A regression based approach of exlporing the glare metrics using real world experiments data

Hongyu Gu, Wenqing Qu, Maximilian van Amerongen
CS-433 Machine Learning, EPFL Lausanne, Switzerland

***Abstract*—Discomfort glare metrics are normally defined by experimental data. With the development of the times and changes in the structure of interior design, a single standard does not meet the needs of all scenarios. Take Daylight Glare Probability (DGP) as an example. It is composed of two parts, one that accounts for saturation and one that accounts for contrast . Although DGP works relatively well in most scenarios, there is still room for improvement in glare prediction for scenarios outside of the range of brightly lit scenarios for which it was originally developed. Based on the experiment organized by LIPID lab at EPFL , new glare metrics are explored in this paper by making use of machine learning techniques.**

I. Introduction

In recent years, as more and more people working indoors, daylight in buildings has become a part of people's daily life and that cannot be ignored. It is not only an environmental parameter, but it is also closely related to people's health and well-being.[1], [2] However, poor daylight conditions can cause a lot of problems for people such as visual discomfort and over-heating. People's visual feeling is influenced by many factors. One way to ensure visual comfort is to reduce occurrences of discomfort glare through good daylight design. Together with physically accurate daylight simulations, discomfort glare prediction metrics can help designer to make decisions in early-design stages. The most important effects, saturation and contrast effect, are always included in existing metrics. However, some existing metrics are no longer applicable, as the design of interior structures changes. For example metrics like Daylight Glare Probability (DGP) were developed from user studies in bright daylit conditions. One disadvantage is that they rely heavily on the vertical illuminance (Ev) term. However, DGP was found to underpredict visual discomfort in post-occupancy evaluations of deep open plan offices for instance, which are increasingly found in cities and usually come with lower vertical illuminances overall. Thus it has lower adaptation levels when exposed to glare sources [3], [4]. To extend discomfort glare metrics to a wider range of scenarios, LIPID lab at EPFL designed and conducted user studies that mimicked such low-light environments.

In this paper, a performance optimization of glare metrics is proposed, so that it is applicable for scenarios outside of the range of brightly lit scenarios. Machine learning techniques will be applied for this purpose. With the experiment data from LIPID lab, new contrast-ratio-based metrics are explored based on regression based machine learning models. More specifically, it is focused on the metric $loggc$, which is a contrast ratio part of Daylight Glare Probability (DGP). A method will be proposed, to come up with a top ten ranking of exponents combinations for the $loggc$. The structure of the $loggc$ will be kept, however the weighting of each parameter will be changed.

II. Background

Hopkinson claimed in 1966, there are two main effects of glare:" One is a contrast effect, which results when a light source, possibly only of moderate brightness, is seen in an environment of much lower brightness and so causes glare by contrast; and the other is a saturation effect, which results when any part of the retina, even the whole retina, is stimulated by light at such a level that the maximum possible rate of neural response from the retinal elements is generated. A snow- covered landscape illuminated by full sun light is completely devoid of contrast, but most people experience acute discomfort due to the saturation of the whole visual response mechanism" [5]. In the early stages of development, glare metrics such as CIE Glare Index (CGI), Daylight Glare Index (DGI), Unified Glare Rating (UGR) were summarized from electric lighting conditions [6], [7]. One disadvantage of those metric were, they only considered the contrast effect of glare. Then, under controlled laboratory conditions, daylit office environments were reconstructed [8]. In combination with extensive user studies the contrast metric Daylight Glare Probability (DGP) was developed, which explicitly includes the vertical illuminance Ev, both as saturation term but also as contrast term. The contrast term, also referred to as $loggc$, is defined as:

$$loggc = \log_{10}(1 + \sum_{i=1}^{n} \frac{L_i^2 \omega_i}{E_v^{1.87} P_i^2}) \tag{1}$$

Recent studies showed that hybrid metrics, like Daylight Glare Probability (DGP), outperform other single-effect metrics [9]. However, contrast-driven glare metrics were found to describe discomfort glare responses better than saturation-driven ones in dim environments with lower adaptation levels [10]. Although DGP works relatively well in most scenarios, there is still room for improvement in glare prediction for scenarios outside of the range of brightly lit scenarios from which it was developed. In low-light scenarios where adaptation level of the eye is lower, it's hypothesized, that the contrast effect takes dominance rather than the saturation effect. To that end, LIPID lab at EPFL devised and carried out a user study to investigate users' discomfort glare evaluations as well as the predictive ability of glare metrics in dim daylight office settings.

Recent years, machine learning methods are more and more used in a variety of industries. Regression analysis is a fundamental concept in the field of machine learning. It falls under supervised learning, wherein the algorithm is trained with both input features and output labels. It helps in establishing a relationship among the variables by estimating how one variable affects the other.

