A low-cost stereo-fisheye camera sensor for daylighting and glare control

Michael Kim¹², Iason Konstantzos³, Athanasios Tzempelikos¹²* ¹Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907 USA ²Center for High Performance Buildings, Ray W. Herrick Laboratories, Purdue University, 140 Martin Jischke Dr., West Lafayette, IN 47907, USA ³Durham School of Architectural Engineering and Construction, University of Nebraska – Lincoln, 1110 South 67th Street, Omaha, NE 68182, USA

* Corresponding author. E-mail: ttzempel@purdue.edu

Abstract. This study presents the development of a window-attached, low-cost fisheye-stereo camera sensor system for reliable daylighting controls in buildings. The stereovision sensor composed of two cameras can measure 3-D position and luminance of any exterior bright sources after photometric and geometric calibration. Also, by applying a neutral density sensor on one of the cameras and substituting the overflown pixels of the luminance map, a complete luminance map with a wide dynamic range can be retrieved. The new sensor exhibited promising results in validation of 3-D positioning and luminance acquisition, and can be used for real-time glare source detection, location and intensity measurements.

1. Introduction
Existing daylighting controls, even in their most advanced forms of glare-based, model-based controls, fail to ensure visual comfort in certain situations. For example, they lack the ability to address glare caused by small but extremely bright sources, such as specular sunlight reflections. This is a common case in dense-urban environments were glass curtain wall buildings are located close to each other [1]. During such instances, photometers cannot detect the extremely bright source observed by the occupants, since a discrete extreme luminance source of a small size does not necessarily lead to high vertical illumination on the eye.

To compensate for the lacking ability of existing control systems to “see” the exterior scene through the window, a High dynamic range image (HDRI) sensor based on a low-cost camera can be used [2,3,6]. A HDRI sensor was recently developed [4], able to capture the complete luminance distribution of the exterior scene and locate potential glare sources. HDRI is a technique that allows retrieval of the full luminance map by combining and calibrating camera photographs of a scene with multiple exposures [5,6] that allows precise measurement of per-pixel luminance from the image. The sensor was placed on the window and was successfully integrated into the building management system, which then controlled motorized roller shades to prevent potential glare sources from reaching the occupants’ eyes in real time. However, since the sensor was mono-vision, it could not estimate the depth (distance) information, which could be important since the algorithm cannot determine if the bright sources are in fact visible from the observer. Another problem with such systems is the limited maximum dynamic range, allowing to capture luminance values of only up to fifth order of magnitude. For some sources
such as the sun or highly specular reflections, this can result in errors in illuminance calculated from the HDR image, to be used as an input in predictive or data-driven models. Using an image sensor with a wider dynamic range can be a solution, however, their cost is rather prohibitive for wide adoption in buildings.

This paper proposes an improved version of the HDRI sensor that resolves the above issues, by adding an additional camera applied with a neutral density (ND) filter to constitute a fisheye-stereovision system. The sensor is able to retrieve 3-D Cartesian coordinates of the exterior scene relative to the sensor and capture the accurate luminance of the solar disk (or extreme bright sources) after calibration, and estimate the respective illuminance. The overall process is shown in Fig. 1.

2. Fisheye-Stereovision System

Two low-cost camera boards (Raspberry Pi Camera Module V2) are the main components of the stereovision system setup (Fig. 2). The camera board is based on the 8-megapixels CMOS sensor (Sony IMX219) and its operation is fully programmable with Raspberry Pi (Raspi) through a Python interface. To capture the full exterior scene through the window, a wide-angle fisheye lens was attached to each camera board to expand the field of view (FOV) to nearly 180 degrees. Two camera boards are horizontally integrated via a 3-D printed mount with the baseline distance of 15cm and connected to a single Raspi using IVMech multiplexer V2. In the left camera, an ND4 filter was applied between the image sensor and the fisheye lens (Figure 2 as in Strumpfel et al. (2004).

