Spatial Frequency and the Performance of Image-Based Visual Complexity Metrics

DORUKALP DURMUS [ID]

Department of Architectural Engineering, Pennsylvania State University, University Park, PA 16802, USA

e-mail: alp@psu.edu

ABSTRACT There is a wide range of visual and spatial complexity measurement methods that aim to quantify perceived image complexity. While image-based calculation methods (edge detection, image compression, contrast) characterize a digital image, visual perception studies focus on fundamental visual mechanisms, such as contrast sensitivity and visual task performance. Despite the evidence from several vision studies, spatial frequency information has not been widely utilized to assess image complexity. Previous studies suggest that image-based performance metrics are limited in explaining perceived complexity due to confounding factors, such as context, memory, familiarity, and expectation. Here, a visual experiment is conducted to assess the performance of image-based metrics and spatial frequency information using 16 abstract and natural images. A new image complexity metric (R_{spt}), based on detectability suprathreshold, was proposed to benchmark the performance of existing measures. Forty-four naïve participants used a 5-point Likert-type scale to judge the visual complexity of the images displayed on a tablet. Results indicate that root-mean-square error (RMSE) and R_{spt} correlate statistically significantly with subjective evaluations. Biological sex did not affect perceived spatial complexity. While RMSE and R_{spt} can potentially be used to estimate the spatial complexity of display images, the performance of spatial frequency information and image assessment measures in immersive viewing conditions require further research.

INDEX TERMS Digital images, image analysis, natural image statistics, spatial complexity, spatial frequency information.

I. INTRODUCTION

Visual (or spatial) complexity is a widely discussed but not precisely defined term. In its simplest form, visual complexity refers to the level of detail within an image. Visual images, such as computer displays or photographs, are widely used to study visual mechanisms since three-dimensional environmental stimuli are reduced to two-dimensional retinal images. The complexity of images can be analyzed through the characterization of the stimuli and its impact on higher-level visual processes. Theories explaining visual complexity, such as *fractal* [1], *fuzzy* [2], and *information theory* [3], are typically based on the physical stimuli itself and computational in nature. For example, Kolmogorov complexity theory [4] measures the computational resources needed to specify an object (the length of the shortest binary computer program that describes it) [5], and it has been applied to image complexity [6], aesthetics [7] and image similarity [8]. In addition

to complexity, *algorithmic information theory* considers the randomness and probability of an algorithm to reproduce a string (sequence of characters) [9]. The computational probability aims to explain the likelihood of a sensory input leading to the perception of the object.

The effect of visual complexity on cognitive and emotional responses can also be examined without a computational approach. For example, people and other animals find visually complex images intrinsically more attractive than simple images [10]. The complexity of the stimuli also impacts the speed and accuracy of visual search [11], scene preference [12], human cognition, and emotion [13], [14]. In addition, abstract images have been previously used to estimate the perceived complexity of architectural spaces [15] and renderings [16], building facades [17], and artwork [18]. In imaging research, the complexity of an image is typically called *spatial complexity* due to the indirect observation of stimuli (observers perceive stimulus that is digitally processed through another channel, such as a display). Spatial complexity measurement methods have also been

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues [ID].

used for satellite image mapping [19], remote sensing for wildlife conservation [20], and quality assessment of imaging systems [21]. However, the use of image-based metrics to quantify the perceived complexity of scenes under projection systems [22] highlighted the need to investigate the accuracy of spatial complexity metrics. Here, the perceived visual complexity of images is investigated through spatial frequency information and spatial complexity metrics.

II. BACKGROUND

Visual complexity has been a topic of interest for psychologists since the late 19th century. Several theories and models have been proposed based on a single visual form, visual arrays, information pickup, visual displays, perceptual learning, and neural circuit theory to explain perceived visual complexity [23]. Both empiricists' and nativists' theories investigated the roots of the perceived quality of images to build a theoretical framework. Empiricist approaches focusing on the discrimination of primitive (basic) attributes of shapes [24], [25] failed to develop a universal metric [26], [27]. On the other side, from a nativity perspective, two main factors influencing the visual complexity of an image have been proposed; familiarity/novelty and spatial frequency information [27].

A. FAMILARITY, NOVELTY, AND INTEREST

Familiarity is an instinctive parameter that affects cognitive functions. Both familiar and familiar objects have been used in visual experiments. The analysis of a series of black and white drawings of real-life objects resulted in three levels of complexity: low complexity for kitchen utensils, fruits, human body parts, clothing, furniture; medium complexity for vegetables; and high complexity for vehicles, birds, animals, insects and musical instruments [28]. Another study showed that participants require more time to encode, mentally rotate, and compare unfamiliar stimuli [29]. Similarly, familiarity and learning (introduced through training) impacted the perceived complexity of unfamiliar shapes [30]. This study also highlighted the shortcomings of image-based complexity metrics.

On the other hand, visual complexity is also associated with visual search tasks, which, in return, is affected by novelty. For example, the search time and accuracy of simple forms depend on the number, spacing, color, shape, and size of the forms constructing an image [31], [32]. While color discrimination showed no effect, spatial frequency discrimination had an intermediate effect on task performance [32].