III. Models and Methods

Data Analysis

As mentioned before, the goal of this paper is, to come up with a systematic approach to find a combination of exponents for the parameters of the contrast metric referred to as $loggc$. The magnitude of this contrast metric should increase with increasing conditional probability of whether somebody experiences discomfort due glare caused by contrast. To do so, data is required.
The used data was retrieved at the Demona facility located in the the EPFL campus. During the experiment 63 human participants were exposed to four different lighting conditions. All four lighting conditions contained a glare source. However, the size and its luminance vary between the different scenes. The scenes are labeled as follows: "1-panel-low, 1-panel-high, 2-panel-low" and "2-panel-high. From equation 1 it can be recognized that the $loggc$ is defined by the parameters Ev,L_i,ω_i and P_i. Therefore, during each experiment, those parameters were measured. More detailed information about the experimental setup can be retrieved from [10].

During this paper, the assumption is made that all participants were exposed to similar lighting scenes. To proof the assumption to be true, the coefficient of variation, for all parameters measured during the different scenes, were calculated. The coefficient of variation (Cv) is defined as the standard deviation of the data divided by its mean. Its calculated values are listed in Table I. It can be recognized that for all parameters, the coefficient of variation does not exceed the value 0.33. This can be considered as a low value and proofs the assumption that all participants were exposed to four similar lighting scenes.

	Ev_{Cv}	Ls_{Cv}	Pos_{Cv}	ω_{Cv}
1 panel low	0.271	0.164	0.138	0.147
1 panel high	0.279	0.177	0.134	0.111
2 panel low	0.300	0.184	0.214	0.150
2 panel high	0.321	0.177	0.195	0.161

TABLE I
Coefficient of variation for different measurements

During each experiment the participants where asked to answer the following question: "At the moment, how would you describe glare in your field of view?". The participants could choose from one of the following answers: "Imperceptible","Noticeable","Disturbing" and "Intolerable". For simplicity the answers "Imperceptible" and "Noticeable" were merged to one answer. The same holds for the answers "Disturbing" and "Intolerable".

In Figure 1, the histograms of the two merged answers for each lighting scene are plotted. It can be recognized that for all four scenes the majority of participants does not experience discomfort due to contrast glare. However, the number of people who felt discomfort, increases in the following order of scenes: "2-panel-low","2-panel-high","1-panel-low","1-panel-high". Furthermore, the number of people who felt discomfort increases from scene to scene by: $3, 5$ and 2 participants.

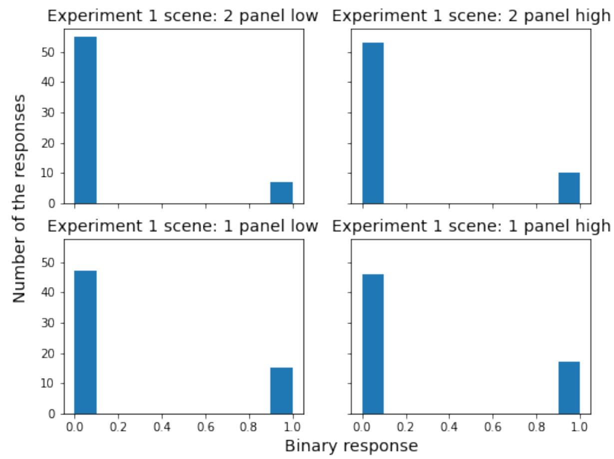

Fig. 1. Histogram of responses belonging to each scene

In Figure 2, the box plot of the nominal $loggc$ of the considered data for the four different scenes are given. It can be recognized that the nominal $loggc$ ranks the different scenes correctly, by assigning the lowest value to scene "2-panel-low" and the highest value to scene "1-panel-high". However, it can be recognized that the difference in the mean of scene "2-panel-low" and "2-panel-high" is minimal. The same holds for the differences in the mean of scene "1-panel-low" and "1-panel-high". Note, the mean of the data is represented by the orange line. Those differences should represent an increase of 2 and 3 participants who started to feel discomfort due to glare. The magnitude of the $loggc$ increases almost by 70% between scenes "2-panel-high" and "1-panel-low". This increase seems out of proportion considering that it represents an increase of 5 participants and comparing it to the previously described increases of 3 and 2 participants. This motivates to redefine the weight of each parameter in the summation of the $loggc$, such that an linear increase of the conditional probability that someone feels discomfort due to contrast glare, given a certain lighting scene, is represented by a linear increase of the magnitude of the $loggc$.

