}} \left(\frac{1}{TN\gamma} + \frac{2vT/L+1}{24(vT/L+1)I_{max}} \right)^{-1/2} \right),$$

where $Q(x) = erfc(x/\sqrt{2})/2$ computes the tail of Gaussian distribution. In Figure 7, we see the analytical result of the lane marking assuming that Y = Z = 1m, R = 1MP/cm^2, I_{max} = 255, SNR γ = 300, and we assume that L = H (each module is square-shaped). In the left figure, we show how shutter interval changes the detection performance: when shutter interval is small, the amount of noise is high; when the shutter interval is large, the amount of blur becomes high, and it becomes increasing difficult to discern the different grayscale levels in the captured frame. In the right figure, we show how vehicle speed affects the decoding performance. As the vehicle speed increases, the size of the module also has to increase so to ensure the same decoding performance.

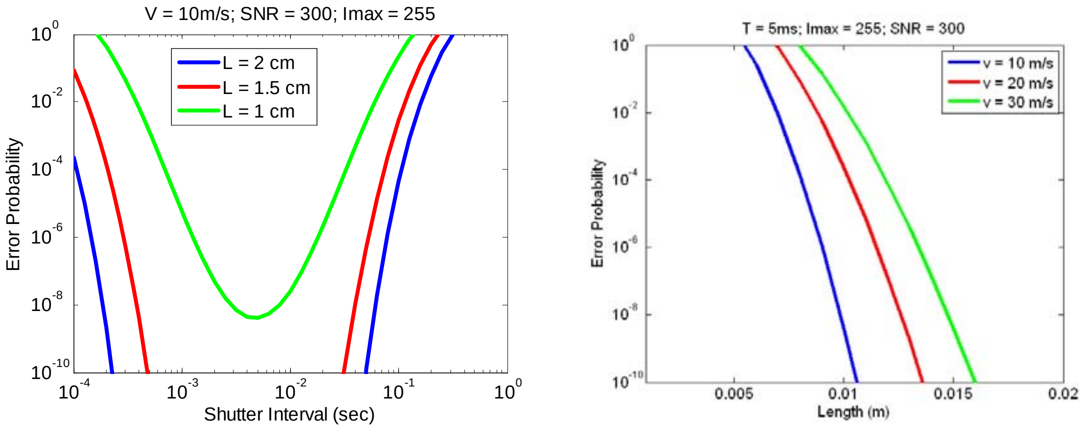

Figure 7 Analytical Result of Direct Encoding

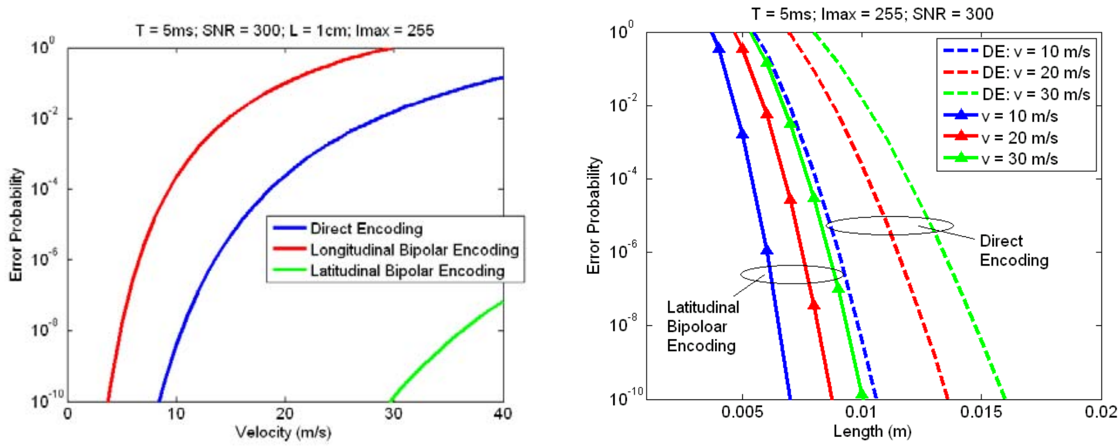

Figure 8 Performance of Latitudinal Bipolar Encoding

MOTION BLUR: BIPOLAR ENCODING

The use of bipolar encoding increases the code distance, in coding theory sense, of the observed frame. At the same time, since each sub-module has less number of pixels in the observed frame, the signal-to-noise ratio for each module is also less. By studying the minimum distance of these codes at the receiver, we obtain, for Longitudinal Bipolar Encoding, the upperbound on error probability is $P_E \le (2^n - 1)[P_1^2 + P_1(1 - P_1)]$, with

$$P_1 = Q\left(\frac{1}{\sqrt{2vT/L+2}}\left(\frac{2}{TN\gamma} + \frac{2vT/L+1}{24(vT/L+1)I_{\max}}\right)^{-\frac{1}{2}}\right).$$

And for Latitudinal Bipolar Encoding, the upperbound on error probability is $P_E \le (2^n - 1)\left[\sum_{i=b+1}^{2b}\binom{2b}{i}P_1^i(1-P_1)^{2b-i} + \frac{1}{2}\binom{2b}{b}P_1^b(1-P_1)^b\right]$, assuming b = vT/L is an integer, and