Complexity is also considered a major determinant of visual interest and pleasure. For example, Berlyne's inverted U-shape curve for pleasure, interest, and complexity [33] predicts higher arousal for medium complexity images, and lower arousal for low and high complexity images, as shown in Fig. 1. This curve acts similar to the spatial frequency slopes (α) found in the spatial frequency patterns, in that divergence from the medium complexity decreases preference ratings. Berlyne's proposed model has been supported
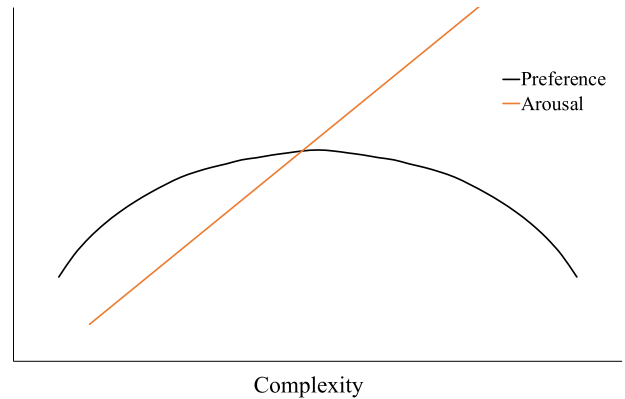

FIGURE 1. A modified version of Berlyne's inverted U-shape curve [33] shows the change in preference and arousal with increased complexity. The inverted U-shape suggests that the medium complexity images are optimal, and further increase in complexity does not equate to pleasure.

by studies investigating abstract patterns [34], paintings [35], and natural images [36], but the model failed to explain preference in the visual complexity of websites [37].

B. SPATIAL FREQUENCY INFORMATION

The physical characteristics of an image can be described using basic attributes, such as contrast, orientation, spatial phase, and spatial frequency. Spatial frequency (f) is a measure of periodic grating across a position within a given distance on the retina [38]. Although the SI unit of spatial frequency is cycles per meter (c/m), it is commonly reported in cycles per degrees (cpd) of visual angle. In vision research, sinusoidal patterns with varying frequencies, amplitudes, and angles, are commonly used to test the visual system's capabilities, especially contrast sensitivity [39], [40] and visual performance [41]. Studies show that spatial contrast sensitivity in adults peaks between 3 cpd and 8 cpd [39], [40], [42].

The experimental observations show that the visual cortex strongly responds to sine-wave gratings, and a multichannel human vision model can explain how neurons of different receptive field sizes create a neural representation of different grating scales [42]. Psychophysical and electrophysiological evidence supporting this hypothesis underlines that the visual cortex has multi-dimensional spatial filters with narrow (non-zero) bandwidths [43]. It is suggested that spatial frequencies of a complex grating are detected independently [44], and adaptation to spatial frequency can enhance contrast sensitivity in the short and long term [45], [46].

Different spatial frequencies can also cause visual discomfort depending on the discomfort groups (sensitivity to physical stimulus). In a study, low and moderate visual discomfort groups found 8 cpd and 12 cpd stimuli unpleasant, whereas high visual comfort group's discomfort peaked around 4 cpd [47]. Another study found increased discomfort with increasing spatial frequency up to 16 cpd for both moderate-high and low visual discomfort groups [48].

The spatial frequency analysis that is based on gratings faced criticism due to its dependence on simpler forms
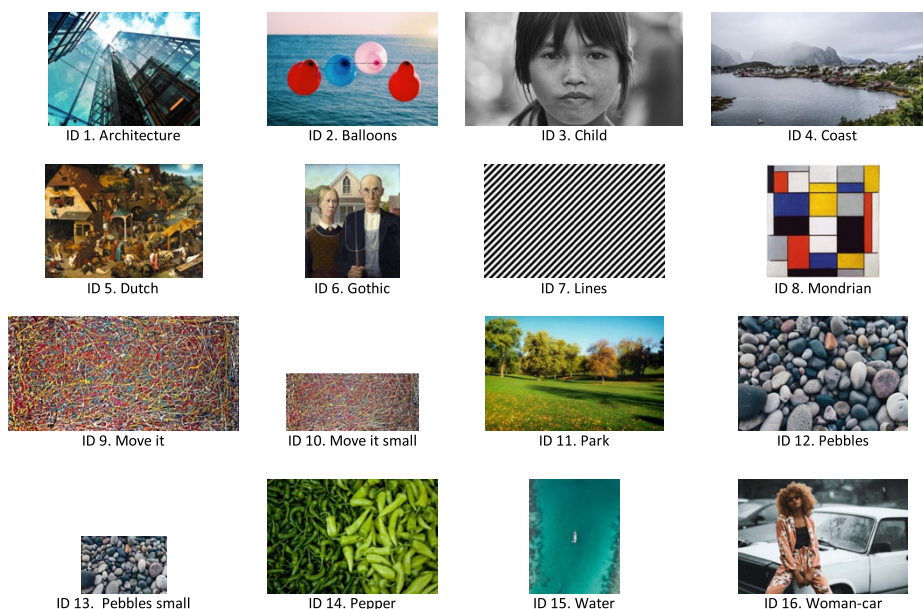
FIGURE 2. images that consist of paintings, environmental scenes, abstract and real objects were used in the experiment. The images are rescaled (normalized to equal height, except *move it small* and *pebbles small*, which are half the height of the other images) for display purposes.

