

ORIGINAL RESEARCH

An algorithm for glare detection via photometric, colorimetric, and global positioning features

Mehran Andalibi^{*1}, Kousuke Kawai², Damon Michael Chandler²

¹Department of Aerospace and Mechanical Engineering, Embry-Riddle Aeronautical University, Prescott, AZ, United States

²Department of Electrical and Electronics Engineering, Shizuoka University, Japan

Received: May 14, 2020

Accepted: June 16, 2020

Online Published: July 19, 2020

DOI: 10.5430/air.v9n1p12

URL: <https://doi.org/10.5430/air.v9n1p12>

ABSTRACT

Biological and automated vision systems which use digital video for navigation depend on the video to be of sufficient quality in order to extract reliable information that can inform the guidance and/or other decision-making processes. Although systems are available for detection and mitigation of digital distortions (e.g., compression, packet loss), detection and mitigation of natural distortions such as glare, rain, and fog have received much less attention. In this paper, we address the issue of glare detection in a single captured frame. We propose an algorithm which uses a combination of simple and efficient photometric, colorimetric, and GPS features to detect the location and spatial extent of glare within captured images. Specifically, feature maps using lightness, saturation, contrast, and color distance are computed, combined, and then, refined based on the sun's predicted location from the GPS information. In addition, we present a new ground-truth database for glare detection, in which the location, extent, and severity of glare was rated by human subjects for a collection of images. Testing on our ground-truth database revealed that the proposed algorithm can reliably detect the locations and spatial extents of glare sources in a variety of images based on subjective ratings and well-known quantitative measures.

Key Words: Glare, Lightness, Saturation, Contrast, Color distance, Global positioning system, Subjective glare database

1. INTRODUCTION

A very common occurrence during navigation is the presence of overly intense light that can overwhelm the driver or the visual sensors of an autonomous ground vehicle, and thereby reduce visibility. When dazzled by a light source, typical reactions by the drivers are (1) flipping down the sunshades, which is not always effective, especially when the glare is caused by a low-positioned external light source, e.g., low sun or direct light from the headlights of an oncoming vehicle;^[1] (2) using one hand to occlude the strong light source from the eyes of the driver, which is not safe and can block the view of some important obstacles on the road;

and (3) wearing sunglasses, which has comfort issues.^[2] For autonomous vehicles, which depend on onboard cameras for navigation, none of these reactions is possible, thus resulting in the partial or full loss of environment perception and subsequent failures.

Although the detection of glare is an important problem, there are relatively few publications on glare detection using digital cameras.^[3–10] The existing approaches to glare detection generally rely on one or two simplistic image properties, followed by ad-hoc thresholding. The obvious and most commonly used feature is light intensity,^[3,4] which can certainly detect the light source(s), but which cannot differen-

*Correspondence: Mehran Andalibi; Email: andalibm@erau.edu; Address: Department of Aerospace and Mechanical Engineering, Embry-Riddle Aeronautical University, Prescott, AZ, United States.

tiate these sources from other bright regions (e.g., clouds), as demonstrated in Figure 1. Note that these bright regions also saturate the camera CCD array, and thus a hard threshold

cannot separate them from the main glare source.



Figure 1. Using light intensity alone to detect the glare regions also captures bright, non-glare regions. Left: Original image. Right: Binarized intensity values using a threshold of 95% of the maximum value

To overcome this limitation, other approaches supplement intensity with hue, followed by spatial filtering or transformations.^[5,6] However, the color of the sun changes based on its elevation; thus, hue alone often fails to delineate the boundaries of the glare region(s), particularly when the image is subject to compression, as demonstrated in Figure 2. In our previous work,^[9] we used saturation instead of hue, taking advantage of the fact that glare regions typically have

low saturation; however, many other regions in the scene also have low saturation. Instead of analyzing the scene, another class of approach has focused on analyzing an image of the driver's face.^[7,8] Unfortunately, this approach cannot be used in autonomous systems nor in scenarios where the driver's environment is different from the actual to-be-driven location (e.g., in remote navigation).

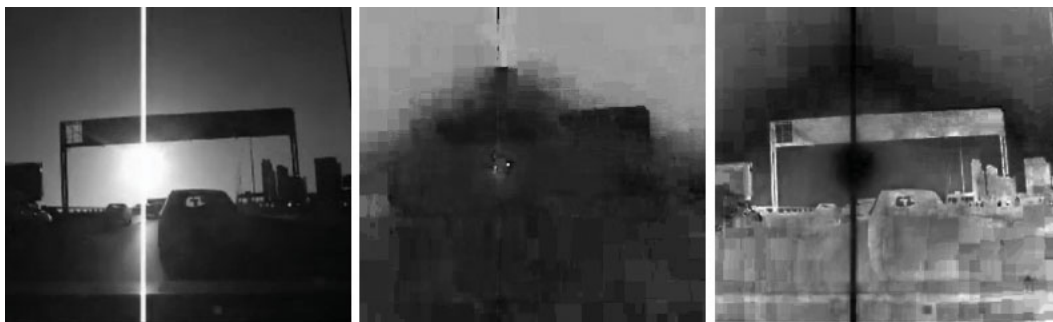


Figure 2. HSV color space components of an image with glare. Left: Intensity values (V channel). Middle: Hue values (H channel). Right: Saturation values (S channel). Previous approaches have used intensity and hue; part of our approach uses intensity and saturation, the latter being more reliable at detecting the spatial extent of each glare region

Deep learning has been utilized in the work by Jatzkowski et al.^[13] to determine whether overexposure (which is not always as detrimental to the driver's vision as glare) exists in an image. Although this work shows relatively high accuracy, it cannot locate the overexposure source in the image and so, it cannot be used for glare suppression. Furthermore, subjective ratings were not reported to confirm the visual impact of the overexposure as judged by different observers. The authors mentioned that there is no large database on glare images and the number of training images (almost 12,000) they

used is not sufficient for training a deep neural network. They also admitted the problem of bias in training a convolution network and the fact that their training set only covers a portion of what can be seen in real-life conditions, which makes it impossible to guarantee this method's accuracy when being tested on different images.

In this paper, we present an algorithm for glare detection that can detect the presence of glare from a single frame, and which can be used in autonomous or remote navigation

settings (i.e., which does not rely on the analysis and co-presence of a human driver). Our approach employs an adaptive combination of features which remain computational simplistic, but which can reliably detect both the location(s) and extent(s) of the glare region(s). We specifically use a combination of: (1) the intensity, saturation, and local contrast of the input frame; (2) the solar azimuth and elevation computed based on the position and heading information from the GPS (used under daylight conditions); and (3) the difference in color from the expected sun's color given the solar elevation information. These data are used to generate a glare occurrence map that indicates the center location(s) and extent(s) of the glare region(s).

The main contributions of this work are as follows: To the best of our knowledge, our approach is the first to use a combination of photometric, colorimetric, and GPS information to perform glare detection. For the photometric/colorimetric approach, we propose a combination of lightness, color distance, saturation, and local contrast that performs more reliably than previous approaches, and which does not require facial analysis of the assumed co-located driver. In addition, to the best of our knowledge, our approach is the first to make use of the sun's azimuth and elevation to further refine the localization process. Furthermore, we have created a large

database of glare images with human subjective markings of the glare center, boundary, and visibility impairment ratio, and we have analyzed the performance of our algorithm on this database. This is the first such database of glare images with subjective ratings, which can be used as a benchmark for evaluation of glare detection algorithms.

This paper is organized as follows. In the Algorithm section, we provide details of the algorithm. In the Ground-Truth Glare Database section, we describe the procedures and images used in our glare-rating experiment. In the Results and Analysis section, we evaluate and discuss the performance of the algorithm. General conclusions are provided in the Conclusions section.

2. METHOD

In this section, we describe the proposed algorithm, which operates by using a combination of image-based and GPS-based features. A block diagram of the algorithm is shown in Figure 3. The input is an RGB digital image (or single video frame), along with GPS information (if available). The output is a map denoting the location and extent of the glare. In the following subsections, we describe the algorithm's two main analysis stages.

