2012 Intelligent Vehicles Symposium
Alcalá de Henares, Spain, June 3-7, 2012

Veiling Luminance Estimation on FPGA-based Embedded Smart Camera

Constantino Grana, Daniele Borghesani, Paolo Santinelli, Rita Cucchiara
DII - Università degli Studi di Modena e Reggio Emilia, Italy
{name.surname}@unimore.it

Abstract— **This paper describes the design and development of a Veiling Luminance estimation system based on the use of a CMOS image sensor, fully implemented on FPGA. The system is composed of the CMOS Image sensor, FPGA, DDR SDRAM, USB controller and SPI (Serial Peripheral Interface) Flash. The FPGA is used to build a system-on-chip integrating a soft processor (Xilinx MicroBlaze) and all the hardware blocks needed to handle the external peripherals and memory. The soft processor is used to handle image acquisition and all computational tasks need to compute the Veiling Luminance value. The advantages of this single chip FPGA implementation include the reduction of the hardware rquirements, power consumption, and system complexity. The problem of the high dynamic range images have been addressed with multiple acquisitions at different exposure times. Vignetting, radial distortion and angular weighting, as required by veiling luminance definition, are handled by a single integer look-up table (LUT) access. Results are compared with a state of the art certified instrument.**

I. INTRODUCTION

Veiling Luminance is the veiling effect produced by bright sources or areas in the visual field that results in decreased visual performance and visibility. This phenomenon is particularly important for a vehicle's driver approaching to the entrance of a tunnel, and typically produces a reduction in driver's ability to perceive the presence of an obstacle or a slowdown caused by the traffic, especially in the first part of the tunnel (called *threshold zone*). The correct estimation of the veiling luminance can be exploited to design a measuring system combined with a power controller, in order to directly control the tunnel lighting, adapting automatically to the environmental conditions outside the tunnel by increasing or decreasing accordingly the lighting intensity inside a tunnel. This has a significant impact both in the energy consumption of the lighting system and in the traffic safety: a correct adaptation can avoid drivers to experience the hazardous "black hole" effect, being a very precious addition to its safety. The European Standard CIE 88/2004 (Italian UNI 11095 [1]) defines the threshold luminance in terms of the minimum amount of luminance necessary to let a driver to see an obstacle within its safety stop distance. This value can be computed as the veiling luminance, defined as the luminance that creates a disturbance on the driver visual within its safety stop distance.

In the literature, several authors studied the close relationship between the driving safety at the entrances of road tunnels, the illumination conditions at the tunnel entrances, the climatic conditions (haze, limpid sky), the speed of the vehicles approaching the tunnel entrances, the road surface (wet, dry). An important role is played by the European regulations CIE 88/2004 that set out recommendations concerning the daytime and the night-time lighting. It also describes the measures to be taken into consideration in order to adapt this lighting to the fluctuation to the external lighting. These works can provide interesting ideas, tools and solutions approaching the problem from the very basic steps of the design of the tunnel lighting system. Wang *et al.* [2], for example, addressed safety assessment criterions for the illumination design at tunnel entrances. In this work the lightness transition index has been established by analyzing the lightness transition character and the visual adaptation mechanism of drivers at tunnel entrances. The operating speed difference was selected as the average index to relate lightness transition character to driving safety. Parise *et al.* [3] focused on the high values that the electrical power system of a roadway tunnel could raise, and in accordance to the European regulations, suggest an adaptive criterion to design the support lighting system in order to mitigate the cost and energetic impact. In addition, they propose to limit the speed value to guarantee the safe conditions related to the maximum designated electric power. Leiato *et al.* [4] present a tool for automate the design of road tunnels lighting systems in which the best light type and its localization, according to a specified design objective along the tunnel, is automatically selected by the developed software framework. All the approaches proposed in the previous works could become even more effective by knowing the external luminance profile at the tunnel entrance instead of using statistic estimation of the climatic condition [3]. Cerotti *et al.* [5] report a work on a WSN-based system for adaptive, closed-loop control of lighting in road tunnels. The ability to match dynamically the lighting levels to the actual environmental conditions improves the tunnel safety and reduces its power consumption. In their work, the regulations determine the lighting inside the tunnel based on the veiling luminance at the entrance. The latter requires a dedicated external veiling luminance sensor.