that are not common in the real world. Although the complex real-world images can be considered a superposition of a large number of fundamental patterns, it is unlikely that models based on simple-stimulus experiments are sufficient to predict the visual quality of complex real-world images [21]. These limitations urged researchers to create datasets that include images of both basic and complex patterns [49] and propose methods to assess the spatial quality of complex images, such as blind/referenceless image spatial quality evaluator (BRISQUE), JPEG, and JPEG2000 [50]. However, image quality metrics do not explicitly asses the perceived spatial complexity of an image. In addition, natural image complexity and the decorrelation problem (strong dependency between intra- and inter-channels in natural images) [21] may impact the accuracy of image quality metrics.

The growing evidence of the human visual system functioning as a frequency analyzer led to the analysis of natural image statistics using spatial frequency information. The regularity (common patterns) in the natural images have been demonstrated to exist in spatial [51] and wavelet domains [52]. The natural statistics underlie the non-randomness in images and enable predicting the visual quality of natural images. For example, the power spectrum of natural images reportedly follows a frequency function of

$$A(f) = f^{-\alpha} \qquad (1)$$

where the amplitude (A) is averaged across all the orientations, f is spatial frequency, and α is the negative slope on log-log coordinates [53]. The slope α varies from image to image, and it typically ranges between -0.7 and -1.5 in achromatic images [53]–[55], 0 and -2 in chromatic images [56].

The studies suggested that slope in natural images peak between -1.1 and -1.3 [53], [55], and deviation from natural scene statistics may cause discomfort [56]. The discomfort caused by non-natural image statistics hints to an evolutionary adaptation mechanism. It should be noted that the natural statistics of images are limited with datasets analyzed in these studies. Therefore, slopes may considerably vary across images.

While spatial frequency theory aims to ground the visual perception of images to a fundamental cognitive framework, the effect of physical stimulus characterization on perceived complexity is still unclear. While the natural statistics and image-based spatial complexity exist in different domains of research, the performance of f statistics has not been previously tested for visual complexity. Here, the performance of spatial frequency information, spatial complexity metrics, and an arbitrary complexity metric have been analyzed.

III. METHODS
A. STIMULI AND COMPLEXITY MEASURES
The performance of nine visual complexity measures have been compared using 16 images selected from open-source repositories. The image dataset included natural images, such as humans, landscapes, paintings, human-made objects, as well as abstract images, such as diagonal achromatic lines, as shown in Fig. 2. All the images were colored, except for two images (*child* and *lines*). The smaller version of two of the images (*pebbles* and *move it*) were added to the set to test the effect of size on perceived and calculated spatial image complexity. Three images consisted of repeating patterns (*pebbles*, *pebbles small*, *pepper*).

TABLE 1. Spatial complexity properties of the images used in the experiment.

| ID (pixels) | D | E | α | R_{spt} | Image compression | | | Edge detection | |
					GIF	JPEG	RMSE	Canny-GIF	Sobel-GIF
Architecture (1280x960)	1.827	7.7175	-1.03	400	647 KB	334 KB	22.26	100 KB	42 KB
Balloons (1280x853)	1.924	7.7542	-1.00	170	240 KB	164 KB	6.58	70 KB	17 KB
Child (1280x664)	1.854	7.6461	-1.06	64	714 KB	125 KB	8.90	51 KB	22 KB
Coast (1280x847)	1.860	7.7361	-0.95	291	456 KB	220 KB	26.91	61 KB	30 KB
Dutch (1000x718)	1.903	7.4905	-1.07	325	641 KB	326 KB	91.07	68 KB	30 KB
Gothic (770x925)	1.863	7.7114	-1.08	226	562 KB	244 KB	43.90	53 KB	28 KB
Lines (1131x707)	1.773	4.9608	-0.72	42	532 KB	313 KB	13.28	68 KB	40 KB
Mondrian (850x845)	1.839	7.3388	-0.99	20	421 KB	179 KB	18.99	48 KB	31 KB
Move it (1000x489)	1.948	7.5204	-0.15	466	598 KB	455 KB	154.75	54 KB	18 KB
Move it small (342x147)	1.888	7.8658	-0.38	301	145 KB	139 KB	204.16	6 KB	1 KB
Park (1280x813)	1.946	7.7558	-1.00	425	743 KB	382 KB	44.46	72 KB	24 KB
Pebbles (1280x853)	1.809	7.7105	-1.26	141	908 KB	308 KB	20.43	99 KB	43 KB
Pebbles small (128x85)	1.806	7.7461	-0.19	104	50 KB	59 KB	68.57	1 KB	561 B
Pepper (1280x851)	1.847	6.8720	-1.46	224	812 KB	293 KB	12.76	85 KB	29 KB
Water (1280x1600)	1.958	6.3486	-0.99	107	768 KB	188 KB	9.52	194 KB	44 KB
Woman-car (1280x853)	1.852	7.7784	-1.37	126	526 KB	169 KB	10.16	81 KB	30 KB

TABLE 2. Spearman's correlation coefficient (and *p*-values) for the visual complexity measures are shown in the correlation matrix.