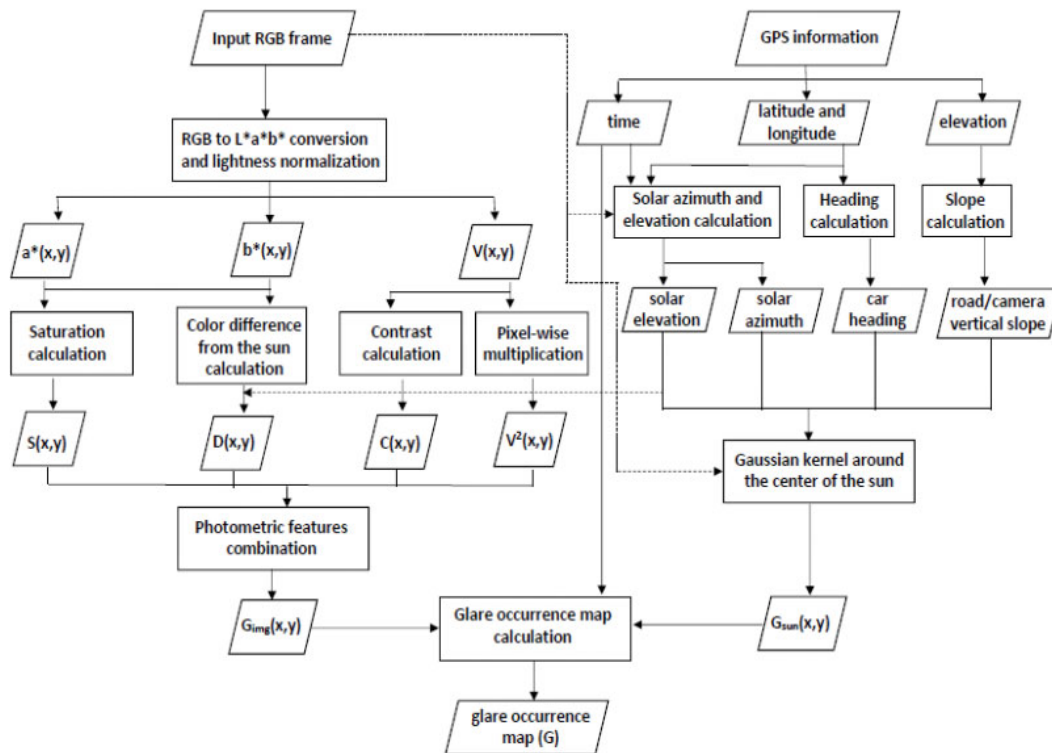


Figure 3. Block diagram of the proposed algorithm. The input is an RGB digital image or single video frame, along with GPS information (if available). The final output is a map denoting the location and extent of the glare. See Section 2 of the text for details.

2.1 Image-based glare occurrence map

Glare regions in an image often have characteristic photometric and colorimetric properties that can be used to detect glare regions given an input image/frame. The most obvious visual feature for glare detection is brightness. Regions in an image, which have the greatest intensity, are good candidates for glare regions. However, high intensity alone can give rise to false positives; thus, in addition to intensity, we also employ saturation and contrast features. Specifically, glare regions not only tend to be of high intensity, but also generally have low color saturation and low luminance contrast. Finally, we observe that the color of the sun can be estimated based on the date, time, and global position; thus, we employ color distance as a fourth feature. In this section, we describe how these features are computed and combined into a predicted glare occurrence map.

2.1.1 Lightness

To estimate the perceived intensity at each location, we use the lightness channel from the CIE 1976 (L^*, a^*, b^*) color space (CIELAB). Let $I_{R'G'B'} \in Z^{W \times H \times 3}$ denote the input RGB digital image with pixel values in the range from $[0,1]$, where W and H denote the image's width and height in pixels respectively. We convert the image into the CIELAB color space assuming sRGB as the input color space and D65 (daylight) as the illuminant.

Let R', G', B' denote the nonlinear $R, G,$ and B color values at a particular location of the input image. The conversion to CIELAB is implemented by first linearizing the R', G', B' values to be proportional to light energy, assuming the sRGB color space:

$$A = \begin{cases} \frac{A'}{12.92} & , A' \leq 0.04045 \\ (\frac{A'+0.055}{1.055})^{2.4} & , A' > 0.04045 \end{cases} \quad , \text{ where } A = R, G, \text{ or } B \quad (1)$$

The linearized I_R, I_G, I_B values are then converted to the CIE XYZ color space as:

$$X = 0.412453 \times R + 0.357580 \times G + 0.180423 \times B \quad (2)$$

$$Y = 0.212671 \times R + 0.715160 \times G + 0.072169 \times B \quad (3)$$

$$Z = 0.019334 \times R + 0.119193 \times G + 0.950227 \times B \quad (4)$$

Finally, the L^*, a^*, b^* values are given by:

$$L^* = 116 \times g\left(\frac{Y}{Y_r}\right) - 16 \quad (5)$$

$$a^* = 500 \times [g\left(\frac{X}{X_r}\right) - g\left(\frac{Y}{Y_r}\right)] \quad (6)$$

$$b^* = 200 \times [g\left(\frac{Y}{Y_r}\right) - g\left(\frac{Z}{Z_r}\right)] \quad (7)$$

where $X_r = 0.950456, Y_r = 1, Z_r = 1.088754$ are the CIE XY Z tristimulus values of the D65 reference white point; and the function g is given by:

$$g(t) = \begin{cases} t^{\frac{1}{3}} & t > 0.008856 \\ 7.787 \times t + \frac{16}{116} & , \text{ otherwise.} \end{cases} \quad (8)$$

For input R', G', B' values in the range $[0, 1]$, L^* has a possible range of $[0, 100]$, and a^* and b^* both have a possible range of $[-110, 110]$.

Note that equations (1) to (8) and all the coefficients used in them are standard conversion formulas between RGB and CIELAB color spaces.

We compute $L^*(x, y), a^*(x, y), b^*(x, y)$ for each pixel (x, y) of the input image. Our brightness feature map, denoted by $V \in Z^{W \times H}$ is given by the lightness values normalized to span the range $[0, 1]$:

$$V(x, y) = \frac{L^*(x, y)}{100} \quad (9)$$

for all pixel coordinates x and y in the ranges and $x \in [0, W - 1]$ and $y \in [0, H - 1]$.

The second column of Figure 4 shows the resulting lightness maps for several example input images (shown in the first column of Figure 4); larger values denote higher lightness.

2.1.2 Saturation

Regions with low color saturation are also good candidates of the central glare sources. We use the a^* and b^* values to obtain the color saturation of each pixel. Let $S \in R^{W \times H}$ denote this saturation feature map, which is computed as follows:

$$S(x, y) = \frac{\sqrt{(a^*(x, y))^2 + (b^*(x, y))^2}}{110\sqrt{2}} \quad (10)$$

where the denominator $110\sqrt{2}$ serves to normalize the $S(x, y)$ values to lie in the possible range of $[0, 1]$. Equation (10) is applied for all pixel coordinates x and y in the ranges $x \in [0, W - 1]$ and $y \in [0, H - 1]$.

The third column of Figure 4 shows the resulting saturation maps for the example input images; note that reverse saturation

maps ($1 - S$) are shown in which larger values denote lower saturation.