$$P_1 \approx Q\left(\frac{1}{\sqrt{2vT/L+2}}\left(\frac{2}{TN\gamma} + \frac{2vT/L+1}{24(vT/L+1)I_{\max}}\right)^{-\frac{1}{2}}\right).$$

An immediate observation is that the Longitudinal Bipolar Encoding cannot outperform Direct Encoding. This is because the minimum distance of Longitudinal Bipolar Encoding remains the same compared to Direct Encoding (Minimum Distance = 2); however, the energy in each sub-module decreases. On the contrary, Latitudinal Bipolar Encoding outperforms both Direct Encoding and Longitudinal Bipolar Encoding. The left side of Figure 8 shows a comparison of performance a function of vehicle speed for the three schemes. For L = H = 1cm, Latitudinal Bipolar Encoding performs well even when vehicles is moving in excess of 30m/s (over 108 km/h). The right side of Figure 8 shows the error performance of Latitudinal Bipolar Encoding only, at varying sizes of the module.

MOTION BLUR: DIFFERENTIAL ORTHOGONAL ENCODING

For simplicity, we will only show the result for Differential Orthogonal Encoding with only 4 sub-modules. In this case, every three adjacent modules in the longitudinal direction are orthogonal to each other. Inter-symbol interference occurs whenever the motion blur has a point spread function with components larger than three times the length L of the module. The upperbound on error probability is $P_E \leq 3(2^n - 1)[P_1^2 + P_1(1 - P_1)]$, and when $b > 2$,

$$P_1 \approx Q\left(\sqrt{\frac{3}{4}}\left(\frac{1}{TN\gamma} + \frac{2b+1}{b+1}\frac{1}{24I_{\max}}\right)^{-1/2}\right).$$

When the motion blur is less, for $b \leq 2$, we have

$$P_1 \approx Q\left(\sqrt{\frac{b+1}{4}}\left(\frac{1}{TN\gamma} + \frac{2b+1}{b+1}\frac{1}{24I_{\max}}\right)^{-1/2}\right).$$

