Predicting Daylight Preferences Using HDRI and Deep Learning

Dongjun Mah – Purdue University, United States of America – dmah@purdue.edu
Athanasios Tzempelikos – Purdue University, United States of America – ttzempel@purdue.edu

Abstract

This paper utilizes High Dynamic Range Imaging and deep learning that utilize pixelwise information from the entire luminance distribution in the field of view to classify daylighting preferences of office workers. Generated luminance and contrast similarity maps were used for training convolutional neural network (CNN) models to classify the occupant's visual preferences. Preference datasets for 11 individuals, collected in real offices, were used to evaluate the preference classification performance. The results showed the superiority of the luminance similarity map as a visual preference indicator compared to common static lighting parameters.

1. Introduction

The spatial luminance distribution within the field of View (FOV) has a high correlation with the human perception of brightness and comfort (Wymelenberg et al., 2010) and is therefore valuable for assessing visual comfort and visual preferences in general. Real-time FOV luminance monitoring is achieved by acquiring per-pixel luminance using High Dynamic Range Imaging (HDRI) sensors (Inanici et al., 2006) combined with wide-angle fisheye lens, now a well-established methodology in lighting research (Pierson et al., 2021).

HDRI measurements have been used to correlate scene luminance characteristics to subjective visual comfort responses (Konstantzos & Tzempelikos, 2017; Konis, 2014; Suk et al., 2017), as well as for real-time daylighting and glare control in buildings (Newsham, 2009; Motamed et al., 2017; Kim et al., 2020). Despite these obvious advantages, the full potential towards human-centered daylighting operations has not been explored. Building systems that learn human preferences and integrate them into building operations can achieve occupant satisfaction as well as energy savings (Xiong et al., 2019; Villa et al., 2013). Detecting visual discomfort scenarios does not necessarily translate to learning or providing preferred conditions for occupants. Instead, learning lighting preferences without considering discomfort scenarios (Xiong et al., 2018) is preferred for optimized visual environment and control in buildings.

But learning and modeling human visual preferences can be extremely complex (Lindelof & Morel, 2008; Xiong et al., 2020). True visual preferences dynamically depend on different environmental, contextual, or (unmeasurable) subjective factors, outside view and aspects such as perceived control (Bakker et al., 2014) or multi-domain interactions, and have rarely been used in building control and optimization applications.

The selection of meaningful variables remains a challenge even when considering only environmental factors. Using simple variables such as horizontal or vertical illuminance, average luminance or simple contrast ratios from captured luminance maps is maybe sufficient for specific glare assessment cases, but not for inferring preferred conditions, partly due to averaging pixel information and neglecting information in different parts of the luminance field. Although HDRI sensors present information through pixelwise luminance maps, there is no agreement on which features can better represent visual preferences of occupants in typical daylighting settings.

This paper presents a novel method for inferring personal daylight preferences using image pixelwise similarity analysis applied in a deep learning framework. Instead of studying how occupants' preferences are affected by instant physical and contextual parameters using numeric scaled responses, we utilize information from the entire luminance distribution in the FOV and extract pair-wise similarity features between HDRI-based luminance maps (different conditions). We also compare models using

Part of
Pernigotto, G., Ballarini, I., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). 2025. Building simulation applications BSA 2024. bu,press.
https://doi.org/10.13124/9788860462022

common lighting variables with the CNN-based models using the new metrics.

2. Pixelwise Similarity Index and Luminance Similarity Maps

A pixelwise similarity index is introduced to quantify the degree of similarity between two luminance distributions. The index is used to generate a new variable named the luminance similarity map, which is an important part of inferring preference between two visual scenes with daylight. It compares pixels at the same location inside two luminance maps, one by one directly (pixel-wise comparison), and it considers both the direction and magnitude of relative luminance change. In contrast to grayscale images containing color data from 0 to 255, luminance maps have much greater variation in pixel intensity. The pixelwise luminance similarity index is:

$$LSim(x_1, x_2) = [sign(x_1 - x_2)] \cdot \left(1 - \frac{|x_1 - x_2|}{\max(x) - \min(x)}\right) \qquad (1)$$