In this paper we describe the design of a Veiling Luminance estimation system based on the use of a CMOS image sensor, fully implemented on FPGA. The FPGA is used to build a system-on-chip integrating a soft processor (Xilinx MicroBlaze). The soft processor handles the image acquisition and all computational tasks need to compute the Veiling Luminance value. The problem of the high

978-1-4673-2118-1/$31.00 ©2012 IEEE

dynamic range images have been addressed with multiple acquisitions at different exposure times. Vignetting, radial distortion and angular weighting, as required by veiling luminance definition, are handled by a single integer look-up table (LUT) access. The advantages of this single chip FPGA implementation include the reduction of the hardware requirements, power consumption, and system complexity.

The paper is organized as follows: section II describes the theoretical aspects related with the veiling luminance and its evaluation. Section III describes the distortion effects introduced by the optical system and how they are corrected. Section IV describes the architecture of the employed vision system. Finally, section V reports the experimental results and section VI draws the main conclusions.

II. The veiling luminance

The veiling luminance is composed by three different contributions:

$$L_v = L_{seq} + L_{atm} + L_{winds} \ . \tag{1}$$

where L_{seq} is the equivalent veiling luminance, L_{atm} is the atmospheric luminance and L_{winds} is the windscreen luminance. It can be directly estimated from the L_{seq}:

$$L_{wind} = 0.4 \cdot L_{seq} \tag{2}$$

whereas the atmospheric veiling luminance L_{atm} is usually estimated from tabulated data or measured before the tunnel launching using a bulky procedure as defined by the UNI 11095 standard [1].

The key quantity to compute for the correct L_v estimation is the equivalent veiling luminance L_{seq}.

A. Approaches to equivalent veiling luminance computation

The equivalent veiling luminance L_{seq} is directly caused by bright glare sources in the periphery of the point of view of the driver that reduces contrast visibility since light scattered in the lens obscures the fovea.

By law, two alternative specifications are available for the equivalent veiling luminance computation (or equivalent straylight luminance), L_{seq}.

- the summation of the luminance contributions of 108 sector of a polar diagram, centered in the middle of the tunnel at 1.5 meters high
- the summation of the luminance contributions of every light source, weighted by the angle with the driver's observation direction

The first approach obtains L_{seq} using a graph (similar to the one depicted in Figure 1) superimposed and centered to the tunnel scene, at a distance equal to the stopping distance. In each area the amount of luminance is considered equal to the straylight in the corresponding center. To compute the resulting value, the formulation proposed is the following:

$$L_{seq} = 0.51 \cdot 10^{-3} \cdot \sum_{i=1}^{9} \sum_{j=1}^{12} L_{ij} \tag{3}$$

where L_{ij} is the mean luminance emitted by the i-th ring at the j-th sector of the polar diagram.

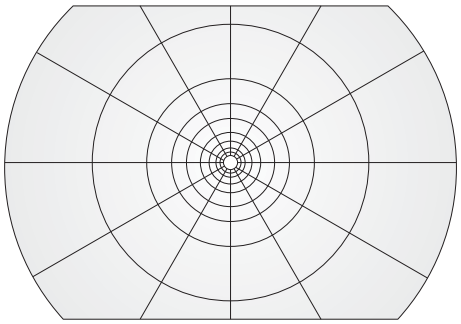


Fig. 1: Polar diagram similar to the one provided by UNI 11095 standard for the L_{seq} estimation process.

Following the indications provided by UNI 11095, the luminances L_{ij} of these regions can be computed manually using some printouts: the tabulated values of luminance for various typical surfaces are provided by empirical studies (CIE 88-1990 [6]), and are considered to represent with a sufficient frequency of time during the year for most common tunnel conditions.

The second approach considers more specifically every light source in the scene. The formulation proposed is the following:

$$L_{seq} = 9.2 \cdot \sum_{i=1}^{N} \frac{E_i}{\theta_i^2} \tag{4}$$

where E_i is the illuminance (i.e. light incident on a surface) produced on the human eye weighted by the angle θ_i between the direction of origin of the i-th light source and the observation direction of the driver.

In order to capture the photometric measures necessary for the second formulation, some specific tools should be required, for example a lux meter. The lux meter gathers illuminance information about its entire solid angle, it is expensive and it is not so stable especially in a real world scenario. Instead of the lux meter, a low cost and standard CMOS-based vision system can be used, which capture in every acquired pixel the illuminance information in the scene in the form of brightness (grayscale value of the acquired image). For this reason, a CMOS camera can be considered a "low cost" lux meter.