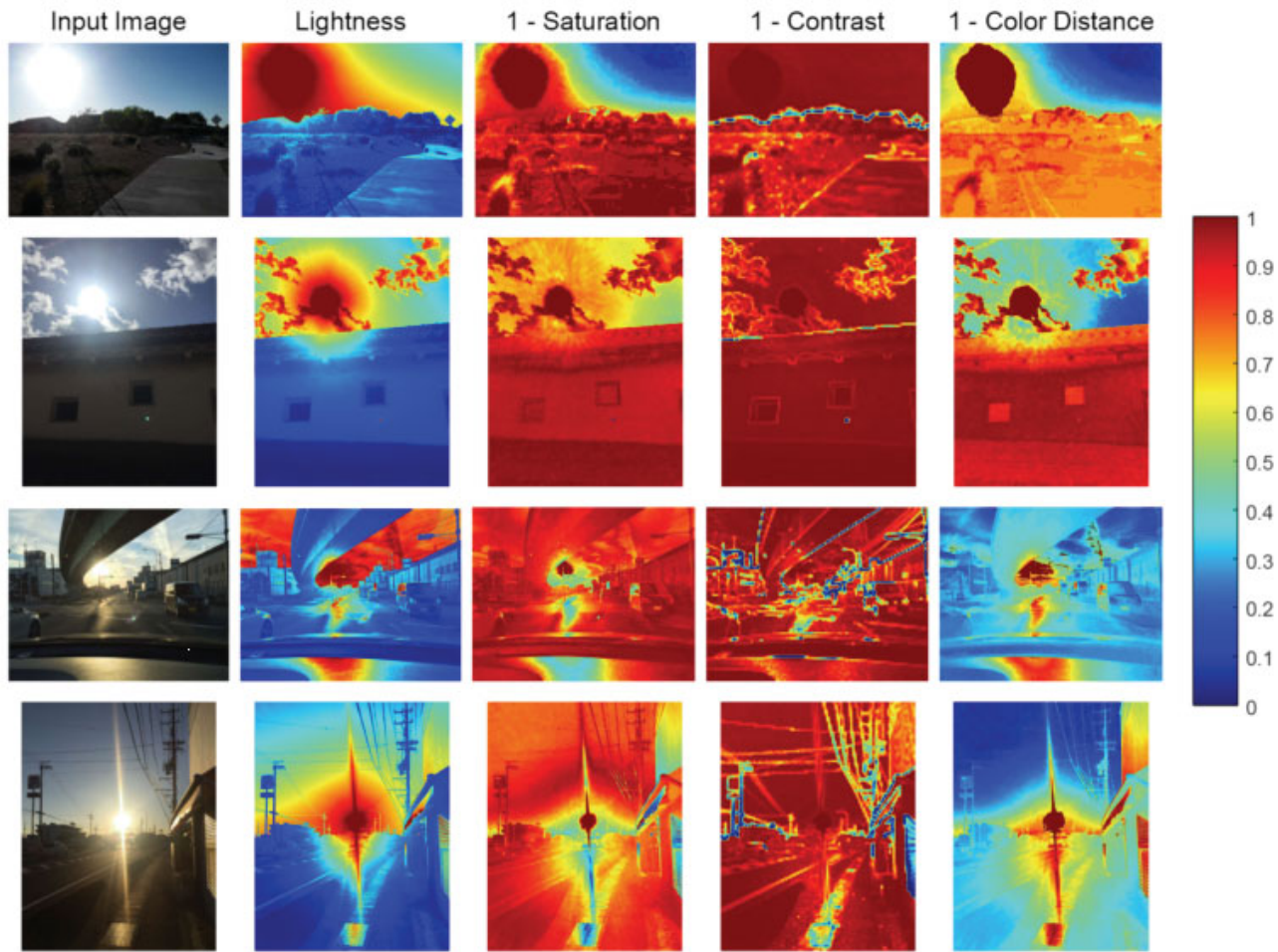


Figure 4. Four example input images and the resulting lightness, saturation, contrast, and color-distance maps. Note that the leftmost three maps are shown with a reverse intensity scale (i.e., $1 - M$, where $M = S, C,$ or D) so that larger map values denote greater predicted glare.

2.1.3 Contrast

Glare regions also tend to appear to be of low luminance contrast, due in part to the limited dynamic range and compressive nonlinearity used in most imaging devices (including the human visual system). We measure the root-mean-square contrast^[11] locally for each 8×8 block of V , collecting the results into a map. Let $C \in R^{W \times H}$ denote this contrast feature map, which is computed as follows:

$$C(x,y) = \frac{\sqrt{\frac{1}{16 \times 16 \sum_{x',y'} [L(x',y') - L(x,y)]^2}}}{\max\{10, \bar{L}(x,y)\}} \tag{11}$$

where $L(x,y) = [0.02874 \times 255 \times V(x,y)]^{2.2}$ denotes the estimated luminance; where (x',y') denote the coordinates corresponding to the 8×8 block centered at (x,y) ; and

$\bar{L}(x,y)$ denotes the mean luminance of that block. We use a lower limit of 10 in the denominator to avoid disproportionately large contrasts. We use non-overlapping blocks, and thus Equation (11) is applied for every 8^{th} pixel coordinate x and y in the ranges and $x \in [0, W - 1]$ and $y \in [0, H - 1]$. (i.e., $x = [0, 7, 15, 23, \dots]$ and $y = [0, 7, 15, 23, \dots]$). The remaining pixels are interpolated via bicubic interpolation.

The fourth column of Figure 4 shows the resulting contrast maps for the example input images; note that reverse contrast maps ($1 - C$) are shown in which larger values denote lower contrast.

2.1.4 Color distance

The color of the sun changes depending on the position of the sun within the sky. At its highest position, the sun appears white, whereas a setting/rising sun appears much more

orange. As we discuss later in Section 2.2, it is possible to compute the elevation angle of the sun given the date, time, and GPS information. We use this angle to compute an estimate of sun's average b^* value, and then compute a color distance map based on the difference between each pixel's b^* and the sun's average b^* .

Let e_{sun} denote the solar elevation angle (see Section 2.2). From e_{sun} , we use the following the empirically obtained mapping to estimate the sun's average b^* , denoted by b_{sun}^* , as follows:

$$\hat{b}_{sun}^* = 0.26 + \frac{71.74}{1 + \left(\frac{e_{sun}}{7.47}\right)^{2.44}} \quad (12)$$

Then, let $D \in R^{W \times H}$ denote a color distance (specifically, a b^* difference) feature map, which is computed as follows:

$$D(x, y) = |b^*(x, y) - \hat{b}_{sun}^*| \quad (13)$$

followed by two adjustments: (1) normalization of D to span the range $[0, 1]$, and then (2) explicitly setting $D(x, y) = 0$ for extremely bright locations indicated by $V(x, y) > 0.975$.

The fifth column of Figure 4 shows the resulting color-distance maps for the example input images; note that reverse color-distance maps ($1 - D$) are shown in which larger values denote smaller distance.

2.1.5 Combination of feature maps

To detect regions that are jointly of high lightness, low saturation, low contrast, and low color distance, we combine the individual feature maps to obtain a single image-based glare occurrence map, which we denote as G_{img} . For this combination, we use a simple pointwise multiplication, as follows:

$$G_{img} = V^2 \times (1 - S) \times (1 - C) \times (1 - D)^2 \quad (14)$$

followed by normalization of G_{img} to span the range $[0, 1]$. Thus, larger values in G_{img} denote regions of high lightness, low color saturation, low luminance contrast, and low color distance, which we estimate to be the most likely candidates of glare regions. Note that the power and multiplication operators in Equation 14 are point-by-point operators.

The second column of Figure 5 shows G_{img} for the same example input images shown previously in Figure 4. The lightness and color-distance maps are generally quite effective at capturing the full extent of the glare, whereas the saturation and contrast maps can serve to refine this prediction. Accordingly, in the product in Equation 14, we use a

power of 2 for the lightness and color-distance maps in order to provide these maps extra influence over the final result.

2.2 GPS-based glare occurrence map

During the daytime, when GPS information is accessible, the physical position of the sun (i.e., the solar azimuth and elevation) can be computed based on the vehicle's latitude and longitude, and the date, time, and time zone. By using this solar azimuth and elevation, and given the direction in 3D space in which the camera is facing, the position of the sun within the image frame can be computed.

Let a_{sun} and e_{sun} denote the solar azimuth and elevation angles, respectively. Let a_{cam} denote the camera's heading angle with respect to North, and let HFOV and VFOV denote the camera's horizontal and vertical fields of view. Let e_{cam} denote the slope of the camera with respect to horizontal (ground), which can be estimated based on the camera's mounting conditions and the vehicle's recent elevation readings.

The image coordinates of the sun's center, denoted by x_{sun} and y_{sun} , can be calculated via:

$$x_{sun} = 1 + \text{round}\left(\frac{W}{2} \left[1 + \frac{e_{sun} - e_{cam}}{\frac{HFOV}{2}}\right]\right) \quad (15)$$

$$y_{sun} = 1 + \text{round}\left(\frac{H}{2} \left[1 + \frac{a_{sun} - a_{cam}}{\frac{VFOV}{2}}\right]\right) \quad (16)$$

where, again, W and H denote the width and height of the image.

By using (x_{sun}, y_{sun}) , we generate a GPS-based glare occurrence map, denoted by $G_{sun} \in R^{W \times H}$, which is an otherwise blank map containing a single Gaussian function centered at (x_{sun}, y_{sun}) . G_{sun} is computed as follows:

$$G_{sun}(x, y) = e^{-\frac{(x_{sun} - x)^2 + (y_{sun} - y)^2}{2\sigma^2}} \quad (17)$$

where σ determines the size of the Gaussian, and thus the size of the sun within the image.