	D	E	α	R_{spt}	GIF	JPEG	RMSE	Canny-GIF	Sobel-GIF
D	1.00								
	-								
E	0.25	1.00							
	(0.36)	-							
α	0.11	0.00	1.00						
	(0.69)	(1.00)	-						
R_{spt}	0.50	0.40	0.22	1.00					
	(0.05*)	(0.01*)	(0.42)	-					
GIF	0.11	-0.21	-0.62	0.13	1.00				
	(0.70)	(0.44)	(0.01*)	(0.63)	-				
JPEG	0.23	-0.17	-0.05	0.66	0.60	1.00			
	(0.39)	(0.52)	(0.85)	(0.005**)	(0.02*)	-			
RMSE	0.31	0.27	0.71	0.53	-0.39	0.14	1.00		
	(0.24)	(0.31)	(0.01*)	(0.03*)	(0.13)	(0.61)	-		
Canny-GIF	0.30	-0.36	-0.51	-0.06	0.67	0.26	-0.51	1.00	
	(0.25)	(0.17)	(0.04*)	(0.81)	(0.004**)	(0.34)	(0.05*)	-	
Sobel-GIF	-0.33	0.09	0.47	-0.23	-0.51	-0.45	0.04	-0.37	1.00
	(0.21)	(0.73)	(0.06)	(0.38)	(0.05*)	(0.08)	(0.88)	(0.16)	-

Although there are several spatial information and compression-based image quality measures, the analysis was limited to visual complexity metrics and spatial frequency information. The spatial characteristics of images and the calculated metrics are given in Table 1.

Inspired by the definition of the Kolmogorov complexity, the size of various compressed image formats has been previously used to estimate the complexity of images [18], [30], [57], [58]. Here, only two of the most popular lossy image compression methods (GIF and JPEG) have been used. GIF compression was performed with a local selective palette of 256 colors and at normal row order. Although a large image size is associated with image complexity (simple images contain more redundant information that can be compressed; therefore they have smaller file sizes than complex images), it should be noted that compressed file size is also affected by other factors, such as luminance and chrominance information [30].

Additional visual complexity metrics, such as root-mean-square error (RMSE) [6], [16], [57], [58], fractal dimension (D) [18], [59], and entropy (E) of the image intensity histogram of a grayscale image [57], were also investigated.

The RMSE (the difference between the original image and the lossy compressed image), spatial frequency slope α, and entropy E were calculated via MATLAB® Image Processing Toolbox™. Fractal dimension D was calculated by a box-counting method via ImageJ, a public-domain image-processing software. Edge detection algorithms commonly used in spatial complexity and quality assessment, such as Canny and Sobel methods [30], [57], [60], were also analyzed. In addition to the existing measures of complexity, a new arbitrary metric (R_{spt}), based on suprathreshold detectability, was introduced to benchmark existing spatial complexity assessment measures. The calculation of the R_{spt} is explained in the following section.

The correlation coefficients between all the complexity measures are given in Table 2. The highest correlation was between spatial frequency slope and RMSE (0.71). JPEG correlated highly with GIF and R_{spt} (0.60 and 0.66, respectively). While there was a high correlation between GIF and Canny-GIF (0.67), GIF negatively correlated with Sobel-GIF (−0.51). The correlation between GIF and spatial frequency slope was also negative (−0.62). The *p*-values of the multiple correlations were tested using the Benjamini and Hochberg

method [61], which is a powerful correction method for false discovery rate. In the level of $p < 0.01$, the statistically significant correlations were found for only two pairs: JPEG - R_{spt} and GIF - Canny-GIF.

B. A NEW VISUAL COMPLEXITY MEASURE

Goodhart's law implies that "when a measure becomes a target, it ceases to be a good measure" [62]. To address the potential misuse of image-quality assessment methods, an arbitrary visual complexity metric (R_{spt}) was introduced as a pseudo-random benchmark.