Given the current properties of the camera (such as focal length, exposure time, ...) and supposing some kind of linear relation between the grayscale value of the acquired image and the real illuminance of the scene, it is possible to come out with an automatic procedure to compute the equivalent veiling luminance in real time.

B. Illuminance Estimation

In the acquisition process, the illuminance information $E(y, z)$ is captured by the CMOS sensor and converted into electrical signals at discrete positions, representing the pixels of the acquired image. In this approximation, every pixel of the image is considered a light source whose gray value represents the amount of illuminance of the relative region of the real world. This approximation should be balanced by

the fact that a CMOS camera is not an ideal lux meter, and that the illuminance we estimate from the grayscale value is not obtained using an ideal reflectance reference target.

In literature, the relation between $E(y,z)$ and the gray level $g(i,j)$ in the image plane is well know, as presented in [7]:

$$g = t \int_{y-\frac{1}{2}\Delta y}^{y+\frac{1}{2}\Delta y} \int_{z-\frac{1}{2}\Delta z}^{z+\frac{1}{2}\Delta z} \int_{\lambda_1}^{\lambda_2} R_\lambda(\lambda) \cdot E_\lambda(y,z)\, \mathrm{d}\lambda \mathrm{d}z \mathrm{d}y \quad (5)$$

where, Δy and Δz are the pixel horizontal and vertical dimensions respectively, i and j are the indexes of the considered pixel in the image plane, $[\lambda_1, \lambda_2]$ is wavelength range of the image illuminance, E_λ is the image spectral illuminance, $R_\lambda(\lambda)$ is the digital sensor spectral responsivity and t is the exposure time. Equation 5 has been obtained considering an array composed of identical photosensors, neglecting any noise and supposing that the image illuminance $E(y,z)$ at a given λ did not change during the exposure time. Simplifying the formulation provided in Equation 5, we can highlight a linear relationship that links the grayscale value g to the illuminance E and the exposure time t, with a negligible error and an (unknown) multiplier parameter R depending on the sensor spectral responsivity and pixel size.

$$g = R \cdot E \cdot t \quad (6)$$

C. Overview of the algorithm

The first step of the algorithm is a necessary preprocessing aimed at the correction of the distortions introduced by the vision system. As it will be detailed in a later Section, a uniform region is acquired using the CMOS camera, then the obtained image is used to estimate and correct radial distortions and vignetting.

The second step is the determination of the parameters of the linear relation between the grayscale value of the image and the illuminance of the acquired scene, with a given exposure time. As shown in the Figures 2a and 2b, the light source is varied in intensity, for each condition an average reading of the illuminance values of a lux meter (positioned near the camera lens) is plotted with the relative average grayscale value of the acquired (and corrected) image, and the same procedure is adopted at different exposure times. The experimental results confirm a linear relationship between these two dimensions: for each useful exposure time provided by the camera, the relative parameters are saved.

Finally the third step computes the final luminance value. Since the aforementioned relation holds for different exposure times, Equation 6 can be used rewritten as illuminance in function of grayscale value:

$$E = R \cdot \frac{g'}{t} \quad (7)$$

where R is a constant parameter, g' is the grayscale value corrected of the distortions and t is the exposure time. Following the formulation proposed by the UNI 11095 standard [1], all the illuminance values computed considering pixels as light sources, weighted by their angle with the driver's point of view can be summed up, obtaining an estimation of the equivalent veiling luminance. In particular, since the device must adapt to different light conditions, a multi-esposures (HDR) procedure is employed: for each pixel in the scene, the illuminance value is computed at the minimum exposure time before saturation (either towards the dark or the bright).

III. Distortions correction

In computer vision, the basic camera model for image formation is the pin-hole camera model. It assumes that each image point is generated as a direct projection of a world point through the optical center. In practice, the model of real cameras is not so ideal, and many factors contribute to introduce different types of distortion that can slightly change the object's appearance. This problem can be sometimes neglected, but not in this case since the estimation of the L_{seq} depends directly on the grayscale values. For this reason, a preprocessing stage over the acquired image is necessary for a correct estimate.