where x_1 and x_2 are pixelwise luminance values in luminance maps 1 and 2 respectively (comparative luminance map data) and $\max(x)$ and $\min(x)$ compute the maximum and minimum luminance values from the entire two luminance maps for normalization. The two luminance maps represent two different scenes (two different visual conditions). To consider the directional relative luminance change in each pixel between two conditions, the sign of luminance difference is also applied to the similarity index equation. An absolute similarity of 1 means that the luminance of those pixels in the two luminance maps is the same. Lower absolute pixel similarity values indicate a difference in luminance; a zero-similarity index between two pixels means that the luminance difference at the same pixel location is the maximum possible (= max-min between the two maps). The -1 similarity value is excluded from further analysis to avoid repetition.

The luminance similarity map is generated by directly comparing two luminance maps and calculating the luminance similarity index pixel-by-pixel. A representative example of generating luminance similarity index maps is shown in Fig. 1. After RGB color data of two HDR images (two different conditions) is converted to luminance maps (Inanici, 2006), the luminance intensity of each pixel is used to calculate $LSim$ values according to Eq. (1). Then all the $LSim$ values are mapped on the camera FOV to produce the entire luminance similarity map in a graphical way that includes the relative luminance change (magnitude and direction).

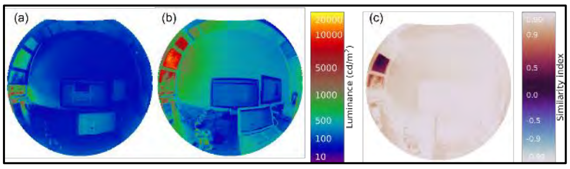

Fig. 1 – Generating a luminance similarity index map (c) from two luminance maps (a and b) corresponding to different conditions

3. A Comparative Luminance Map Dataset for Evaluating Daylight Preferences

To evaluate the ability of different models (and variables) to infer daylighting preferences, comparative luminance map datasets were created with simultaneous occupant feedback. HDR images were captured at the eye level of 11 office occupants in identical, side-by-side private offices with large windows (Xiong et al., 2019) and controllable shades, under various daylight and interior luminance conditions without glare. Calibrated Canon Rebel T2i cameras equipped with fisheye lens were used to generate reliable luminance maps. The highest DGP observed in the data set was 0.35. Additionally, conditions with vertical illuminance exceeding 2760 lux were also excluded. The daylight conditions were changed every 10 minutes by adjusting the position of window shades, and shortly after the 11 occupants were asked about their visual preference between current and previous conditions. Electric lights were automatically controlled to maintain the required work plane illuminance (300 lux).

The collected HDR images were converted to luminance maps and directly compared pixel-by-pixel to generate luminance similarity index maps. Since the HDR images were acquired under comfort conditions, a resolution of 330 x 330 was selected for this study considering the computation load required to produce the similarity maps. The luminance maps were converted to 11 comparative luminance map datasets (one for each occupant) by grouping successive luminance maps into pairwise comparative data (corresponding to current vs previous condition) and linking corresponding occupant binary visual preference data (preferring current or previous visual condition) to each pair.

Since the occupants responded with their preference between successive conditions every 10 minutes, part of the collected pairwise comparative data was considered to be the test data for assessing the classification performance of the trained models. As shown in Fig. 2, 40 % of pairwise HDR images with preference responses were used as a test dataset and used only once for evaluating the trained visual preference classification models. For the training data, the remaining 60 % of original HDR pairwise images were used and augmented by the automatically captured HDR images with 2 min-intervals, since the users have maintained their visual preference response during this time -except if they override the system in the meantime. Then, to avoid overfitting, the training dataset was randomly divided into 5 to 1 ratio to generate the validation dataset. The training dataset consists of around 500 luminance map pairs and the test dataset is about 40 pairs.