Feature selection

To achieve the previously mentioned goal of the this paper, the problem is reformulated as a feature selection problem. To do so, the retrieved data was used to generate for each experiment different $loggc$ values with varying values for the exponents of the parameters. For the parameters L_i, ω_i, E_v

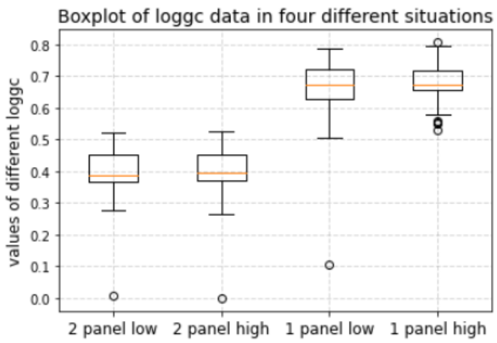

Fig. 2. Boxplot of $loggc$ data in four different situations

and P_i, 10 different exponents, linearly spaced on the interval $[0.5, 10]$ have been used to calculate different $loggc$ values. Therefore, for each experiment, a total of $10^4 = 10000$ $loggc$ values have been calculated. For the rest of the paper, it will be referred to them as the "features". After having generated the different features, a feature selection technique is required to select the best performing one. A common approach is, to select features based on fitting machine learning models to them and evaluate how well the model fits the features.

As mentioned before, the $loggc$ should represent a linear mapping between the features and the probability of whether someone experiences glare due to discomfort, given a certain lighting scene and accordingly a certain value for $loggc$. This conditional probability will represent our observation y and will be denoted as $p(\text{discomfort}|X)$.

Since machine learning models will be used for feature selection and a linear relation between the features X and the observations y is desired, linear regression models are used to evaluate the generated features.

Linear regression

Linear regression aims to find a weight vector β that maps the features X linearly to the observations y by minimizing the least squares error $\epsilon^T\epsilon$. This can be written as:

$$\min_{\beta} \epsilon^T \epsilon \quad y = X\beta + \epsilon \tag{2}$$

Four different lighting scenes are considered in the experimental setup. It was shown, that during the experiment for each participant the separate lighting scenes were similar. Therefore, each lighting scene can be characterized by the mean of the considered $loggc$ based on all experimental participant data. Furthermore, each lighting scene comes with a conditional probability $p(\text{discomfort}|X)$. For the four scenes the conditional probabilities have been calculated to be equal to $p(\text{discomfort}|X) = [0.113, 0.159, 0.2419, 0.269]$. Having defined the observations $y = (\text{discomfort}|X)$, the feature selection procedure will be as followed. The feature matrix, containing the 10^4 generated features, will be split into the four different scenes. Afterwards, it will be iterated through the total number of features. During each iteration one column will be extracted from the split feature matrix. This column represents the calculated $loggc$ values for all participants for one combination of parameter exponents. Afterwards, for each scene the mean of the extracted $loggc$ values is calculated and it is stored in the feature vector X. In the last step, a linear regression model is trained, and it is evaluated how well the features fit the trained model.

This motivates to define metrics that indicate how well the features fit the trained model.

Feature evaluation

The goal is, to find features that are linearly spaced and match the linearly spaced observations y the best. To be able to compare the different features, during each iteration the feature vector X has been normalized by the value of the mean of the scene $1 - panel - high$. Therefore, during each iteration, the last value of the feature vector X was equal to one. Furthermore, the desired slope of the model is approximately equal to: $\frac{p(\text{discomfort}|\text{1-panel-high})}{X(\text{1-panel-high})}$. Therefore, the difference between the slope of the model and the desired slope will be used as an evaluation metric.
Furthermore, a common metric to evaluate regression models is equal to mean square error (MSE). It represents the total sum of the difference between the predicted observations $\hat{y}_i$ and the actual observations y. Therefore, the MSE is defined as:

$$MSE = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{3}$$

Previously, it was shown that during the experiments, the four different lighting scenes were similar for each participant. Therefore, the coefficient of variation of the $loggc$ values, within each scene should be small. This motivates to use the sum of the coefficients of variations of the $loggc$ values within each scene as an evaluation metric.

A. Feature selection

After defining the different feature evaluation metrices, feature selection can be performed. In the considered case this was done as follows: First the top 1000 features were selected based on the difference between the slope of the model and the desired slope. From those, the top 200 were selected with the minimal MSE. Finally, the top 10 features with the minimal sum of coefficients of variations were extracted.

B. Results

The final top 10 combination of exponents, derived by the proposed method, are listed in Table III-B. It can be observed that many of the top 10 proposed exponents combinations, weight the luminance (Ls) and the vertical illuminance (Ev) similarly in the calculation of the $loggc$. This coincide with the made assumption that discomfort in low light conditions is caused by contrast glare, which can be described by the luminance (Ls) and the vertical illuminance (Ev). The box plot of the, in Table III-B listed first exponents combination, is shown in Figure 3.Comparing it to the box plot of the nominal $loggc$ given by Figure 2, it can be recognized that there is a bigger difference between the means of the different scenes. It is easier to differentiate between the different scenes, since the increase in conditional probability of feeling discomfort due to glare, is better represented by the difference of the means of the data.