A. Radial distortion

The radial distortion is a common non-linear distortion due to the lens (in particular wide-angle or low-cost lenses) and causes rectilinear (straight) lines that do not pass through the optical center to be represented as curved lines. As reported in [8], given a distorted image point $p_d = (x_d, y_d)$, we can obtain the undistorted image point $p_u = (x_u, y_u)$ as follows:

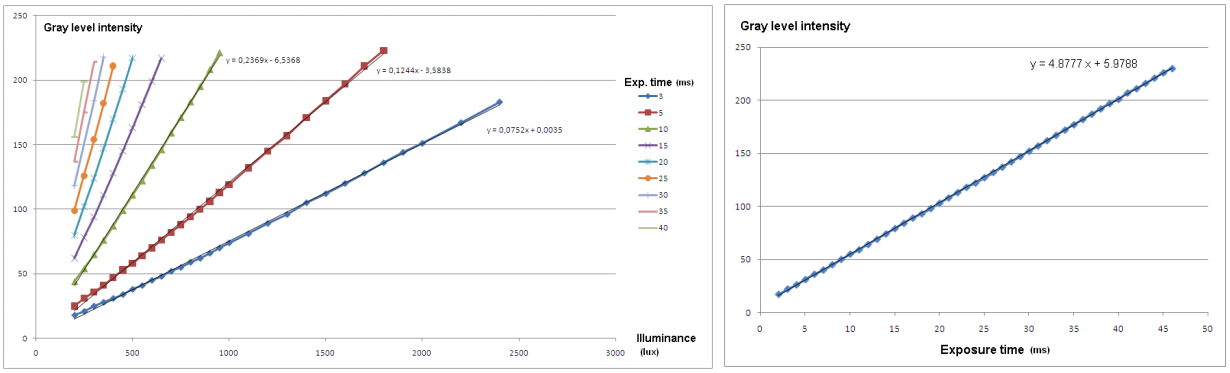
$$\begin{aligned} x_u = \;& c_x + (x_d - c_x)\left(1 + k_1 r_d^2 + k_2 r_d^4 + \ldots\right) + \\ & p_1\left[2\,(x_d - c_x) + r_d^2\right] + 2p_2\,(x_d - c_x)\,(y_d - c_y) \\ y_u = \;& c_y + (y_d - c_y)\left(1 + k_1 r_d^2 + k_2 r_d^4 + \ldots\right) + \\ & p_2\left[2\,(y_d - c_y) + r_d^2\right] + 2p_1\,(x_d - c_x)\,(y_d - c_y) \end{aligned} \quad (8)$$

where (c_x, c_y) are the coordinates of the center of distortion and $r_d = \sqrt{x_d^2 + y_d^2}$. As mentioned in [9], this formulation can be approximated and simplified by neglecting the tangential components and by considering only the lower-order components: in this way, more than 90% of the radial distortion can be corrected. The final formulation employed in this work is the following:

$$\begin{aligned} x_u = \;& x_d + (x_d - c_x)\left(k_1 r_d^2\right) \\ y_u = \;& y_d + (y_d - c_y)\left(k_1 r_d^2\right) \end{aligned} \quad (9)$$

The parameters of the radial distortion are computed in a semi-automatic fashion using the method described in [9]. Briefly, the method exploit the distortion itself to evaluate its magnitude: one or more straight lines (candidate lines) are automatically extracted from the image by means of the Hough Transform, then an iterative variation of the estimated distortion parameters is performed in order to maximize the straightness of the candidate lines. This method proved to be sufficiently robust to correct the radial distortion with the used setup.

Once the points have been mapped from the distorted to the corrected image, they can be used to compute the correct angular weight (as proposed by UNI 11095 standard [1]).



(a) (b)

Fig. 2: Linear relationship between grayscale value and the illuminance of the acquired scene for different exposure time values (a) and linear relationship between gray level intensity and exposure time (b)

B. *Vignetting correction*

The lens attached to the camera front introduce a strong darkening effect at the edges of the image, known as *vignetting*. Vignetting effect refers to a position-dependent loss of light in the output of an optical system, due mainly to the blocking of a part of the incident ray bundle by the effective size of the aperture stop; thus gradual fading-out of an image at points near its periphery results [10].

Vignetting correction begins by setting up a uniform white illumination source with a known input intensity level over a reference object with low specular reflection [11]. To accomplish this task, an integration sphere should be necessary, in order to have an uniform light source without distortions. This tool, very common in many photometric activities, is also very expensive, so it has been decided to approximate the sphere with an handmade *integration cube* (Figure 3), internally covered by mirrors and with two opal glass faces. The light of a source controlled in voltage passes through the input opal glass face, and the internal mirrors contribute to make the light look uniform all over the surface of the output opal glass face. The experiment for measuring illumination intensity response is carried out in a dark room to avoid ambient light interference. The camera is pointed toward the reference surface and intensity response at each pixel position is recorded.