Note that σ cannot be chosen arbitrarily; if σ is too small, G_{sun} could miss a considerable portion of the glare region (leading to false negatives), whereas if σ is too large, G_{sun} might include unwanted non-sun regions (leading to false positives). To estimate an appropriate value for σ , we use G_{img} as follows:

- (1) First, we binarize G_{img} using a threshold of 0.975 to capture only those regions with high glare probability

as estimated by the image-based features. We denote the largest blob of this binarized image by G_{img}^{blob} .

- (2) Next, we locate the two points in G_{img}^{blob} with maximum spatial distance. We denote this distance by d_{smax} , and we set σ as $\sigma = \max(\frac{d_{smax}}{1.4}, [\frac{\min(W,H)}{16}])$.

Before application of (17), we perform the following check to ensure that the sun's center is within the ranges $[0, W - 1]$ and $[0, H - 1]$:

$$a_{sun} - a_{cam} < \frac{HF_{OV}}{2} \text{ and} \tag{18}$$

$$e_{sun} - e_{cam} < \frac{VFOV}{2} \tag{19}$$

If these conditions are met, G_{sun} is computed and then normalized to span the range $[0, 1]$. Otherwise, or if GPS information is not available, or under nighttime conditions, G_{sun} is not computed.

The third column of Figure 5 shows G_{sun} for the same example input images shown previously in Figure 4.

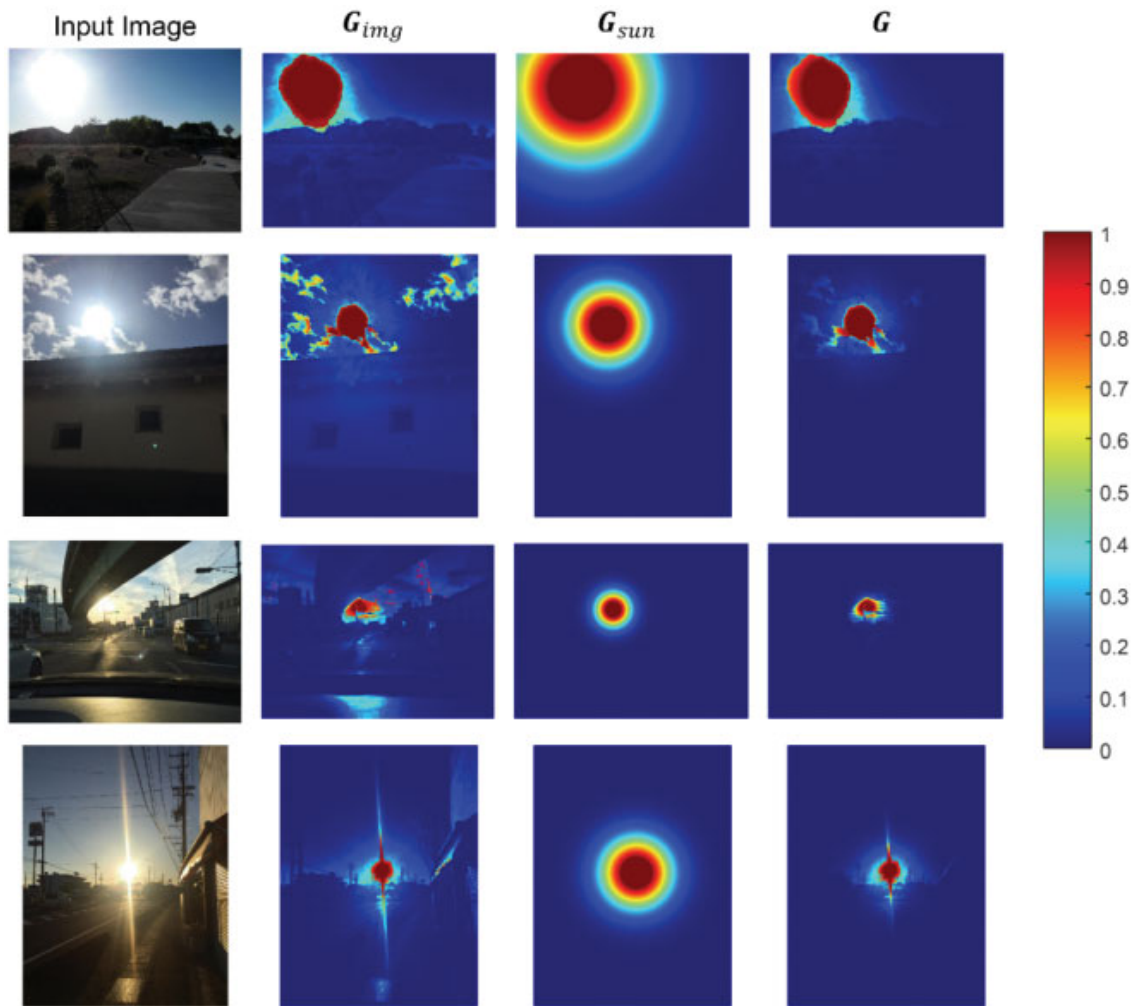


Figure 5. Four example input images and their corresponding intermediate and overall glare occurrence maps. Second column from left: Glare occurrence maps based on image features. Third column from left: Glare occurrence maps based on GPS information. Leftmost column: Overall glare occurrence map. Larger map values denote greater predicted glare

2.3 Overall glare occurrence map

Finally, we combine G_{img} and G_{sun} to generate an overall glare occurrence map, which we denote by $G \in R^{W \times H}$. There are numerous ways in which the maps can be combined, and indeed, the requirements of each particular naviga-

tion setting can dictate the proper way to use and/or combine the maps. For example, in an autonomous or remote navigation setting, one reasonable goal would be to mitigate the effects of the glare by obscuring the sun from the camera, or by selectively darkening the glare regions either digitally of

physically. In this case, the location, extent, and intensity of each glare region is needed, and thus an overall map from which this information can be extracted would be useful.

We have tested a variety of fast and simple methods of combining G_{img} and G_{sun} . We have found that the following simple product tends to yield the good overall results across a variety of scenes:

$$G = \begin{cases} G_{img} \times G_{sun}, & \text{if } G_{sun} \text{ is available} \\ G_{img}, & \text{otherwise.} \end{cases} \quad (20)$$

Note that the multiplication is a pointwise operation performed on individual pixels of the maps.

Figure 5 shows G_{img} , G_{sun} , and G for the example images shown previously in Figure 4. Observe that the combination of G_{img} and G_{sun} is generally effective at refining the result to better highlight the single strongest source of glare while reducing erroneous detections.