The first step of the calculation of R_{spt} is converting an image into grayscale and then binarizing it through Matlab® Image Region Analyzer application, which uses an adaptive thresholding method [63]. A series of low to high complexity images were visually judged by the author to identify the smallest detectable region. A detection threshold of 1 in 25000 pixels was found reasonably accurate for a variety of images. The number of suprathreshold regions (R_{spt}), approximating the perceptual complexity of images, is calculated

$$R_{spt} = \sum_i [PR_i > PR_{total}/25000] \quad (2)$$

where PR is the size of a region, i is the region index, and PR_{total} is the total number of pixels in the image. R_{spt} quantifies the number of detectable regions in an image and considers an image with a high number of regions to be more complex. The arbitrary suprathreshold metric R_{spt} was tested against a set of images previously used in visual clarity and blur perception research [64], as shown in Table 3. An empty white or black image ($R_{spt} = 0$) and Jackson Pollock's *move it* ($R_{spt} = 466$) are the known boundaries of the R_{spt} scale. All of the tested images (i.e., natural scenes and paintings) lie within these boundaries, and approximately $R_{spt} < 100$ denotes visual simplicity and $R_{spt} > 300$ denotes high complexity.

TABLE 3. R_{spt} complexity metric values for sample images.

Painting	Width x height (mm)	R_{spt}	Spatial perceptual information (SI)
Composition A by Piet Mondrian	582 x 578	20	34.5
American Gothic by Grant Wood	580 x 700	226	69.6
Dutch Proverbs by Pieter Bruegel the Elder	822 x 582	325	94.5

C. PARTICIPANTS AND PROCEDURES

Forty-four naïve participants (23 female, 17 male, four undisclosed/other) with normal color vision as tested by the Ishihara pseudoisochromatic plates took part in the experiment. Participants viewed 16 images on a tablet (Apple iPad® 2) with a peak brightness of around 410 cd/m². They hold the tablet around half an arm's length (approximately 50 cm), which equates to approximately 27° visual angle.

Participants were instructed to judge the visual complexity of the images using a 5-point Likert-type scale with a neutral midpoint. The scales were "very complex," "complex," "medium," "simple," and "very simple." The order of the images was randomized for every participant. Participants were allowed to move back and forth between images and change their judgments (anchoring was allowed). Although there was no minimum or maximum time limit to make judgments, participants did not spend more than three minutes per image.

IV. RESULTS AND DISCUSSION

A. QUANTITATIVE IMAGE COMPLEXITY ASSESSMENT

Participants' subjective evaluations were tested for normality with the Shapiro–Wilk test at $p = 0.01$. The test showed that participants' subjective visual complexity evaluations were not normally distributed. Based on the non-normal distribution, interobserver differences were tested with the Kruskal-Wallis test. The null hypothesis was rejected ($\chi^2 = 267.77$, $p < 0.001$, df = 703), which hints a statistically significant variation in the interobserver judgments, as shown in Fig. 3. However, performing multi-sample comparisons in the Kruskal-Wallis test can inflate Type I errors (rejection of a true null hypothesis). Therefore, the rejection of the null hypothesis should be taken with a grain of salt.
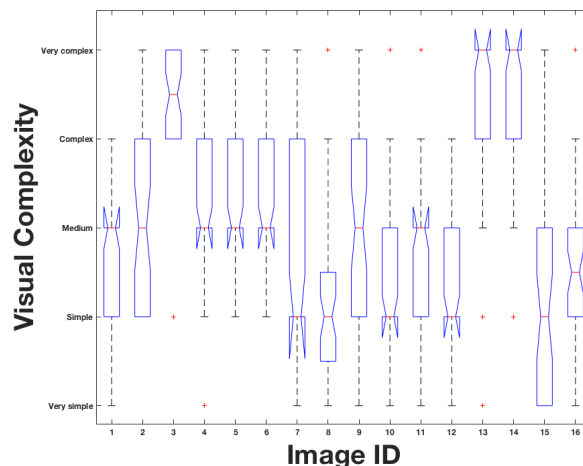

FIGURE 3. Interobserver differences were tested with the Kruskal-Wallis test, and the null hypothesis was rejected ($\chi^2 = 267.77$, $p < 0.001$, df = 703).

Spearman's rank correlations between subjective judgments and computation measures for all images are given in Table 4. There were only two cases of statistically significant relationship between subjective assessments and metrics; RMSE ($\rho = 0.92$, $p < 0.001$) and R_{spt} ($\rho = 0.62$, $p = 0.01$).

The influence of biological sex was investigated by looking into female ($n = 23$) and male responses ($n = 17$). Spearman's rank correlation for image complexity metrics was not different for biological sex. The only small difference was the

TABLE 4. Spearman's rank correlation coefficient was used to test the image complexity metrics and subjective assessments of all participants.

Spatial complexity measure	Spearman's ρ
Fractal dimension D	0.18
Entropy E	0.32
Spatial frequency slope α	0.34
R_{st}	0.62*
GIF	-0.24
JPEG	0.30
RMSE	0.92***
Canny-GIF	-0.40
Sobel-GIF	-0.17

increased statistical significance for R_{spt} for males compared to females, as shown in Table 5.

TABLE 5. Spearman's rank correlation coefficient was used to test the image complexity metrics and subjective assessments of male and female participants.