![Figure 1. Overview of fisheye-stereo vision sensing system.](image)

![Figure 2. Raspberry Pi 3B+ with IVMech V2 camera multiplexer.](image)

![Figure 3. Fisheye-stereo integrated with the 3-D printed mount.](image)

3. Fisheye-Stereovision Sensor Calibration

3.1. Photometric calibration

Photometric calibration is an essential process to acquire reliable luminance readings from HDR images largely composed in three steps; 1) camera response function (CRF) retrieval, 2) calibration factor (CF) estimation and 3) vignetting correction [6]. The CRF is the key component of HDR imaging, which refers to a non-linear curved function associating radiometric intensity to absolute values of luminance. In the present study, a CRF estimation algorithm proposed by [5] was used, using its OpenCV (computer vision library for C++/Python) implementation. Input to the CRF estimation is a set of low-dynamic images with different exposures. After creating HDR images with the estimated CRF, the calibration
factor (CF) that converts the HDR pixel intensities is calculated by fitting the HDR intensities to the absolute luminance measured by scientific-grade luminance spot-meter on the diffusive, neutral-colored surface. The luminance of a gray-colored square placed at the camera optical center (retrieved from the geometric calibration) was calculated from the un-calibrated HDR images and fitted to the luminance spot-meter reading by a multiplication factor. To reduce measurement and temporal error, CF was averaged from ten different measurements. The luminance measured by the spot-meter ranged from 4,800 cd/m² to 7,540 cd/m². The final part of the photometric calibration is the correction of the vignetting effect, a radial decrease of luminance observed in fisheye cameras of fast apertures that comprises the measurement accuracy of luminance. The process [6] was performed to obtain a 5th-degree polynomial correction function, which was then applied in HDR images via RADIANCE [7].

From the CF retrieved from the non-filter camera, the range of reliable luminances was estimated from the fastest shutter speed used in HDR creation. The HDR merging algorithm weights calibrated intensities from the low-dynamic images based on the isosceles triangle-shaded function that peaks at mid-intensity (128 for 24-bit RGB image) of the image. Thus, the luminance value corresponding to the mid-intensity calculated from the lowest and the highest exposures can be considered reliable, assuming proper photometric calibration. As shown in Figure 4, the measurable luminance ranges by two cameras are $165 \sim 117,000$ cd/m² (non-filtered) and $11,745 \sim 1.31 \times 10^6$ cd/m² (ND4). Since the overlapping ranges from 11,745 cd/m², bright sources above this range can be identified by both cameras. Although the luminance of the solar disk observed on earth can reach $1.6 \times 10^9$ cd/m², the ND4 filtered camera is able to capture most of the instances considering the observed luminance under solar incidence angles on building facades and typical visible transmittance of modern window systems. In any case, values in the order of $10^9$ cd/m² are much higher than any discomfort thresholds. The optical filter density above 4.5 will be sufficient for Raspi camera board to capture solar the disk in all circumstances.

![Figure 4. Measurable luminance range by fisheye-stereo vision.](image)

### 3.2. Geometric calibration

Geometric calibration of fisheye-stereo estimates camera parameters that determine the 2-D projection of real-world scenes into images. In the generic camera model [8] used in this paper (Fig. 5), there are two types of parameters for each camera (intrinsic parameters $K$ and distortion coefficients $d$) and stereovision-specific extrinsic parameters $[R|\ell]$. In the fisheye eye distortion model, the incoming ray of the scene, $P$, is refracted by a factor of $\theta_d$ (a polynomial function of $\theta$), where $\theta$ is the angle between the camera principal axis and the incoming ray from the scene $P$. The intrinsic parameters $K$ determine the pixel-scaling of the scene, skewness between the horizontal and vertical pixel array, and principal points (equivalent to fisheye distortion center). Extrinsic parameters $[R|\ell]$ are the rotation and translation between two camera frames constituting the stereovision. The calibration requires photographs of flat checkerboard patterns with a known square size, taken from many different views. 30 different image pairs were taken from 25.85mm-square checkerboards and put into an OpenCV implementation of fisheye-stereo calibration.
4. Positioning of glare sources