An immediate observation from these equations is that, when the vehicle velocity is high, the value b is large, and P_1 converges quickly as a function of b. In other words, the effect of motion blur to the decoding performance is limited when the marking uses Differential Orthogonal Encoding. Figure 9 shows the performance of Differential Orthogonal encoding as a function of the module size. Note that such performance is almost independent of the vehicle velocity. This is because, unlike other schemes that depend highly on the ability to resolve the exact individual grayscale value for their maximum likelihood decoders (which becomes increasingly difficult as blur increases due to the quantization level of individual pixel of the camera output), the maximum likelihood performance of Differential Orthogonal Encoding depends only on the code distance after the differential detector (which is fixed regardless of blur). Also, the performance of the code is limited by the number of pixels present in each sub-module, due to the limited resolution of the camera. Certainly, this result idealistically assumes that the receiver has the capability to perfectly identify which pixels belong to which sub-modules. Synchronization error depends on module size and vehicle speed, and it can further degrade the performance of the overall system. At 30m/s, to reach bit error rate (BER) of 10^{-10}, Differential Orthogonal Encoding requires L = 3.6mm, this corresponds to a density of $77kb/m^2$. If a lane making is 15cm wide and 1m long, and inter-lane spacing is 1m, then at 30m/s, this is equivalent to a data rate of 5.7kbps per lane.

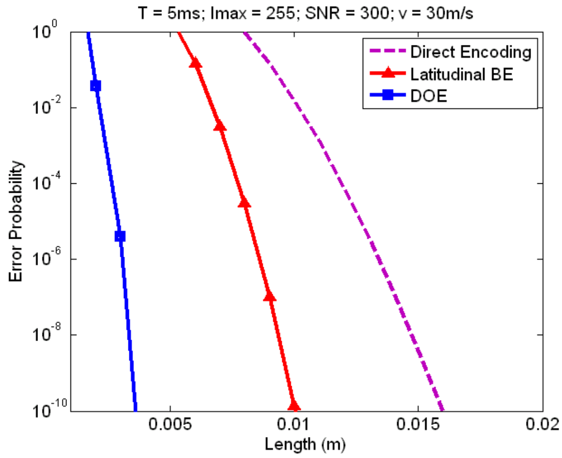

Figure 9 Performance of Differential Orthogonal Encoding

REAL-TIME EXPERIMENT

We tested the performance of the various encoding techniques in a real-time system as shown in Figure 10. We captured video of lane marking using black and white camera, and sent raw data to a PC over a FireWire cable. Instead of moving the camera, we mount a lane marking onto the runway of a treadmill, which has capability of rotating the runway at up to 10mph. The camera is located approximately at 92cm away in the longitudinal direction, and 134cm in the vertical direction, from the imaging position. In Figure 10, we show the captured images of the same lane marking, when the treadmill speed and shutter interval is set to different values. As the speed and shutter interval increases (from left to right of the figure), we see the amount of motion blur and the contrast of the resulting image also increase.