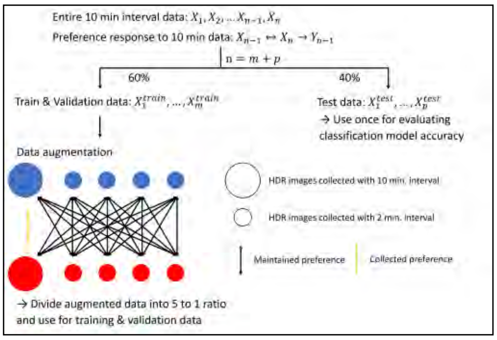

Fig. 2 – Schematic procedure for generating comparative preference test and training data

4. Performance Evaluation of Daylighting Preference Learning Models Using Common Variables

In this section we evaluate the ability of different methods and variables to infer preferred visual scenes with daylight. Using the comparative luminance map dataset, simple and advanced methods using commonly used lighting variables are compared with a deep learning model that uses the luminance similarity index maps.

4.1 Common Lighting Variables Used to Predict Daylight Preferences

To evaluate the efficiency of luminance similarity index-based metrics, a comparison is first made with common lighting parameters. The 7 "reference" parameters listed in Table 1 are selected since they have been extensively used to predict lighting preferences and comfort in daylighting settings. DGP, average luminance of entire scene, and average luminance of 40° horizontal band are computed directly by Evalglare, while the rest of the parameters can be computed by masking the window part.

Table 1 – Selected reference variables

Lighting metrics used as model variables
1. Percentage of pixels exceeding 2000 cd/m² (p_{2000})
2. Average luminance of entire scene ($avlum$)
3. Daylight glare probability (DGP)
4. Standard deviation of window luminance (std_{win})
5. Average luminance of 40 ° horizontal band ($avlum_{B40}$)
6. Maximum luminance in window divided by 200 cd/m² ($wmax200$)
7. Average luminance in window divided by 200 cd/m² ($wav200$)

4.2 Logistic Regression Model Trained with Reference Variables

To estimate the need for using more complex variables or advanced models to infer personal daylight preference, a logistic regression model is first trained using Eq. (2) with each of the reference parameters:

$$p = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2)}} \quad (2)$$

where β_i are model parameters and x_i is the selected reference variable from Table 1. Each logistic regression model was trained using each reference parameter computed from each pair of HDR images in the dataset. The classification accuracy for each person was tested by computing the ratio of correctly classified test data over the entire test data.

The classification results with each variable are shown in Fig. 3 for each of the 11 subjects (noted in the x-axis). In some cases, for example for subjects 1 and 4, the logistic regression model could classify visual preferences quite well for most variables. In other cases, the selection of variables significantly affected the prediction accuracy. Most importantly, the classification performance varies significantly between subjects. Although this is somewhat expected because of the natural preference inconsistency between individuals, sometimes this model fails to predict any preference at all (although the subject reported specific preference trends). Considering that the classification problem in this study is a binary classification problem, the trained model with about 60 % accuracy means that it is not possible to train the model with the selected reference parameter.

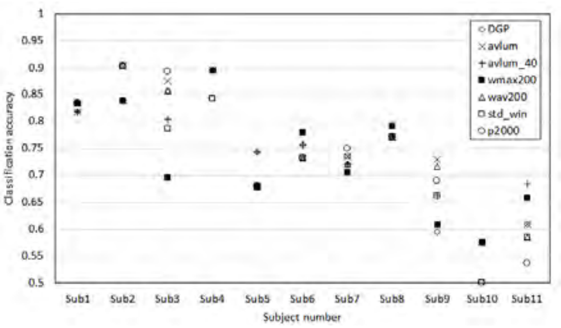

Fig. 3 – Preference classification results for each subject: logistic regression model trained with selected reference parameters

The limited and variable classification performance can be attributed to the simplicity of the logistic regression model or the insufficiency of the selected variables. To eliminate the first possibility, the performance of more advanced (deep learning-based) models is evaluated with the same reference variables.