Spatial complexity measure	Spearman's ρ (female)	Spearman's ρ (male)
Fractal dimension D	0.19	0.19
Entropy E	0.39	0.14
Spatial frequency slope α	0.27	0.27
R_{st}	0.60*	0.67**
GIF	-0.27	-0.12
JPEG	0.26	0.35
RMSE	0.88***	0.88***
Canny-GIF	-0.36	-0.34
Sobel-GIF	-0.19	-0.22

The effect of size on the perception of spatial complexity was also analyzed with the Wilcoxon-Mann-Whitney two-sample rank-sum test. The large and small variations of two images (*move it* and *pebbles*) were tested separately. The difference between *move it* and *move it small* was not statistically significant ($U = 957.5$, $n_1 = n_2 = 44$, $z = 0.10, p = 0.46$). Similarly, the difference between *pebbles* and *pebbles small* was not statistically significant ($U = 790$, $n_1 = n_2 = 44$, $z = 1.56$, $p = 0.12$). However, the effect sizes for both of the results were small ($r = 0.01$ and $r = 0.23$, respectively). The image size did not have a significant effect on female or male participants' subjective evaluations.

The correlation between image complexity metrics and subjective evaluations of paintings (*Mondrian, Dutch, Gothic, move it, move it small*) and natural images (*architecture, balloons, child, coast, park, pebbles, pebbles small, pepper, water, woman-car*) are given in Table 6. Although there was a high correlation between subjective ratings and several metrics, the correlations were not statistically significant (D $\rho = 0.79$ $p = 0.20$, α $\rho = 0.79$ $p = 0.20$, R_{spt} $\rho = 0.79$ $p = 0.20$, RMSE $\rho = 0.95$ $p = 0.07$, Sobel-GIF $\rho = -0.79$ $p = 0.20$). The only statistically significantly correlation with subjective evaluations was with RMSE ($\rho = 0.87$ $p = 0.001$) for natural images.

In previous studies, medium correlation with JPEG and mixed results for GIF correlation were recorded [18], [57].

TABLE 6. Spearman's rank correlation coefficient was used to test the image complexity metrics and subjective assessments of paintings and natural images.

Spatial complexity measure	Spearman's ρ (paintings)	Spearman's ρ (natural images)
Fractal dimension D	0.79	-0.42
Entropy E	0.47	0.41
Spatial frequency slope α	0.79	0.33
R_{st}	0.79	0.43
GIF	0.00	-0.16
JPEG	0.16	0.43
RMSE	0.95	0.87*
Canny-GIF	0.00	-0.09
Sobel-GIF	-0.79	0.44

Compression file-formats correlated significantly with subjective complexity judgments of representational paintings [57], whereas edge detection methods correlated moderately with subjective measures for both environmental scenes and paintings [30]. The results found in this study do not strongly support these findings. Instead, results support evidence from another study where RMS measures outperformed JPEG compression [58]. The discrepancy in these results can be attributed to differences in experimental procedures. Training and elaborate explanations of the sought terms allow participants to be more consistent in their responses [65]. Precautionary methods (e.g., training) may provide different results compared to studies where data are collected using a less restrictive method. The wording of the research questions or statements can also cause bias [66]. In this study, the question was limited to "visual complexity," but there was no training provided prior to the experiment.

B. QUALITATIVE IMAGE COMPLEXITY ASSESSMENT
In a preliminary study, a small group of participants were presented the same stimuli (16 images) to examine the perception and definition of visual complexity. When the complexity was defined as "ease of remembering details," participants reported that they would provide different responses than when there was no definition (e.g., *Mondrian* might be easy to look at, but hard to remember, therefore complex). Participants seem to have a non-universal and intrinsic definition of complexity.

In the actual experiment, participants were asked to judge only "visual" aspects of the presented images rather than encoding the semantic complexity. Some participants reported that certain features of an image were *whole objects* (i.e., the face of a human) and they did not consider the whole objects to be complex. Some participants thought that abstract paintings (e.g., *move it*) have no identifiable objects in them, and they considered the abstract structure to be a background; therefore, they did not find it complex.

Some images (e.g., *lines, architecture*) caused nausea and dizziness, and they were considered complex due to the reaction they caused, not due to their structural formation. This finding supports the notion that while the human visual system is attuned to horizontal and vertical gratings [67],

they may cause somatic and perceptual side-effects [68]. On the other side, one of the participants reported that the *park* was "predictable," therefore participants considered it to be relaxing and not complex.

Images with repetitive objects (e.g., *pebbles*, *lines*) caused the highest disagreement among participants. *Pebbles small* was considered complex for some participants, but not for others. Although most participants did not report considering color as an influencing factor, a high level of variation in *woman-car* image hints the role of color contrast in visual cognition. This is likely due to the influence of color contrast on visual performance [69].

Three paintings (*Mondrian*, *Dutch*, *Gothic*) used in a previous experiment [64] showed high correlations to their intended use. Participants typically judged *Dutch* image to be the most complex in the survey (even more complex than *move it*). This is possibly due to participants' consideration of the number of interesting objects (i.e., relatable objects such as people), rather than abstract shapes (grating). *Gothic* was considered medially complex, and *Mondrian* was generally judged to be visually simple, with some noted exceptions.