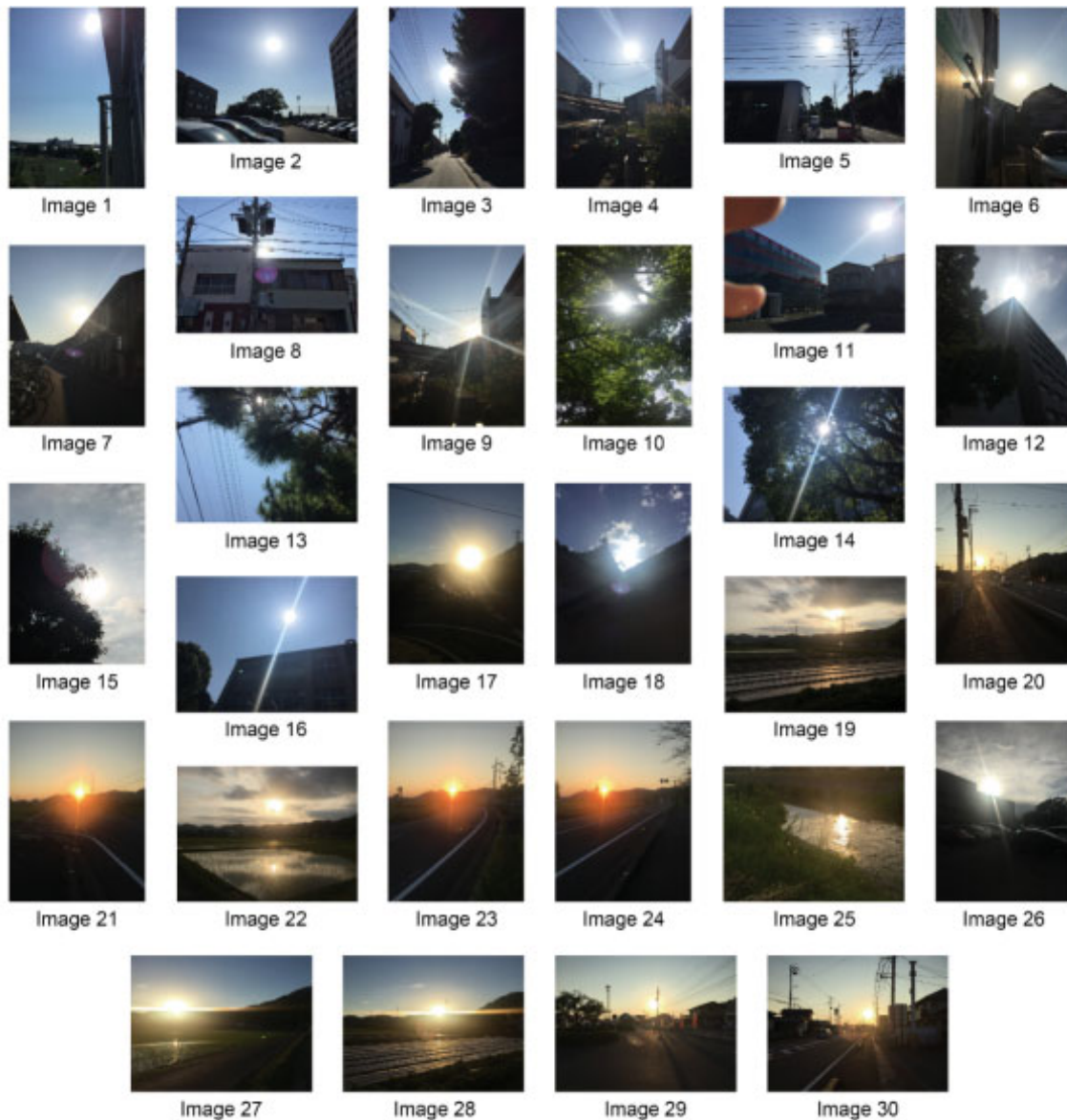


Figure 6. Thirty original images used in the glare rating experiment and contained in the database

3. GROUND-TRUTH GLARE DATABASE

Because glare is a visual sensation, it is necessary to obtain human judgments of glare as perceived when viewing the

captured image; this condition is a reasonable representation of the condition faced when using the captured images for navigation or decision-making. To this end, and to serve as a

ground-truth dataset for both our current method and for future use by other researchers, we created a database of images containing various amounts of glare along with subjective data. Specifically, for each image, human subjects labeled the center and boundary of the primary glare region, and they provided a rating of the perceived impairment caused by the glare.

3.1 Database stimuli

Thirty images were used as stimuli in the rating experiment and are contained in the database. All of the images were captured outdoors, during the daytime, in Shizuoka Prefecture, Japan, by using the rear-facing camera of an iPhone 6 (iSight) at its default exposure settings. The images were of size 3264×2448 or 2448×3264 pixels, with 24 bits of color resolution per pixel. The images were taken to expressly capture the sun and/or its reflections at various positions, colors, shapes, and extents, across a range of times (early morning until late afternoon). Figure 6 shows the 30 images.

The images were displayed on a DELL Inspiron-24 Model 3459’s monitor in a darkened room. The viewable area of the monitor was 52.7 cm wide and 29.7 cm tall, and the screen resolution was set to 1920×1080 pixels. The monitor yielded minimum and maximum luminances of 0.34 and 250.9 cd/m^2 , respectively; an overall luminance gamma of 2.58; and CIE (x, y) chromaticity coordinates of (0.636, 0.329), (0.306, 0.609), (0.149, 0.061), and (0.305, 0.337) for the R, G, B primaries and white point, respectively; these measurements were obtained via a Datacolor Spyder 5 PRO (Lawrenceville, New Jersey, USA). Subjects were instructed to use a viewing distance of approximately 60 cm.

Eight male subjects participated in the experiment: KK (the second author), YK, DI, YO, YK2, YZ, KK2, and YM. All subjects had self-reported normal or corrected-to-normal vision. The subjects ranged in age from 22 to 31 years old.

3.2 Database procedures

The images were presented to the subjects by using the GNU Image Manipulation Program;^[12] zooming and panning were allowed, but other adjustments were prohibited. The subjects were instructed to provide for each image the following data:

- (1) *Center*: Subjects specified the center location of the primary source of glare by using the pencil tool to mark a dot on an separate layer in GIMP.
- (2) *Boundary*: On a separate layer in GIMP, subjects used the pencil tool to trace the boundary of what they considered to be the primary glare region; subjects were also instructed to close the boundary.
- (3) *Impairment*: A visual impairment rating (VIR) on a scale from 0-10, where 0 denotes no impairment in

viewing the scene, and 10 denotes complete impairment in viewing the scene.

The 30 images were divided into two sets of 15 images; each subject viewed and marked one set of 15 images. Although no time limit was imposed for the experiment, most subjects finished each set within 60 minutes.

3.3 Database results

The results for the images in the database are provided in Figure 7 and Table 1. Figure 7 shows the average perceived glare occurrence map (computed from the boundary data) for each corresponding original image shown previously in Figure 7. The corresponding average center locations (computed from the center data) and average VIR data are listed in Table 1.

Table 1. Average center locations and average visual impairment ratings for the 30 images contained in the glare database; also shows are standard deviations (SDs) of the means.

Image	Avg. x	SD x	Avg. y	SD y	Avg. VIR	SD VIR
1	1514.5	4.36	300.3	4.03	3.3	1.09
2	1748.5	4.51	680.3	4.79	4.6	1.19
3	1088.5	22.52	1220.8	4.50	4.6	1.08
4	1357.5	2.65	784.8	4.19	4.1	1.14
5	1827.3	4.57	678.0	2.31	5.1	1.43
6	1521.0	2.00	1323.0	2.94	3.4	1.92
7	1295.5	6.61	1268.5	13.72	7.0	0.71
8	1575.5	12.58	992.0	9.06	3.9	1.14
9	1529.5	6.66	1559.8	4.57	5.9	0.54
10	1176.8	13.74	1051.5	3.42	5.4	1.08
11	2840.3	9.54	438.8	6.70	5.5	1.12
12	1409.0	4.69	810.0	6.06	4.4	1.56
13	1503.5	1.00	249.5	6.45	2.6	0.96
14	1818.8	4.11	769.0	2.58	5.3	1.09
15	1556.3	5.44	1889.8	6.13	2.9	1.43
16	2051.0	1.83	732.3	2.22	4.9	1.02
17	1479.0	3.83	1348.8	10.05	8.4	1.08
18	1310.3	6.18	1276.8	6.18	7.4	0.65
19	1967.3	8.77	681.5	5.92	5.1	0.54
20	770.0	1.83	1411.8	2.87	2.9	2.07
21	1268.0	3.37	1308.3	6.65	6.5	0.79
22	1764.8	4.99	714.0	5.35	4.4	1.78
23	1169.0	6.06	1254.8	4.99	7.9	1.52
24	887.3	2.87	1264.3	3.50	6.3	0.43
25	2139.8	4.79	1016.8	48.43	4.5	1.50
26	991.0	1.15	1201.0	9.59	8.8	1.09
27	811.8	9.00	946.5	15.42	7.3	1.03
28	1726.5	5.20	1011.3	9.91	7.9	1.14
29	1784.3	10.11	867.8	2.36	6.8	1.09
30	1820.5	4.80	1047.3	4.50	6.8	1.09

Specifically, from the subjective boundary data from all subjects, we computed for each image a so-called average perceived glare occurrence map. This map was generated by first generating a binary map from each subject’s closed

boundary, in which the center of the boundary was set to a value of 1, and outside of the boundary was set to zero. Next, the per-subject binary maps for each image were averaged to generate the glare occurrence map. A value of 1 indicates that all subjects considered the corresponding location to contain glare, whereas a value of 0 indicates that no subject considered the corresponding location to contain glare. As

can be seen from the results, this judgment varied between subjects; in terms of the size of the boundary, the average correlation coefficient between any given subject and the average of all subjects was approximately 0.7. However, for all images there was always a core area in which all subjects agreed (denoted by a value of 1 in the resulting maps).