4.3 ANN Model Trained with Reference Variables

An Artificial Neural Network (ANN) model that uses the reference variables to estimate personal visual preference was evaluated next. As shown in Fig. 4, two constant values, representing one reference parameter, were computed from current and previous luminance maps and used as input to the model. This ANN model contains three hidden layers, and each hidden layer consists of 100 neurons. After all the neurons in the hidden layers extract relationships between two input values, two output values were computed. The model classifies the person's visual preference by selecting the greatest between these two values. In addition, a second ANN model (Fig. 4) was trained using all 7 reference parameters to check if the classification accuracy could be increased. The only difference is the number of input values (2 x 7 = 14 values were input to the hidden layers in this case). The training was monitored using the cross entropy loss function:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\cdot\log(1 - p_i)] \quad (3)$$

where y_i is the binary preference label (e.g., current vs previous condition), p_i is the probability that the person prefers the current condition, and N is the data size. A stochastic gradient decent algorithm with 0.9 momentum was used as optimizer with batch size = 16. In addition, 10^{-3} was used both for learning rate and for L2 regularization strength. Then, 100 epochs were predefined for training the ANN models, monitoring validation loss to avoid overfitting. Following one of the recommended early stopping criteria (Prechelt, 2012), the training procedure was terminated when the validation loss at a specific epoch (t) was greater than the loss at previous epoch (t-5) twice.

The ANN model classification results (Fig. 4) showed better overall classification performance compared to logistic regression while the performance variation among variables was reduced. In addition, the ANN model trained with all the reference variables (marked with red markers) performed similar or better than models trained with single variables for every subject. However, the classification accuracy is still low for several subjects even with this complex model. This indicates that the static lighting variables cannot really estimate daylight preferences; they cannot include important visual information located in different areas of the visual scene and they cannot express the change in perceived luminance distribution (similar values of input parameters have different preference labels and vice versa). As discussed next, the pixelwise luminance similarity-based metrics present a clearer characterization of personal daylight preferences.

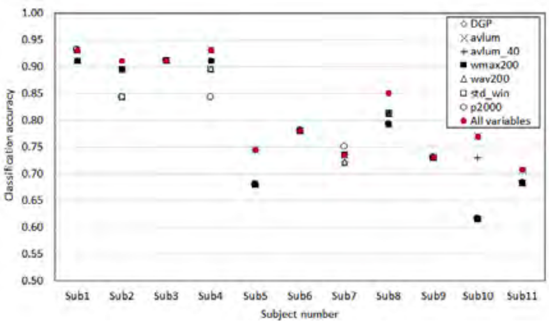

Fig. 4 – Preference classification results for each subject: ANN model trained with selected reference parameters and with all parameters together

5. Using The New Luminance Similarity Index Maps to Infer Daylighting Preferences

5.1 CNN Model Architecture Trained with Luminance Similarity Index Maps

In contrast to the reference variables, similarity index maps extracted from luminance similarity maps are in 2D array data format. A Convolutional Neural Network (CNN) model structure is therefore developed to preserve pixelwise information and patterns as much as possible. Luminance similarity index maps are both used solely and together to study their ability to classify personal visual preference. As shown in Fig. 5, the CNN model consists of 4 convolutional layers and 2 fully connected layers. Unlike typical CNN model architectures, which consist of convolutional and maxpool layers, only convolutional layers were used. This is because computing maximum values in the maxpool layers might result in erasing important similarity index patterns, which can be the combination of positive and negative similarity index values. Instead, convolutional layers with stride 2 were selected, which not only preserve necessary information but also reduce CNN model parameters considering computational efficiency. Stride 1 is only used in the first convolutional layer to capture pixelwise information in detail as much as possible. All the layers used 32 kernels with size of 3 x 3 to extract helpful features for preference classification. After the input luminance similarity index maps move through the 4 convolutional layers, the initial 330 x 330 size is reduced to 42 x 42 and input to the fully connected layers to link all the parameters with each other. By selecting the greatest between the two output values, the CNN model will classify the occupant's preference.

Similar to ANN model training, loss function was evaluated using cross entropy, a stochastic gradient decent algorithm with 0.9 momentum was used as optimizer with batch size = 16, and 10^{-3} was used both for learning rate and for L2 regularization strength. 50 epochs were set for training in this case and the same early stopping criteria were applied to prevent overfitting.