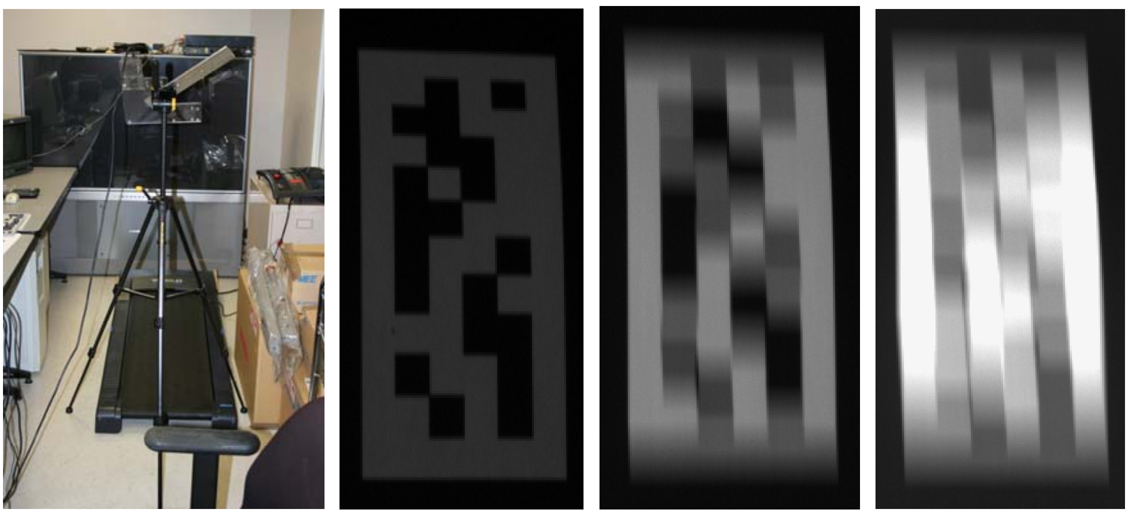

Figure 10 Experiment Set up and Observed Images

We implemented simple algorithms to decode Differential Orthogonal Encoding, Latitudinal Bipolar Encoding and a simple form of Longitudinal Bipolar Encoding. For Longitudinal Bipolar Encoding, to reduce the complexity of the maximum-likelihood decoder, we encode different bits at different latitudinal positions so that each longitudinal slide has only a single bit of information. We encode exactly 10 bits of information in the lane marking so that the overall sizes of the markings of all three encoding schemes are the same. At the bottom left of Figure 11, we see that Latitudinal Bipolar Encoding fails to decode as the lane marking moves faster than 4mph. This is because, to reduce complexity of receiver to enable real-time decoding functionality, our decoder performs bit-by-bit detection, rather than using maximum likelihood of the whole 10 bits. The Longitudinal code avoids motion blur by having one bit of information in each longitudinal strip. In the bottom right figure, we see the effect of synchronization on decoding performance. As the sizes of the modules shrink, our receiver fails to associate the right pixels to the right sub-module. The Longitudinal Bipolar Code fails first in this case due to the relatively small module length in the latitudinal direction of each sub-module.