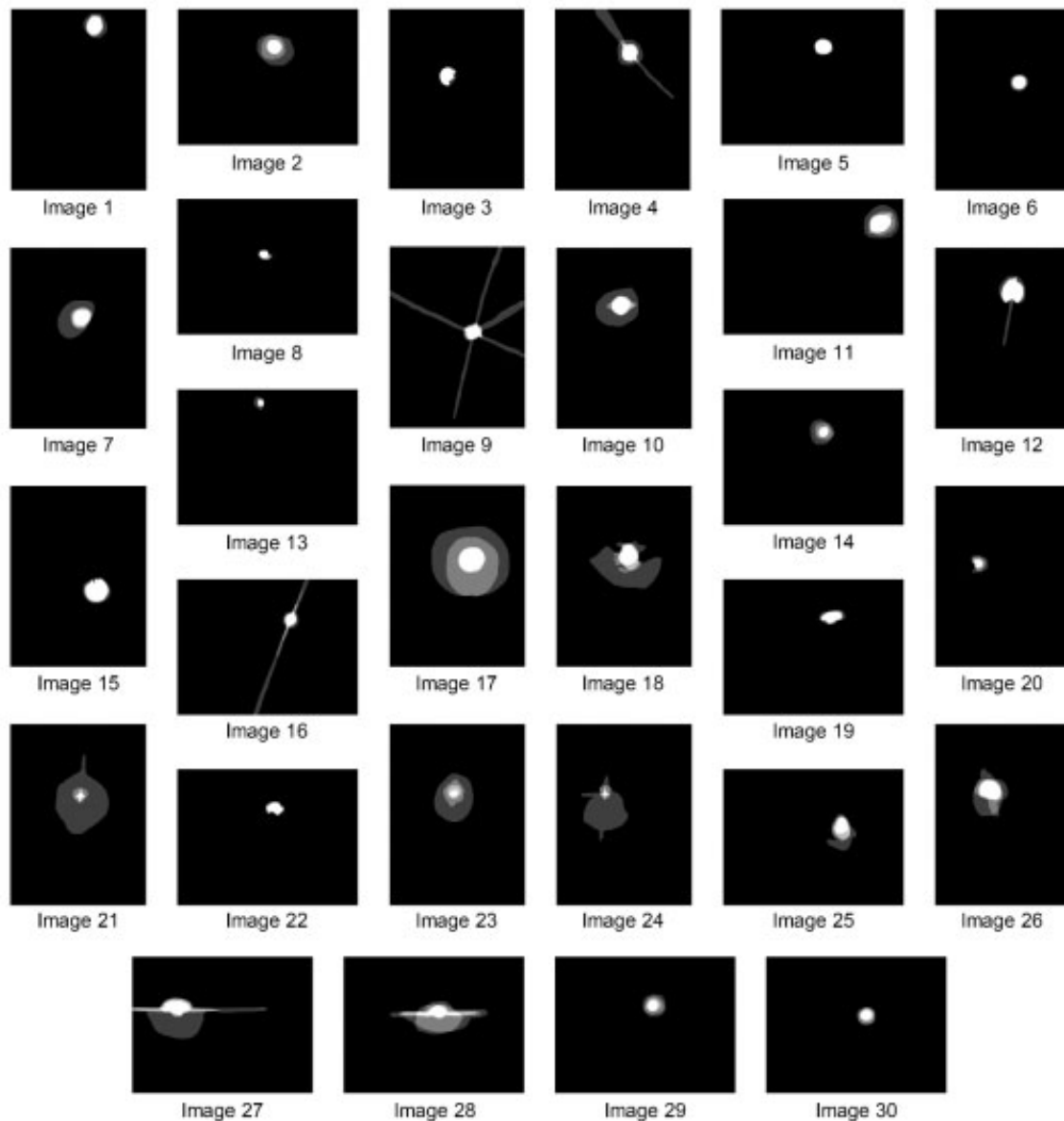


Figure 7. Subjective glare occurrence maps obtained for the 30 images contained in the glare database

From the center data from all subjects, we computed for each image the average x-coordinate and average y-coordinate, thus yielding a ground-truth center location for each image (see Table 1). Overall, the subjects were in high agreement with each other in regards to the center locations; the correlation coefficient between any given subject and the average of

all subjects was consistently greater than 0.9.

From the VIR data, for each image, we computed the average VIR and the standard deviation of this average across subjects (also shown in Table 1). The subjects generally agreed with each other; the largest standard deviation (of approximately 20%) occurred for image #20; the lowest standard

deviation (4%) occurred for image #24. The average correlation coefficient between any given subject and the average of all subjects was 0.8.

4. RESULTS AND ANALYSIS

In this section, we present and discuss the results of the proposed algorithm in terms of its ability to estimate the glare occurrence maps in the database described in the previous section. (We do not test the prediction of the center locations nor the VIR ratings; although it may certainly be possible to predict these latter data, the proposed algorithm was not designed for this purpose.) For the results described here,

the original 3264×2448 -pixel or 2448×3264 -pixel input images were resized via bicubic resizing with antialiasing to 816×612 pixels or 612×816 pixels, respectively.

4.1 Qualitative results

Figure 8 shows the input images, the average ground-truth markings, the intermediate maps (G_{img} and G_{sun}), and the overall glare occurrence map (G) for the images in the database. As demonstrated by these results, the proposed algorithm is generally effective at capturing the center and extent of the glare.

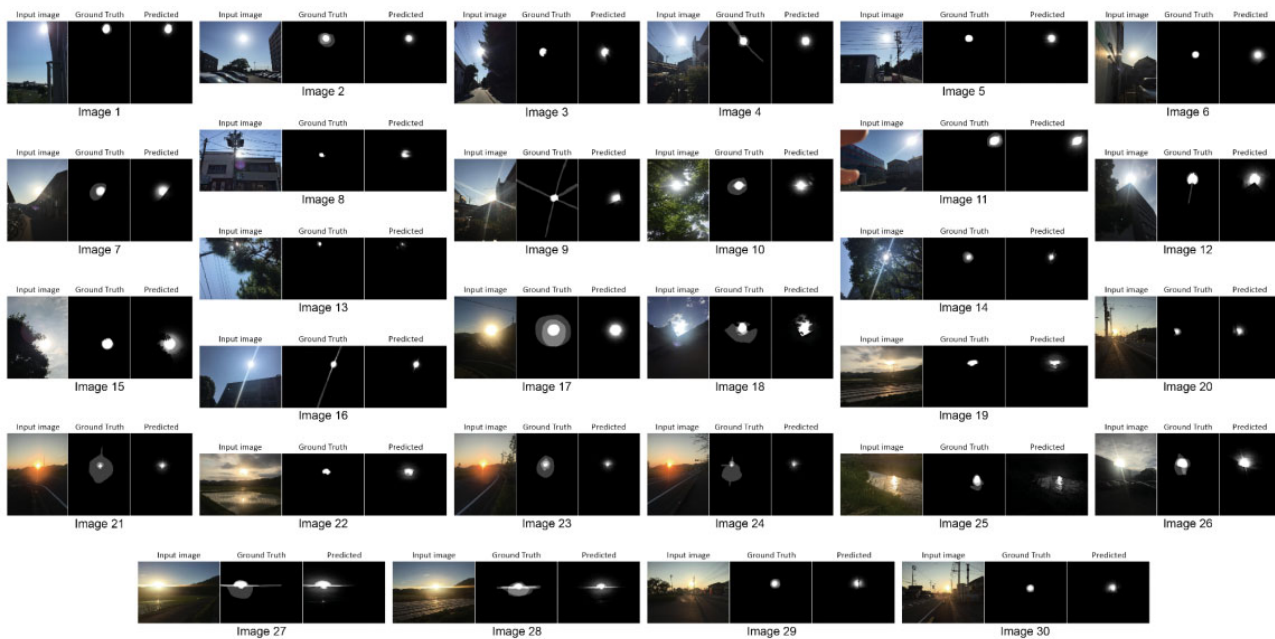


Figure 8. Original images, subjective glare occurrence maps, and predicted glare occurrence maps for the 30 images contained in the glare database

Regarding detection of glare region boundaries, predictions by our algorithm match the ground-truth data the best for images #1, #3, #5, #7, #8, #10, #11, #15, #19, #20, #22, #27, #28, #29, and #30. For images #3, #5, #7, #8, #15, #19, #20, #22, #27, #29, and #30, the algorithm can even capture the soft boundary around the primary glare region. However, for images #2, #9, #12, #16, #17, #18, #21, #23, and #24, the algorithm cannot find the entire extent of the glare region. A common property of these latter images is the existence either (1) long sun rays; or (2) large halo regions around the sun with an orange/red color; these features are suppressed in our algorithm due to the use of a Gaussian kernel of limited size for G_{sun} (suppresses long sun rays) and the use of an experimentally-determined color for the sun which is closer to yellow than red (suppresses the halo). Also, note that for images #27 and #28, the algorithm cannot fully capture both

the halo regions and the long rays at the same time, although it is able to achieve a reasonable balance.