C. IMPLICATIONS FOR IMMERSIVE VIEWING CONDITIONS

Subjective evaluations of 16 display images supported the notion that image-based complexity metrics are limited in quantifying perceived complexity with a caveat. While RMSE performed well in predicting the perceived visual complexity, it may not be possible to use the root-mean-square analysis in realistic (immersive) environments due to its dependence on a reference condition (undistorted image). In immersive viewing conditions, the reference point of a subjective evaluation is typically stored in participants' short or long-term memory. The spatial memory reference is often distorted in time, even within a very short period [70]. The lack of a meaningful reference in realistic, immersive environments limits the use of reference-based image assessment measures.

Another important difference between displayed images and immersive environments is the visual angle. The images displayed in these experiments forces participants to make judgments based on narrow to medium field of view. Therefore, the effects of background and surrounding field may not be taken into account in the image assessment metrics. Visual acuity peaks in the fovea, but the optical quality of the retinal image reduces slowly with peripheral angle and spatial frequency sensitivity changes with visual eccentricity [71]. The changes in the visual system with the visual angle hints the importance of experimental conditions in image quality assessment.

The discrepancies between the metrics and subjective assessments can also be linked to other endogenous and exogenous elements. Image-based metrics do not consider internal factors, such as context, memory, familiarity, and expectation [30]. There are several external factors, such as illumination levels [72], spectral power distribution of a light source [73], and adaptation [74] that may impact subjective and objective judgments of visual perception for both display images and real scenes [75]. The limitations are highly relevant for integrative lighting systems that are based on the modeling of the spatial and spectral sensitivity of the visual system. For example, it is possible to conceptualize a real-time integrative system that detects the physical characteristics of the built environment using sensors (CCD or CMOS for spectral, spatial, and luminance imaging of the built environment [76], [77]) to optimize the light output for energy efficiency, visual comfort, and visual performance. Such a system can estimate the perceived spatial quality and complexity of the built environment and enable making predictions through computational models based on machine learning and optimization algorithms. However, the performance of such an integrative system depends on the accuracy and precision of the mathematical models and parameters.

V. CONCLUSIONS

Image quality assessment measures have been widely used to assess the quality of display images. The performance of these measures varies depending on the dataset and experimental methods. A visual experiment with 16 images that consist of paintings, natural scenes, and abstract structures was conducted to assess the performance of spatial frequency information and spatial complexity measures. Images used in this experiment contained unusually repetitive structures, varying sizes, and chromatic characteristics (i.e., not all the images were natural). The wide variety of characteristics in the physical stimuli enabled a rich exploration, where root-mean-square error (RMSE) and newly proposed R_{spt} correlated significantly with subjective assessments. However, visual complexity judgments are likely influenced by higher-level processes, such as grouping, object identification, and emotional responses.

It should be noted that reported findings are limited to visual complexity. Although image complexity and image quality are interrelated, they are not interchangeable concepts. Future research should investigate the use of image-based quality assessment measures and spatial frequency information for other dimensions of visual perception, such as overall image quality and visual clarity.

In this experiment, spatial frequency slope did not achieve a statistically significant correlation with subjective evaluations of visual complexity. There may be two main reasons for the lack of correlation: high variation in the image dataset and experimental conditions. Spatial frequency information is widely used to estimate contrast sensitivity and visual performance for natural and simple images. However, the dataset used here was not limited to natural images.

RMSE and R_{spt} performed well in this study, and it is likely that other studies will find statistically significant results for other complexity measures. The image complexity metrics are typically used to quantify the perceived complexity of images displayed on a screen. However, visual tasks in the real-world can be harder to complete and model [78].

Moreover, immersive environments illuminated by an integrative lighting system based on light projection [79], [80] requires further investigation. Future research aims to address the effects of exogenous elements, such as the spectrum and intensity of the illumination and spatial frequency information, on visual perception.

REFERENCES

[1] A. E. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Trans. Image Process.*, vol. 1, no. 1, pp. 18–30, Jan. 1992.

[2] M. Cardaci, V. Di Gesù, M. Petrou, and M. E. Tabacchi, "A fuzzy approach to the evaluation of image complexity," *Fuzzy Sets Syst.*, vol. 160, no. 10, pp. 1474–1484, May 2009.

[3] J. Rigau, M. Feixas, and M. Sbert, "An information-theoretic framework for image complexity," in *Proc. 1st Eurographics Conf. Comput. Aesthetics Graph., Vis. Imag.*, 2005, pp. 177–184.

[4] A. N. Kolmogorov, "On tables of random numbers," *Sankhyā, Indian J. Statist., Ser. A*, vol. 25, no. 4, pp. 369–376, Dec. 1963.

[5] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.

[6] H. Yu and S. Winkler, "Image complexity and spatial information," in *Proc. 1th Int. Workshop Qual. Multimedia Exper. (QoMEX)*, Jul. 2013, pp. 12–17.