Fig. 3: Integration cube

Subsequently, a correction factor at each pixel position is calculated by the following form [12]:

$$F(x,y) = \frac{\max(I_{ref})}{I_{ref}(x,y)} \tag{10}$$

where $I_{ref}(x,y)$ is an intensity value at (x,y) pixel position, $\max(I_{ref})$ is the maximum value of $I_{ref}(x,y)$, and $F(x,y)$ is a correction factor to be stored in a LUT at (x,y) position. Thereafter, an image captured with the same camera can be corrected by multiplying pixel values of the image with corresponding correction factors stored in the LUT:

$$I'(x,y) = I(x,y) \cdot F(x,y) \tag{11}$$

where $I(x,y)$ and $I'(x,y)$ denote intensity values at (x,y) pixel position before and after vignetting correction.

The correction procedure is inspired to the technique propose by [12]. In their approach, an hyperbolic cosine function is considered the most suitable to mimic the physical manifestation of vignetting effect:

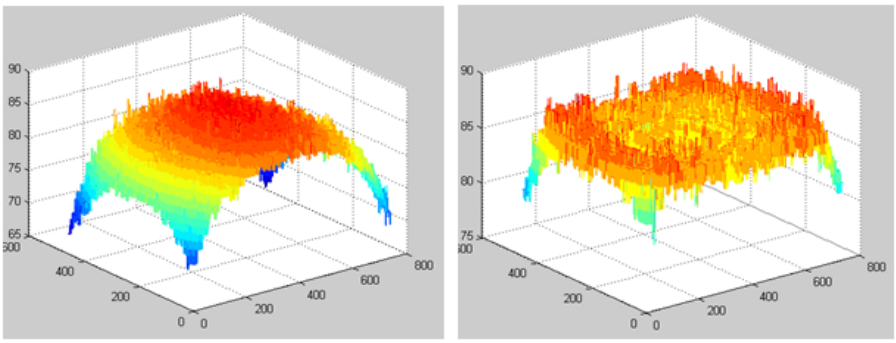
$$f(x,y) = cosh(r_x(x-x_0)) \cdot cosh(r_y(y-y_0)) + c \tag{12}$$

where r_x and r_y denote falloff rates along x and y axis of the image, while x_0 and y_0 denotes the image center.

Instead, in this implementation we have preferred to use a paraboloid function:

$$f(x,y) = a(x-x_0)^2 \cdot b(y-y_0)^2 + c \tag{13}$$

where a, b, c, x_0 and y_0 are the parameter to determine in order to model the intensity plateau. Analogously to [12], a nonlinear model fitting using the Levenberg-Marquardt optimization algorithm has been employed to find out the model parameters, and thus the image $F(x,y)$ of correction factors. The results of the vignetting correction are shown in Figure 4



(a) (b)

Fig. 4: Gray level spatial distributions of an image obtained from a uniform source with a known intensity level before (a) and after (b) vignetting correction.