It worth mentioning that sun rays do not occupy large areas of the camera field of view and thus will not be as impairing as the center region of the glare. Since the halo regions have smaller light intensities, they are not as dazzling as the center glare region. Our algorithm detects these center regions in most of the images (as the Gaussian kernel intensity also drops as one moves away from the center) but might not detect the entire halo region. Therefore, using a simple Gaussian kernel for G_{sun} might be preferred over more advanced kernels (e.g. fan filters that might also detect rays) due to its smaller computational burden.

4.2 Overall prediction performance

Table 2 lists the quantitative results of the proposed algorithm on each of the 30 images in the database. We evaluated the ability of the algorithm to predict the ground-truth glare maps in terms of linear correlation coefficient (CC), Kullback Leibler divergence (KLD), and root mean-squared error (RMSE) between the ground-truth map and the predicted

map. We also evaluated the ability of the algorithm to act as a binary classifier to determine the presence or absence of glare at each location) by using Precision, Recall, and F_{α} . For the binarization needed in the latter evaluation, we used thresholds of 0.5 and 0.75 (with maps ranging from 0 to 1), corresponding to at least 50% and at least 75% of the subjects agreeing that a particular location contained glare.

Table 2. Linear correlation coefficient (CC), Kullback Leibler divergence (KLD), root mean-squared error (RMSE), precision, recall, and F-value between the ground-truth and predicted maps

Image	CC	KLD	RMSE	T = 0.5			T = 0.75		
				Prec.	Recall	F-val	Prec.	Recall	F-val
1	0.977	0.249	0.022	0.981	0.962	0.975	0.988	0.972	0.983
2	0.929	0.250	0.041	1.000	0.406	0.672	0.990	0.835	0.932
3	0.900	0.614	0.043	0.812	0.979	0.861	0.927	0.970	0.941
4	0.898	1.662	0.049	0.839	0.917	0.863	0.899	0.989	0.927
5	0.915	0.558	0.039	0.926	0.981	0.944	0.928	0.949	0.935
6	0.802	0.831	0.063	0.515	1.000	0.614	0.784	0.972	0.838
7	0.956	0.228	0.034	0.806	0.973	0.855	0.946	0.926	0.939
8	0.752	0.901	0.050	0.431	0.999	0.532	0.571	0.991	0.665
9	0.748	6.451	0.074	0.748	0.983	0.812	0.895	0.961	0.916
10	0.934	0.292	0.047	0.715	0.988	0.788	0.916	0.905	0.912
11	0.956	0.345	0.039	0.959	0.786	0.893	0.979	0.931	0.962
12	0.862	0.842	0.077	0.689	0.973	0.763	0.919	0.991	0.942
13	0.770	0.960	0.026	0.961	0.414	0.667	1.000	0.265	0.519
14	0.918	0.336	0.030	0.983	0.446	0.701	0.944	0.805	0.893
15	0.783	0.912	0.106	0.476	0.987	0.575	0.800	0.970	0.849
16	0.861	6.160	0.043	0.939	0.712	0.849	0.970	0.834	0.920
17	0.843	0.604	0.118	1.000	0.245	0.493	0.996	0.886	0.957
18	0.828	0.787	0.085	0.757	0.932	0.807	0.666	0.990	0.747
19	0.818	0.986	0.056	0.731	0.807	0.755	0.930	0.844	0.899
20	0.899	0.531	0.033	0.862	0.942	0.887	0.635	0.978	0.719
21	0.714	1.723	0.066	0.991	0.673	0.857	0.774	0.944	0.823
22	0.815	0.997	0.057	0.680	0.980	0.757	0.719	0.987	0.791
23	0.849	0.631	0.057	1.000	0.269	0.524	0.999	0.305	0.568
24	0.647	3.122	0.057	0.941	0.646	0.817	0.734	0.696	0.721
25	0.717	1.733	0.087	0.841	0.591	0.737	0.911	0.568	0.758
26	0.877	0.593	0.067	0.771	0.736	0.759	0.897	0.922	0.905
27	0.899	0.599	0.070	0.914	0.941	0.923	0.845	0.966	0.882
28	0.870	0.490	0.074	1.000	0.255	0.507	0.968	0.608	0.808
29	0.926	0.367	0.031	0.976	0.713	0.869	0.938	0.753	0.867
30	0.900	0.508	0.034	0.887	0.722	0.824	0.881	0.821	0.860
Average	0.852	1.175	0.056	0.838	0.765	0.763	0.878	0.851	0.846
Std. Dev.	0.082	1.517	0.023	0.161	0.251	0.135	0.116	0.191	0.116

As shown in Table 2, the algorithm generally demonstrates good predictive performance in terms of CC, KLD, and RMSE. The best correlation coefficients ($CC > 0.92$) occurred for images #1, #2, #11, #7, #10, and #29. Images #1, #2, #11, #7, and #10 also yielded the best KLDs ($KLD < 0.3$).

Images #1, #13, #7, #14, and #29, #20, #30 yielded the best (lowest) RMSEs ($RMSE < 0.035$). The worst result in terms of CC was obtained for image #24, due to the inability of the algorithm to capture the full spatial extent of the glare cast below the sun indicated by one of the subjects. The worst

result in terms of KLD was obtained for image #9 due to the algorithm not being able to fully capture the long rays. In terms of RMSE, the worst result was obtained for image #17, due to the inability of algorithm to fully capture the relatively large diameter of the glare indicated by two subjects.

In terms of the ability of the algorithm to determine the presence or absence of glare, the performance is also generally quite good. On average, F_{α} is greater than 0.75 for detecting the presence/absence of glare using a binarization threshold of 0.5; and F_{α} is greater than 0.84 for detecting the presence/absence of glare using a binarization threshold of 0.75. For the $T = 0.5$ condition, the best results occurred for images #1, #5, #11, #20, and #27. The worst results occurred for images #17 and #28 (the algorithm does not fully capture the extent of the glare surrounding the sun). For the $T = 0.75$ condition, the results are satisfactory for most of the images with the exception of underestimates of the glare extents in images #13 and #23.

4.3 Contributions of the individual features

To examine the contributions of each of the individual feature maps used in the algorithm to the overall prediction performance, we generated predictions by using only one or two of the feature maps at a time. Specifically, we generated predictions by using only (G_{img} , only G_{sun} , only (G_{sun} and V , only (G_{sun} and S , only (G_{sun} and C , and only (G_{sun} and D). Because the final maps predicted by using only selective feature maps generally predicts broader glare areas than the ground-truth maps, values less than 0.25 were set to zero, and the remaining values were rescaled to span the range [0, 1]; this scheme generally allowed the individual predictions to perform much better than by using only the raw predictions. For the performance measure, we used correlation coefficient. Evaluations are shown in Table 3.

The first column of Table 3 shows the CC values for the full algorithm (repeated from Table 2). The second and third column show the CC values obtained using only G_{img} (i.e., no G_{sun}) and only G_{sun} (i.e., no G_{img}), respectively. Observe that although there certainly are cases in which only G_{img} or G_{sun} can perform equal or even better than the combination used in the proposed algorithm, always using just G_{img} or always using just G_{sun} does not yield consistent results across the images. For these images, the combination yields a higher CC approximately 60% of the time.

The fourth through seventh columns of Table 3 show the CC values obtained using a product of G_{sun} and only one of the constituent features maps from G_{img} (V or S or C or D). Again, although there certainly are cases in which G_{sun} and an individual map can perform equal or even better than

the combination used in the proposed algorithm, much more consistent performance across all of the images is obtained by using the full combination. Comparing the average CC values for the individual constituent maps, V and D appear to have a greater contribution to the overall performance as compared to S and C . The proposed combination provides a reasonable balance between leveraging the strengths of the individual maps and compensating for the weaknesses of sometimes-erroneous detections on certain images/regions.