	Ls	Pos	Ev	Omega	MSE	sum of standard deviations
1	5	4	3	4.5	0.0490	0.644
2	4.5	4	2.5	4	0.0591	0.664
3	3.5	3.5	3	3	0.0409	0.714
4	3	3.5	1.5	2.5	0.0526	0.740
5	5	2.5	4	5	0.0591	0.799
6	5	4.5	3	4.5	0.0560	0.807
7	5	4.5	3.5	4.5	0.0244	0.817
8	1.5	3	0.5	1	0.0449	0.845
9	4.5	4	3	4	0.0302	0.851
10	2	5	0.5	1	0.0242	0.889

TABLE II
TOP 10 COMBINATIONS OF EXPONENTS

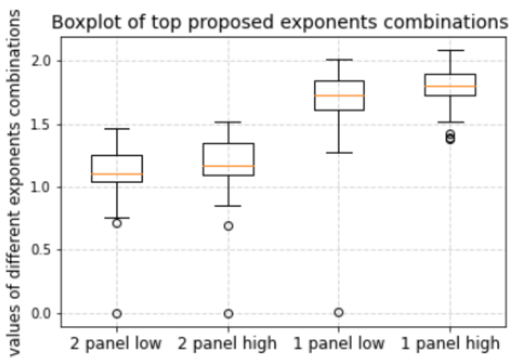

Fig. 3. Boxplot of top proposed exponents combinations exponents combination

C. Conclusion and recommendations

After generating the different new features, a method to extract top 10 features is proposed. With these new features, new contrast-ratio based glare metrics are generated. According to the previous experiment data, these new metrics can linearly rank different situations. Compared with previous $loggc$, it shows a more linear performance among the four scenes("1-panel-low", "1-panel-high", "2-panel-low" and "2-panel-high"). Although it shows linear performance, it can only be used to sort different scenarios and can not be used for predicting the actual conditional probability of whether someone is experiencing discomfort due to glare. In the course of the research, it would perform be interesting to merge $\frac{Ls}{Ev}$ to one contrast term and observe how the performance is influenced. Furthermore, it would also be interesting to replace the vertical luminance Ev with the average luminance and observe how the performance changes. For further research, more experimental data and more experimental scenarios are indispensable. In combination with new data of different scenarios, the performance of the derived exponents can be tested and evaluated. If the derived exponents do not perform well on new data, the filtering bounds can be adjusted to come up with different exponents combinations. Overall it can be concluded that further research is required.

REFERENCES

[1] M. B. Aries, M. P. Aarts, and J. van Hoof, "Daylight and health: A review of the evidence and consequences for the built environment," *Lighting Research & Technology*, vol. 47, no. 1, pp. 6–27, 2015.

[2] H. S. Jim, A. I. Hoogland, N. C. Brownstein, A. Barata, A. P. Dicker, H. Knoop, B. D. Gonzalez, R. Perkins, D. Rollison, S. M. Gilbert *et al.*, "Innovations in research and clinical care using patient-generated health data," *CA: a cancer journal for clinicians*, vol. 70, no. 3, pp. 182–199, 2020.

[3] M. Hirning, G. Isoardi, and I. Cowling, "Discomfort glare in open plan green buildings," *Energy and Buildings*, vol. 70, pp. 427–440, 2014.

[4] R. Talami and J. A. Jakubiec, "Early-design sensitivity of radiant cooled office buildings in the tropics for building performance," *Energy and Buildings*, vol. 223, p. 110177, 2020.

[5] D. Hopkinson and H. HARRIS, "Rare phosphoglucomutase phenotypes," *Annals of human genetics*, vol. 30, no. 2, pp. 167–181, 1966.

[6] H. Einhorn, "A new method for the assessment of discomfort glare," *Lighting Research & Technology*, vol. 1, no. 4, pp. 235–247, 1969.

[7] P. Chauvel, J. Collins, R. Dogniaux, and J. Longmore, "Glare from windows: current views of the problem," *Lighting Research & Technology*, vol. 14, no. 1, pp. 31–46, 1982.

[8] J. Wienold and J. Christoffersen, "Evaluation methods and development of a new glare prediction model for daylight environments with the use of ccd cameras," *Energy and buildings*, vol. 38, no. 7, pp. 743–757, 2006.

[9] K. Chamilothori, J. Wienold, and M. Andersen, "Adequacy of immersive virtual reality for the perception of daylit spaces: comparison of real and virtual environments," *Leukos*, vol. 15, no. 2-3, pp. 203–226, 2019.

[10] G. Quek, J. Wienold, and M. Andersen, "User evaluations of contrast-dominant discomfort glare in dim daylit scenarios: Preliminary findings," 09 2021.