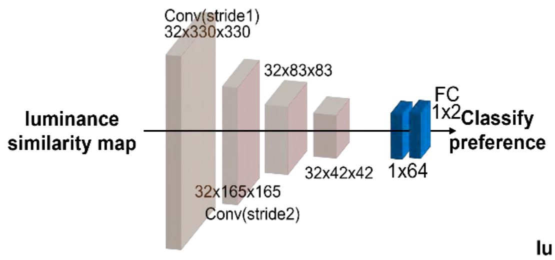

Fig. 5 – CNN model architecture, trained with luminance similarity index maps

5.2 Classification Performance of CNN Model with Luminance Similarity Index Maps

The CNN models using luminance similarity index maps show excellent classification performance compared to other models (Fig. 6). More importantly, the CNN models performed much better than the ANN especially for Subjects 5-11, where the previous models performed poorly. This proves that the luminance similarity map, which contains a great amount of pixel-wise information from each pair of daylighting conditions, is a superior preference indicator when used in powerful deep learning models, in contrast to constant variables used in previous studies.

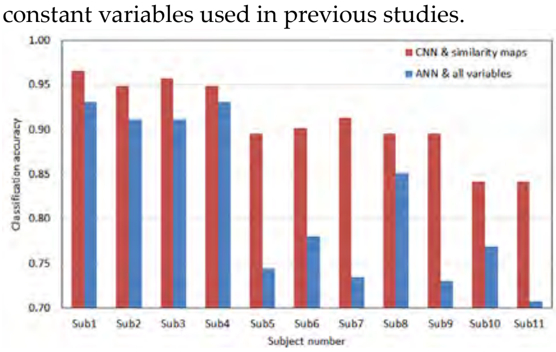

Fig. 6 – Preference classification results for each subject: CNN model trained with luminance similarity index maps. The results of ANN model trained with all reference variables are shown for comparison

The training results of Fig. 7 show that the CNN model with the similarity index maps presents stable training, fast convergence and excellent test accuracy. That is the case for all 11 occupants in the dataset. Therefore, the similarity index is a superior metric when inferring personal visual preference in daylight scenes; in fact, the luminance similarity map concept is valuable even when compared with other powerful CNN models considering raw luminance maps.

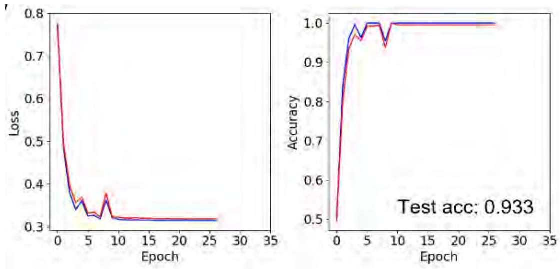

Fig. 7 – Representative CNN training results for subject #6

5.3 Impact of Similarity Index Sign on Preference Classification Performance

The sign function in the new similarity index indicates the relative change in luminance (e.g., from higher to lower and vice versa) for every pixel in the map. To examine if this information is useful (and necessary), we compared the CNN model classification performance using similarity indices without the sign. The results of Table 2 show that the model performance was drastically reduced -in some cases worse than the models trained with reference parameters. Therefore, preserving the direction of change in luminance distribution is necessary when inferring personal daylight preference, and the luminance similarity index is the appropriate metric to capture that information. Excluding this information would eventually confuse the classification model during all stages of training.

Table 2 – Preference classification accuracy for each subject without the sign function (indicating relative change in luminance) and relative reduction compared to CNN models using luminance similarity index maps.

Classification accuracy	Sub1	Sub2	Sub3	Sub4	Sub5	Sub6	Sub7	Sub8	Sub9	Sub10	Sub
Without sign	0.59	0.61	0.63	0.51	0.68	0.54	0.53	0.56	0.60	0.56	0.5
Relative reduction	60.7%	64.6%	65.2%	53.8%	75.9%	59.5%	57.9%	62.3%	66.5%	66.3%	69.5

6. Discussion

Satisfaction with visual conditions dynamically depends on different environmental, contextual and subjective factors. Especially for daylighting conditions, the dynamic nature of daylight, related to outside views, requires a continuously updated evaluation of comparative preferences. The developed luminance similarity index that captures dynamic changes in luminance patterns, utilized in a powerful CNN model, showed an impressive preference classification performance under fixed contextual settings. Outside view preferences are outside the scope of this study; however, there is strong evidence that outside view quality, perception and preference affect satisfaction with the overall visual environment (Giraldo Vasquez et al., 2022; Chinazzo et al., 2019). This study did not consider interaction effects in the preference learning framework. The predictive framework can be used to study if personal daylighting preferences can change with different views and other environmental (Te Kulve et al., 2018; Belia et al., 2021; Pittana et al., 2023) or contextual factors.