Table 3. Linear correlation coefficient (CC) performance for individual features

Image	All	G_{img}	G_{sun}	G_{sun} and only the map listed below			
				V	S	C	D
1	0.977	0.947	0.762	0.895	0.802	0.780	0.982
2	0.929	0.842	0.890	0.896	0.893	0.891	0.842
3	0.900	0.905	0.662	0.856	0.689	0.761	0.830
4	0.898	0.911	0.792	0.859	0.813	0.801	0.916
5	0.915	0.937	0.777	0.851	0.799	0.808	0.957
6	0.802	0.599	0.780	0.835	0.809	0.788	0.834
7	0.956	0.826	0.843	0.939	0.854	0.876	0.878
8	0.752	0.743	0.595	0.721	0.609	0.668	0.715
9	0.748	0.670	0.637	0.755	0.654	0.672	0.712
10	0.934	0.780	0.858	0.919	0.866	0.894	0.881
11	0.957	0.970	0.771	0.897	0.802	0.777	0.979
12	0.862	0.864	0.702	0.831	0.721	0.719	0.893
13	0.770	0.544	0.522	0.761	0.522	0.522	0.521
14	0.918	0.782	0.936	0.911	0.949	0.962	0.920
15	0.783	0.573	0.699	0.796	0.714	0.746	0.780
16	0.862	0.901	0.682	0.765	0.718	0.685	0.850
17	0.843	0.782	0.859	0.789	0.844	0.856	0.796
18	0.828	0.796	0.741	0.778	0.764	0.769	0.800
19	0.820	0.567	0.624	0.765	0.698	0.646	0.793
20	0.899	0.861	0.754	0.842	0.839	0.809	0.864
21	0.714	0.654	0.629	0.621	0.633	0.639	0.633
22	0.815	0.635	0.623	0.723	0.723	0.641	0.819
23	0.849	0.713	0.851	0.792	0.797	0.855	0.823
24	0.647	0.556	0.620	0.598	0.631	0.630	0.620
25	0.717	0.794	no sun	0.729	-0.011	-0.044	0.158
26	0.877	0.810	0.827	0.871	0.844	0.833	0.882
27	0.899	0.894	0.386	0.675	0.398	0.409	0.865
28	0.870	0.752	0.654	0.750	0.653	0.664	0.895
29	0.926	0.734	0.949	0.950	0.970	0.950	0.958
30	0.900	0.831	0.849	0.886	0.907	0.862	0.912
Average	0.852	0.772	0.734	0.809	0.730	0.729	0.810
Std.Dev.	0.082	0.125	0.128	0.090	0.188	0.191	0.163

4.4 Relative computation times

Finally, we briefly analyzed the relative run-times required for computing the various features used in the algorithm. Although our current implementation of the algorithm consists only of an unoptimized Matlab version, it is still informative to compare the relative amounts of time required for each major computation of the algorithm. This timing analysis was performed on a desktop PC with the following specifica-

tions: CPU: Xeon X5670 @ 2.9 GHz; RAM: 12 GB DDR3; Software: Matlab 2016b (Version 9.1). The code was not explicitly written to use multiple cores. The timing results are shown in Table 4; these results represent the average run-times (and the standard deviations) for the first 10 images in the database.

Table 4. Run-times for the various feature computations in the proposed algorithm

	Computation	Avg. Time (sec)	Avg. Time (% of total)	SD
G_{img}	RGB to $L^*a^*b^*$ (including Lightness map)	0.218	40.6%	0.007
	Saturation map	0.004	0.7%	0.001
	Contrast map	0.198	36.9%	0.009
	Color difference map	0.007	1.2%	0.001
	G_{sun}	0.098	18.3%	0.005
	G (total time)	0.536	100%	0.026

As can be seen from the data in Table 4, the two most time-consuming computations are the color-space conversion from RGB to $L^*a^*b^*$, and the computation of the contrast map. The exponentiations (power operators) in Equations (1) and (8) account for much of the time required for the color-space conversion; if these can be approximated or possibly implemented differently, significant time-savings could likely be realized. The computation of the contrast map is also relatively time-consuming due both to the exponentiation and the block-based local contrast computation; the latter was implemented via two nested loops in Matlab, which is notoriously inefficient. Although the current total time of this Matlab implementation is approximately 0.5 seconds per image, we believe that significant speedups could be realized

by using parallel computations and a different programming language.

5. CONCLUSIONS

In this paper, we have presented an algorithm for glare detection and a database for its evaluation. The proposed algorithm uses a combination of photometric, colorimetric, and GPS features to detect the location and spatial extent of glare within captured images. Specifically, feature maps using lightness, saturation, contrast, and color distance are computed, combined, and then refined based on the sun's predicted location from the GPS information. Testing on the database revealed that the proposed algorithm can reliably detect the locations and spatial extents of glare sources in a variety of images based on subjective ratings and well-known quantitative measures. Furthermore, our testing revealed that no single feature alone could provide consistent detection performance across all of the images in database; the proposed combination of all of the feature maps, however, was able to achieve a relatively consistent performance.

6. FUTURE WORK

We have considered the following goals for our future work: The first goal is real-time implementation of our algorithm in driving vehicles in order to effectively darken the windshield and suppress the glare source without entirely occluding the driver's field of view. The second goal is to create a large database for glare images from a variety of real-life situations that can be used by other researchers and to train a deep neural network for locating the glare source in the captured frame and then to compare the accuracy and computational speed of deep learning vs. our photometric and GPS features.

REFERENCES

- [1] Devassy A, Gopinath NB, Ramachandran A. Coordinated, progressive vehicular headlight glare reduction for driver safety using wireless sensor networks. In International Conference on Connected Vehicles and Expo, ICCVE 2014, Vienna, Austria. November 3-7 2014, pp. 14-15. IEEE. <https://doi.org/10.1109/ICCV.2014.7297531>
- [2] Friedland H, Snyckerski S, Palmer EM, et al. The effectiveness of glare-reducing glasses on simulated nighttime driving performance in younger and older adults. *Cognition, Technology, and Work*. 2017; 19: 571-586. <https://doi.org/10.1007/s10111-017-0442-2>
- [3] Pharadornpanitchakul S, Chairsrichaoren R. Danger detection from head light glare through vision estimation. In Communications and Information Technologies (ISCIT), 2013 13th International Symposium on, pp. 661-664. IEEE. <https://doi.org/10.1109/ISCIT.2013.6645944>
- [4] Shehata M, Pervez M, Cai J, et al. Real time static glare identification in ITS. In Practical Real World Technologies for Communications and Embedded Platforms. 2006.
- [5] Cho CS, Song J, Park JI. Glare region detection in night scene using multi-layering. In The Third International Conference on Digital Information Processing and Communications. 2013. pp. 467-469. The Society of Digital Information and Wireless Communication.
- [6] Kumon H, Haseyama M. Glare detection for night wet road surfaces and driver visibility improvement by using multiple onboard cameras. *The Journal of the Institute of Television Engineers of Japan*. 2013; 67(3): J95-J103. <https://doi.org/10.3169/itej.67.J95>
- [7] Liu X, Leu A, Bacara D, et al. Robust dazzling detection in a novel visual based dazzling avoidance system. In 2014 IEEE Intelligent Vehicles Symposium Proceedings, pp. 833-838. IEEE.
- [8] Aiteanu D, Zeisler J, Hainfellner M. Selective darkening of the windshield using high-speed video analysis for driving assistance. In Optimization of Electrical and Electronic Equipment (OPTIM), 2012

- 13th International Conference on, pp. 1465–1470. IEEE. <https://doi.org/10.3169/itej.67.J95>
- [9] Andalibi M, Chandler DM. Automatic glare detection via photometric, geometric, and global positioning information. *Electronic imaging*. 2017: 77–82. <https://doi.org/10.2352/ISSN.2470-1173.2017.19.AVM-024>
- [10] Madan N, Geetha K. The effectiveness of glare-reducing glasses on simulated nighttime driving performance in younger and older adults. *Computational Vision and Bio Inspired Computing*. 2018; 28: 883–891.
- [11] Moulden B, Kingdom FAA, Gatley LF. The standard deviation of luminance as a metric for contrast in random-dot images. *Perception*. 1990; 19: 79–101. PMID:2336338. <https://doi.org/10.1068/p190079>
- [12] The GIMP Development Team (2019). GIMP, Version 2.10.12, Available from: <https://www.gimp.org>, Accessed 2019-06-12
- [13] Jatzkowski I, Wilke D, Maurer M. A Deep-Learning Approach for the Detection of Overexposure in Automotive Camera Images. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2030-2035). IEEE. 2018 November. <https://doi.org/10.1109/ITSC.2018.8569692>