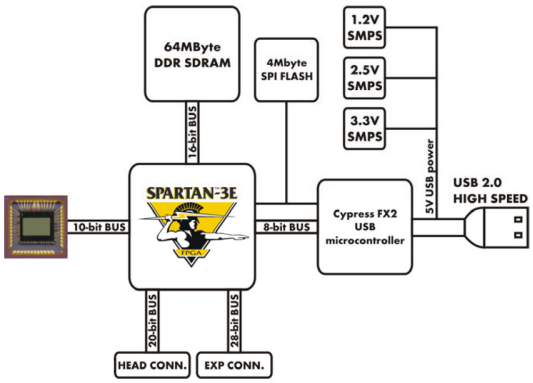


Fig. 5: Camera structure block scheme

IV. Architecture and Design

A. Hardware organization

The hardware employed to realize the whole system is the Cameleon camera, from OptoMotive [13]. This is an innovative USB camera based on FPGA system-on-chip. It is a ready-to-use solution together with the full source code of the complete system and the powerful HSDK (hardware-software development kit). The open architecture nature of this camera allows easy customization and integration of new features and functions. The system on chip is based on the MicroBlaze embedded processor soft core, it is a highly configurable reduced instruction set computer (RISC) optimized for implementation in Xilinx FPGAs. In this application it runs at 50 MHz clock frequency with data and instruction caches, integer divider and 64-bit multiplier units enabled. The soft processor is used to handle image acquisition and all computational tasks need to compute the veiling luminance value. The hardware system (Figure 5) is composed of 2 boards: the base board and the head board. The base board includes power supply, FPGA, DDR SDRAM, USB controller and SPI Flash. The whole hardware system is USB powered with triple DC-DC converters for maximal power efficiency (SMPS - Switching Mode Power Supply). The core is a 1.6 Mgates Spartan-3E FPGA, with 64 MB DDR SDRAM. The 16-bit wide DDR data bus running at 100 MHz offers a 400 MB/s peak bandwidth. The USB controller is the Cypress CY7C68013A (also known as FX2) which offers up to 36 MB/s of bandwidth in bulk mode. The FX2 is also connected to FPGA through I2C which is used as a command interface. The board also includes non-volatile SPI FLASH memory. The SPI bus is shared by FPGA and FX2 which is necessary to enable USB firmware upgrade and FPGA booting. The head sensor board houses the image sensor. It is the Aptina MT9V034 1/3"'CMOS monochrome sensor, 752 x 480 pixels, 64 frames per second at full resolution, with a global shutter. Global shutter means that exposure of all pixels starts and stops at the same time preventing motion blur for fast moving objects. The pixel clock is set to 27 MHz. The sensor uses 10-bit parallel interface for reduced system power consumption and electromagnetic noise.

B. Device utilization summary

The digital architecture described in this paper has been implemented and tested on the Xilinx FPGA available on the Camelon board, the Spartan-3E FPGA. It is composed by 1.6 million system gates, an array of 76x58 configuration logic blocks (CLBs) providing 14,752 Slices (1 CLB = 4 Slices) and a maximum of 231 Kbits distibuted RAM. Furthermore, it provides 36 18x18 bits multiplier blocks and 36 18-Kbit selected-RAM blocks for an amount of 648 Kbits RAM. Here we report the synthesis report, that is the number of instantiated components for the whole design.

```
Device utilization summary:
---------------------------
Selected Device : 3s1600efg320-4

 Number of Slices:              8868  out of  14752    60%
 Number of Slice Flip Flops:   10777  out of  29504    36%
 Number of 4 input LUTs:       16185  out of  29504    54%
    Number used as logic:      11099
    Number used as Shift regs:  694
    Number used as RAMs:       4392
 Number of IOs:                  111
 Number of bonded IOBs:           67  out of    250    26%
    IOB Flip Flops:               39
 Number of BRAMs:                 34  out of     36    94%
 Number of GCLKs:                 15  out of     24    62%
 Number of DCMs:                   3  out of      8    37%
```

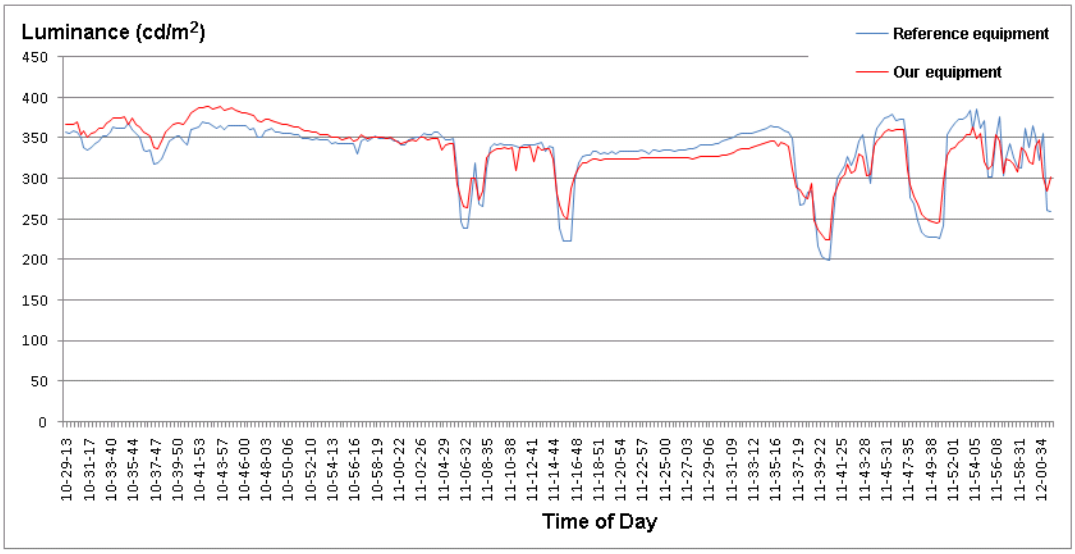


Fig. 6: Estimation of the veiling luminance at different time of day.