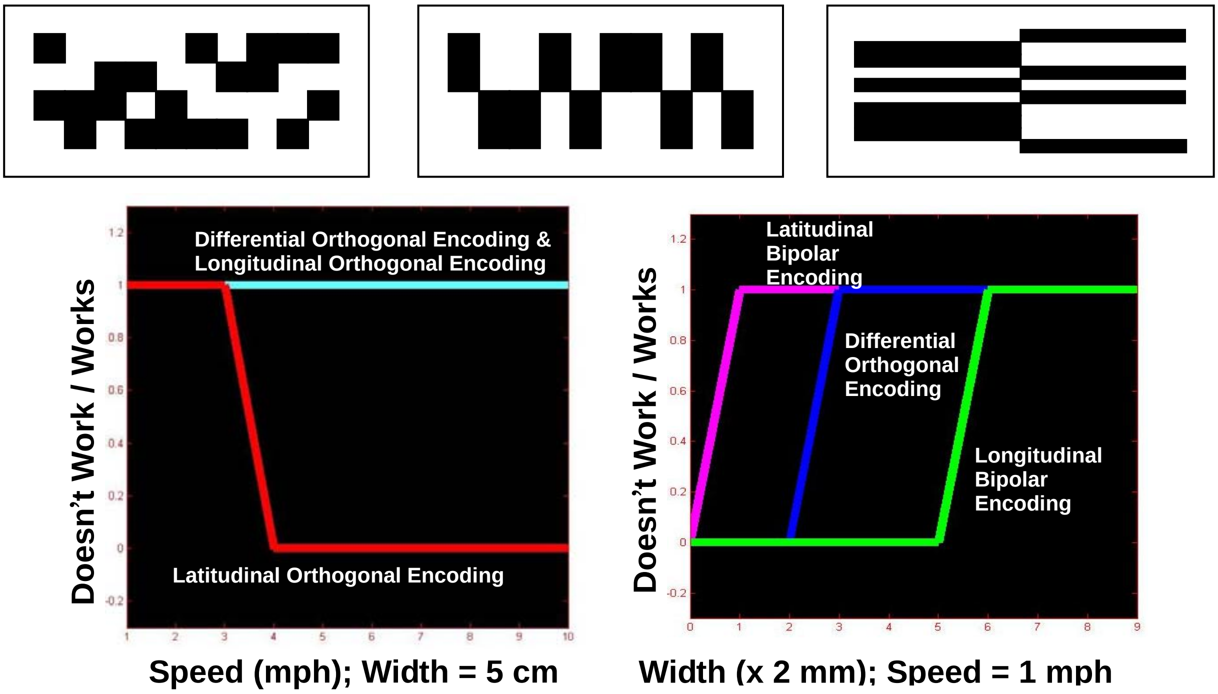

Figure 11 Experiment Set up and Observed Images

CONCLUSIONS

Through theoretical analyses and real-time experiments, we have demonstrated how four encoding techniques, Direct Encoding, Longitudinal Bipolar Encoding, Latitudinal Bipolar Encoding and Differential Orthogonal Encoding, are applied to store digital information in lane markings. Information in lane markings can supplement location information to existing geo-positioning systems, or to create new advertising services for transportation authority and local merchants. Besides lane markings, it is also possible to apply this technology for object or location tagging in any systems with high mobility. From theoretical analysis, amongst the four encoding schemes, Differential Orthogonal Encoding provides the best decoding performance in terms of vehicle speed and module size, allowing encoding density of $77kb/m^2$ for camera with $1MP/cm^2$ sensor located at 1m away, moving at 30m/s (108km/h). In real-time experiment, synchronization became an issue as the sub-modules became too small in size. Latitudinal Bipolar Encoding does not perform as well in theoretical analysis compared to Differential Orthogonal encoding; however, since bipolar encoding only divide a module into two, rather than four in the case of Differential Orthogonal Encoding, the decoder can better coupe with synchronization issue. Hence, at low speed (velocity at 1mph), Latitudinal Bipolar Encoding has superior performance. At high velocity, nonetheless, the encoding scheme fails due to its inability to deal with motion blur. Differential Orthogonal Encoding is the only code that works when vehicles moves at 100km/h. As future work, it is important to design synchronization pattern that helps decoder to properly associate a pixel value to a specific module. Training pattern is also important for decoder to identify true white and true black regardless of mobility, background lighting and possible localized reflection. Error correcting codes should also be considered so that data can still be corrected received after the road suffers wears and tears. Finally, further investigation should be conducted to ensure that the encodings works even when the blur direction is not purely in the longitudinal direction.

REFERENCES

[1] Wilson, H. W., Stearns, M. D., Koopmann, J. and Yang, C. Y. D., "Evaluation of a road-departure crash warning system." DOT HS 10 854, December 2007.

[2] Pomerleau, D., Jochem, T., Thorpe, C., Batavia, P., Pape, D., Hadden, J., McMillian, N., Brown, N. and Everson, J., "Run-off-road collision avoidance using IVHS countermeasures." DOT HS 809 170, December 1999.

[3] Shladover, S. E., "Automated driving: progress and prospects," *Proc. 17th World Congr. On Intell. Transport. Sys.*, Stockholm, Sweden, September 2009.

[4] De La Fortelle, A., "Cybercars: fully automated cars for the city of tomorrow," *Proc. 17th World Congr. On Intell. Transport. Sys.*, Stockholm, Sweden, September 2009.

[5] Höger, R., "Driver centric and context-sensitive automation to enhance safety and improve fuel efficiency," *Proc. 17th World Congr. On Intell. Transport. Sys.*, Stockholm, Sweden, September 2009.

[6] Wilson, A., "Auto cameras benefit from CMOS imagers", *Vision System Design*, May 2005.

[7] Czerniak, R. J., Reilly, P. R., *Applications of GPS for Surveying & Other Positioning Needs in Departments of Transportation*. National Cooperative Highway Research Program Synthesis 258. National Academy Press, June 1998.

[8] Hadamard, J. "Résolution d'une question relative aux déterminants." Bull. Sci. Math. 17, 240-246, 1893.