When the sensor identifies a bright source in the exterior exceeding a certain luminance threshold, its relative position to the camera can be estimated in 3-D cartesian coordinates via triangulation. For the 3-D positioning, the stereo-correspondence problem needs to be solved, referring to the task of finding a set of points in one image which can be identified as the same points in another image. One of the simplest and fastest methods of solving stereo-matching is template matching based on normalized cross-correlation (NCC). This method works well if the variance between corresponding pixels is low. However, there are several factors in fisheye-stereovision that can potentially create high variations within corresponding pixels; slightly different camera intrinsic parameters and fisheye distortion create different shapes and sizes of between the corresponding pixels. Also, mismatching can happen if multiple bright sources are present at the same time. A solution for the above issues is un-distortion and rectification of the fisheye-stereo images.

Correcting the fisheye distortion transforms a fisheye image into a distortion-free perspective image, where the straight lines in the real world remain straight. This can be done by back-projecting the pixels into rays and re-projecting them with any distortion-free perspective projection matrix. In the present study, the same perspective projection is used for both cameras to make the un-distorted image have equal scaling and principal points, and zero-skewness. Stereo-rectification aligns the camera frames to have the parallel principal axis purely translated in a horizontal direction with the baseline distance. Rectification simplifies the stereo-matching because it makes corresponding epipolar lines to appear as horizontal lines at the same height in the stereo image pairs. Note that any non-occluded corresponding pixels between two images lie on corresponding epipolar lines, which is a stereovision characteristic as illustrated in Fig. 6. In the rectified stereo image, the search space for template matching reduces into a single horizontal pixel array, at the same height of the template cropped from the image. In the present study, pixels above 30,000 cd/m² are identified as a source of disability glare [9], although this threshold is not absolute. Both luminance maps were converted into binary mask assigning pixels above the threshold as 1 and 0 for the other pixels. Then pixels assigned with 1 were grouped by connected-components labeling with 2-connectivity (neighboring pixels in 8 directions are grouped together). From the labeled binary mask, the bounding boxes for grouped glare source pixels are formed allowing paddings in four directions. Templates for the glare sources were then cropped from the ND4-filtered image (left) using the bounding boxes. Then the NCC was computed along the scanline (corresponding epipolar line to the template) to find the matching point with highest NCC (Fig. 7).

From the correspondence solved, estimation of the 3-D position of the scene is computed from disparity, the horizontal pixel difference between the corresponding pixels between two images (Fig. 8). 3-D position of any corresponding pixel with disparity \( d \) can be estimated by Equations (1) - (3), where \([X, Y, Z]\) are the 3-D coordinates of the corresponding scene in left-camera frame, \( x \) and \( y \) are pixel horizontal and vertical pixel displacement from the left-camera image’s principal point, \( f \) is the focal length of the new camera projection, \( b \) is the baseline distance between two cameras, and \( d \) is the disparity between corresponding pixels.
5. **Overflow correction for the solar disk**

A correction of the solar disk pixels is required because of “overflow” problems. The basic idea of overflow correction is replacing overflown pixels of the non-filtered luminance map with corresponding pixels of the ND-filtered luminance map containing higher luminance values. Like 3-D positioning, stereo-matching of the corresponding pixels is key to success. Finding correspondence for the solar disk can be simple for horizontal stereo images since image projection of the solar disk will be practically identical in the images due to its relatively infinite distance from the camera. Assuming the intrinsic parameters of two cameras are identical and two camera frames are completely leveled, the solar disk will appear at exactly the same pixel coordinates away from the principal point of each image. The relative squared error (RSE) between pixel-scaling factors (intrinsic parameters), the distortion function \( \theta_d \) in \([0, \pi/2]\), and rotational matrix \( R \) and \( I_{3 \times 3} \) (identical camera angles) are extremely small (6.0E-0.5, 8.3E-5, and 1.35E-5 each). Thus, it can be safely assumed that the solar disk will appear at the same image coordinates if the principal points of the images are aligned, simplifying the template matching needless of un-distortion and rectification. The overall overflow correction logic is as follows: 1) from the ND-filtered luminance map, detect the solar disk pixels using threshold of 10 \( 7 \) cd/m\(^2\), determined by a cloud-covered reading from the luminance map; 2) extract the template from the pixels same as done in 3-D positioning; 3) compute NCC in a padded square pixel area centered at equal image coordinates of the template after aligning principal points of two images; and 4) find the pixel with the highest NCC and replace the overflown pixels with the template (Fig. 9).