Our study is focused on personal daylighting preferences using CNN-based preference classification, which is more challenging than predicting comfort limits under constant luminance distributions. However, our results cannot be generalized like visual comfort metrics. In addition, our training data set was rather limited. It was used to develop the proof of concept and test the ability of luminance similarity maps as personal preference input variables. More daylight scenes with larger variation of luminance patterns are needed for a more complete demonstration of the luminance similarity index concept in real settings.

Finally, in our experimental setting we had the camera next to the person, in order to capture the human FOV and extract lighting metrics (used as reference parameters). The results of this study showed that there is no need to extract these static parameters since they cannot really predict daylighting preferences. The sensor does not necessarily need to match the occupant FOV for the purpose of this work, although estimating or re-projecting the camera-captured luminance distribution to the occupant FOV is possible (Kim & Tzempelikos, 2021, 2022).

7. Conclusion

This study presents a new approach for inferring personal daylight preferences using a new composite luminance similarity index and deep learning techniques. Information from the entire luminance distribution in the FOV was used to extract pair-wise similarity features between HDRI-based luminance maps (different conditions). The luminance similarity index considers both the direction and magnitude of relative luminance change instead of instantaneous metrics.

Comparative visual preference datasets for 11 individuals were generated using collected pairwise HDR images. The generated luminance similarity maps were directly used for training convolutional neural network (CNN) models to classify the occupant's visual preferences. The results showed the superiority of the luminance similarity index map as a preference indicator variable. CNN models trained with luminance similarity index maps showed impressive classification accuracy for all tested subjects in the dataset. Static lighting variables cannot really estimate daylight preferences. Preserving the direction of change in luminance distribution pixel-wisely, is necessary when inferring personal daylight preference.

Acknowledgement

We would like to thank the Center for High Performance Buildings and the Lyles School of Civil Engineering at Purdue for the support.

References

Bakker, L. G., E. C. M. H. Oeffelen, R. C. G. M. Loonen, J. L. M. Hensen. 2014. "User satisfaction and interaction with automated dynamic facades: a pilot study." *Building and Environment.* 78: 44–52. https://doi.org/10.1016/j.buildenv.2014.04.007

Bellia, L., F. R. A. Alfano, F. Fragliasso, B. I. Palella, and G. Riccio. 2021. "On the interaction between lighting and thermal comfort: An integrated approach to IEQ." *Energy and Buildings* 231: 110570. https://doi.org/10.1016/j.enbuild.2020.110570

Chinazzo, G., J. Wienold, and M. Andersen. 2020. "Influence of indoor temperature and daylight illuminance on visual perception." *Lighting Research & Technology* 52(3): 350-370. https://doi.org/10.1177/1477153519859609

Inanici, M. 2006. "Evaluation of high dynamic range photography as a luminance data acquisition system." *Lighting Research & Technology* 38(2): 123–134. https://doi.org/10.1191/1365782806li164oa

Kim, M., and A. Tzempelikos. 2021. "Semi-automated luminance map re-projection via high dynamic range imaging and indoor space 3-D reconstruction." *Automation in Construction* 129: 103812. https://doi.org/10.1016/j.autcon.2021.103812

Kim, M., and A. Tzempelikos. 2022. "Performance evaluation of non-intrusive luminance mapping towards human-centered daylighting control." *Building and Environment* 213: 108857. https://doi.org/10.1016/j.buildenv.2022.108857

Kim, M., I. Konstantzos, and A. Tzempelikos. 2020. "Real-time daylight glare control using a low-cost, window-mounted HDRI sensor." 2020. *Building and Environment* 177: 106912. https://doi.org/10.1016/j.buildenv.2020.106912