V. Experimental results

In order to evaluate the performance of the developed system, a comparison with a third part reference equipment has been done. The reference equipment provides a certified evaluation of the veiling luminance of the observed scene. The equipments have been aligned on the same target, obtaining measurements as fair as possible. The observed scene was a street in an average sunny day with some moments of cloud cover, which were reported by both systems. In particular, in the chart in Figure 6, a substantial correspondence between the estimation values of both veiling luminance equipments is reported. The average deviation between the two datasets is 4.51%. This value must be analyzed considering two primary sources of error: an observed 3% of error from the CMOS image sensor that indicates the fluctation of the average gray level obtained by observing a stable scene, and 2% of error due to the employed light meter.

VI. Conclusions

In this work, the design and development of a low cost veiling luminance estimation system based on a CMOS image sensor, fully implemented on FPGA has been proposed. The project is motivated by the fact that the veiling luminance is particularly important for a vehicle's driver approaching to the entrance of a tunnel, since it impacts on the driver's ability to perceive the presence of an obstacle or a slowdown caused by the traffic, and thus its safety. The formulation used to compute the luminance value follows the European Standard CIE 88/2004 (Italian UNI 11095 [1]). The proposed solution, based on an FPGA implementation presents advantages in terms of size, energy efficiency and costs that make it suitable to develop stand alone device capable to operate on the field close the street tunnels. The device for the veiling luminance computation, once combined with a power controller, can be used to increase or decrease accordingly the lighting intensity inside a tunnel.

References

[1] N. U. 11095, *Illuminazione delle gallerie*, Dec. 2003.

[2] W. Yan, G. Zhong-yin, and L. Zhi-gao, "Safety analysis for illumination design at tunnel entrance and exit," in *Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on*, vol. 3, may 2010, pp. 255 –259.

[3] G. Parise, L. Martirano, and S. Pierdomenico, "An adaptive criterion to design the lighting system in the road tunnels," in *Industry Applications Conference, 2007. 42nd IAS Annual Meeting. Conference Record of the 2007 IEEE*, sept. 2007, pp. 1244 –1248.

[4] S. Leitao, E. Pires, and P. de Moura Oliveira, "Road tunnels lighting using genetic algorithms," in *Intelligent System Applications to Power Systems, 2009. ISAP '09. 15th International Conference on*, nov. 2009, pp. 1 –6.

[5] M. Ceriotti, M. Corra, L. D'Orazio, R. Doriguzzi, D. Facchin, S. Guna, G. Jesi, R. Cigno, L. Mottola, A. Murphy, M. Pescalli, G. Picco, D. Pregnolato, and C. Torghele, "Is there light at the ends of the tunnel? wireless sensor networks for adaptive lighting in road tunnels," in *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, april 2011, pp. 187 –198.

[6] C. I. d. l. CIE, *Guide for the lighting of road tunnels and underpasses*, Vienna, Austria, USA, 1990.

[7] B. Jahne, *Practical Handbook on Image Processing for Scientific and Technical Applications, Second Edition*. Boca Raton, FL, USA: CRC Press, Inc., 2004.

[8] C. Slama, *Manual of Photogrammetry, Fourth Edition*. Falls Church, Virginia, USA: American Society of Photogrammetry and Remote Sensing, 1980.

[9] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani, "A computer vision system for in-house video surveillance," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 152, no. 2, pp. 242–249, Apr. 2005.

[10] E. Hecht, *Optics*, 4th ed. Addison-Wesley, 2001.

[11] G. Wyszecki and W. Stiles, *Color science: concepts and methods, quantitative data and formulae*, 2nd ed. John Wiley & Sons, Inc., 2000.

[12] W. Yu, "Practical anti-vignetting methods for digital cameras," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 4, pp. 975–983, Nov. 2004.

[13] . Optomotive., "Optomotive, mehatronika d.o.o., v murglah 229, si-1000 ljubljana, slovenia," http://optomotive.si.