![Figure 8. 3-D Positioning from the disparity.](image)

![Figure 9. Overflow correction for solar disk pixels.](image)

6. **Sensor validation**

The 3-D positioning and complete HDR mapping with overflow correction were validated. The 3-D positioning was validated by comparing estimated stereo-depth to the known distance of an artificial glare source (reflector) from the window measured with a calibrated laser distance meter. The RSE of estimated depth was 4.20% in 13 measurements ranging from 2~10 m (Fig. 10). Since we are not able to directly measure the luminance of the sun, overflow correction was validated through vertical illuminance (E\(_v\)) by integrating the pixel luminance with the configuration factor. The E\(_v\), calculated from the overflow-corrected luminance map was compared to scientific-grade photometer readings, attached right next to the fisheye-stereovision sensor (Fig. 3). In 44 points measured with the sun visible, the overflow correction improved the accuracy significantly, reducing the RSE from 43% to 5.2% (Fig. 11).
7. Discussion and conclusion
The proposed low-cost fisheye stereovision sensor can detect potential glare sources, locating them with reasonable accuracy in real time. Also, overflow correction of the solar disk pixels allows lowering the cost by using non-wide dynamic range image sensors and can replace conventional window photosensors while adding more capabilities. Future work includes examining and improving the robustness of the sensor in more challenging conditions, including positioning of highly slanted glare sources or handling cases where the number of detected glare sources is very high. The window-attached fisheye-stereovision sensor allows the implementation of more advanced control algorithms with real-time glare source detection, location and intensity measurements, important for reliable daylight control operation and measurement-aided simulation. In addition, accurate luminance mapping can be used for visual preference-based modeling and control.

References
[1] Suk J Y, Schiler M and Kensek K 2016 Absolute glare factor and relative glare factor based metric: Predicting and quantifying levels of daylight glare in office space Energy Build. 130 8–19
[2] Wu Y, Kämff J H and Scartezzini J L 2019 Design and validation of a compact embedded photometric device for real-time daylighting computing in office buildings Build. Environ. 148 309–22
[3] Wu Y, Kämff J H and Scartezzini J L 2017 Characterization of a quasi-real-time lighting computing system based on HDR imaging Energy Procedia 122 649–54
[4] Kim M, Konstantzos I and Tzempelikos A 2018 A New Control Framework For The Visual Environment Based On Low-Cost HDR Luminance Acquisition. Proc. of 5th International High Performance Buildings Conference at Purdue, July 2018.
[5]Debevec P E and Malik J 1997 Recovering high dynamic range images proceeding of the SPIE: Image Sensors vol 3965 pp 392–401
[6] Inanici M N 2006 Evaluation of high dynamic range photography as a luminance data acquisition system Light. Res. Technol. 38 123–36
[7] Ward G 1994 The RADIANCE lighting simulation and rendering system ’94 SIGGRAPH conference pp 459–72
[8]Kannala J and Brandt S S 2006 A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses IEEE Trans. Pattern Anal. Mach. Intell.
[9]Jakubiec J A and Reinhart C 2014 Assessing Disability Glare Potential of Reflections from New Construction Transp. Res. Rec. J. Transp. Res. Board 2449 114–22
[10]Stumpfle J, Tchou C, Jones A, Hawkins T, Wenger A, Debevec P 2004 Direct HDR capture of the sun and sky. Proc. of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa, 145-149