Konis K. 2014. "Predicting visual comfort in side-lit open-plan core zones: Results of a field study pairing high dynamic range images with subjective responses." *Energy and Buildings* 77: 67–79. https://doi.org/10.1016/j.enbuild.2014.03.035

Konstantzos, I., and A. Tzempelikos. 2017. "Daylight glare evaluation with the sun in the field of view through window shades." *Building and Environment* 113:65–77. https://doi.org/10.1016/j.buildenv.2016.09.009

Lindelöf, D., and N. Morel. 2007. "Bayesian estimation of visual discomfort." *Building Research & Information* 36(1): 83–96. https://doi.org/10.1080/09613210701544061

Motamed, A., L. Deschamps, and J. L. Scartezzini. 2017. "On-site monitoring and subjective comfort assessment of a sun shadings and electric lighting controller based on novel High Dynamic Range vision sensor." *Energy and Buildings* 149: 58–72. https://doi.org/10.1016/j.enbuild.2017.05.017

Newsham, G., and C. Arsenault. 2009. "A camera as a sensor for lighting and shading control." *Lighting Research & Technology* 41(2): 143–163. https://doi.org/10.1177/1477153508099889

Pierson, C., C. Cauwerts, M. Bodart, and J. Wienold. 2021. "Tutorial: Luminance Maps for Daylighting Studies from High Dynamic Range Photography." *LEUKOS* 17(2): 140–169. https://doi.org/10.1080/15502724.2019.1684319

Pittana, I., Morandi F., Cappelletti F., Gasparella A., and Tzempelikos A. 2023. "Within- and cross-domain effects of environmental factors on students' perception in educational buildings." *Science and Technology for the Built Environment* 29(7): 678-697. https://doi.org/10.1080/23744731.2023.2239080

Prechelt, L. 2012. "Early stopping—but when?" In *Neural Networks: Tricks of the Trade*, edited by Grégoire Montavon, Geneviève B. Orr, and Klaus-Robert Müller, 53–67. Berlin: Springer.

Suk, J. Y., M. Schiler, K. Kensek. 2017. "Investigation of existing discomfort glare indices using human subject study data." *Building and Environment* 113: 121–130. https://doi.org/10.1016/j.buildenv.2016.09.018

te Kulve, M., L. Schlangen, and W. v. M. Lichtenbelt. 2018. "Interactions between the perception of light and temperature." *Indoor Air* 28: 881–891. https://doi.org/10.1111/ina.12500

Van Den Wymelenberg, K., M. Inanici, and P. Johnson. 2010. "The Effect of Luminance Distribution Patterns on Occupant Preference in a Daylit Office Environment." *LEUKOS* 7(2): 103–122. https://doi.org/10.1582/LEUKOS.2010.07.02003

Vasquez, N. G., R. F. Rupp, R. K. Andersen, and J. Toftum. "Occupants' responses to window views, daylighting and lighting in buildings: A critical review." *Building and Environment* 219: 109172. https://doi.org/10.1016/j.buildenv.2022.109172

Villa, C., and R. Labayrade. 2013. "Multi-objective optimisation of lighting installations taking into account user preferences – a pilot study." *Lighting Research & Technology* 45(2): 176–196. https://doi.org/10.1177/1477153511435629

Xiong, J., A. Tzempelikos, I. Bilionis, and P. Karava. 2019. "A personalized daylighting control approach to dynamically optimize visual satisfaction and lighting energy use." *Energy and Buildings* 193: 111–126. https://doi.org/10.1016/j.enbuild.2019.03.046

Xiong, J., A. Tzempelikos, I. Bilionis, N.M. Awalgaonkar, S. Lee, I. Konstantzos, S.A. Sadeghi, and P. Karava. 2018. "Inferring personalized visual satisfaction profiles in daylit offices from comparative preferences using a Bayesian approach." *Building and Environment* 138: 74–88. https://doi.org/10.1016/j.buildenv.2018.04.022

Xiong, J., N. M. Awalgaonkar, A. Tzempelikos, I. Bilionis, and P. Karava. 2020. "Efficient learning of personalized visual preferences in daylit offices: An online elicitation framework." *Building and Environment* 181: 107013. https://doi.org/10.1016/j.buildenv.2020.107013