[7] J. Rigau, M. Feixas, and M. Sbert, "Conceptualizing Birkhoff's aesthetic measure using Shannon entropy and kolmogorov complexity," in *Computational Aesthetics*. Goslar, Germany: Eurographics Association, 2007, pp. 105–112.

[8] N. Nikvand and Z. Wang, "Generic image similarity based on kolmogorov complexity," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 309–312.

[9] R. Solomonoff, "The application of algorithmic probability to problems in artificial intelligence," *Mach. Intell. Pattern Recognit.*, vol. 4, pp. 473–491, Jan. 1986.

[10] D. E. Berlyne, "Curiosity and exploration," *Science*, vol. 153, no. 3731, pp. 25–33, 1966.

[11] T. van Marlen, M. van Wermeskerken, and T. van Gog, "Effects of visual complexity and ambiguity of verbal instructions on target identification," *J. Cognit. Psychol.*, vol. 31, no. 2, pp. 206–214, Feb. 2019.

[12] T. Heath, S. G. Smith, and B. Lim, "Tall buildings and the urban skyline: The effect of visual complexity on preferences," *Environ. Behav.*, vol. 32, no. 4, pp. 541–556, Jul. 2000.

[13] A. N. Tuch, J. A. Bargas-Avila, K. Opwis, and F. H. Wilhelm, "Visual complexity of websites: Effects on users' experience, physiology, performance, and memory," *Int. J. Hum.-Comput. Stud.*, vol. 67, no. 9, pp. 703–715, Sep. 2009.

[14] R. Pieters, M. Wedel, and R. Batra, "The stopping power of advertising: Measures and effects of visual complexity," *J. Marketing*, vol. 74, no. 5, pp. 48–60, Sep. 2010.

[15] A. Olivia, M. L. Mack, M. Shrestha, and A. Peeper, "Identifying the perceptual dimensions of visual complexity of scenes," in *Proc. Annu. Meeting Cogn. Sci. Soc.*, 2004, pp. 1–7.

[16] S. Rockcastle, M. Amundadottir, and M. Andersen, "Contrast measures for predicting perceptual effects of daylight in architectural renderings," *Lighting Res. Technol.*, vol. 49, no. 7, pp. 882–903, Nov. 2017.

[17] K. Chamilothori, *Perceptual Effects of Daylight Patterns in Architecture*. Lausanne, Switzerland: EPFL, 2019.

[18] A. Forsythe, M. Nadal, N. Sheehy, C. J. Cela-Conde, and M. Sawey, "Predicting beauty: Fractal dimension and visual complexity in art," *Brit. J. Psychol.*, vol. 102, no. 1, pp. 49–70, Feb. 2011.

[19] L. D. Cola, "Simulating and mapping spatial complexity using multi-scale techniques," *Int. J. Geographical Inf. Syst.*, vol. 8, no. 5, pp. 411–427, Sep. 1994.

[20] A. Lim, J. D. Hedley, E. LeDrew, P. J. Mumby, and C. Roelfsema, "The effects of ecologically determined spatial complexity on the classification accuracy of simulated coral reef images," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 965–978, May 2009.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[22] D. Durmus and W. Davis, "Visual clarity and blur acceptability in complex illuminated images," in *Proc. Light, Energy Environ.* Singapore: OSA, 2018, Paper SW2D.4, doi: 10.1364/SSL.2018.SW2D.4.

[23] D. C. Donderi, "Visual complexity: A review," *Psychol. Bull.*, vol. 132, p. 73, Jan. 2006.

[24] C. J. Johnson, A. Paivio, and J. M. Clark, "Cognitive components of picture naming," *Psychol. Bull.*, vol. 120, no. 1, p. 113, 1996.

[25] M. García, A. N. Badre, and J. T. Stasko, "Development and validation of icons varying in their abstractness," *Interacting Comput.*, vol. 6, no. 2, pp. 191–211, Jun. 1994.

[26] S. J. McDougall, O. de Bruijn, and M. B. Curry, "Exploring the effects of icon characteristics on user performance: The role of icon concreteness, complexity, and distinctiveness," *J. Experim. Psychol., Appl.*, vol. 6, no. 4, p. 291, 2000.

[27] A. Forsythe, "Visual complexity: Is that all there is," in *Proc. Int. Conf. Eng. Psychol. Cogn. Ergonom.*, 2009, pp. 158–166.

[28] J. G. Snodgrass and M. Vanderwart, "A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity," *J. Exp. Psychol., Hum. Learn. Memory*, vol. 6, no. 2, p. 174, 1980.

[29] C. E. Bethell-Fox and R. N. Shepard, "Mental rotation: Effects of stimulus complexity and familiarity," *J. Experim. Psychol., Hum. Perception Perform.*, vol. 14, no. 1, p. 12, 1988.

[30] A. Forsythe, G. Mulhern, and M. Sawey, "Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing," *Behav. Res. Methods*, vol. 40, no. 1, pp. 116–129, Feb. 2008.

[31] B. Julesz and J. R. Bergen, "Human factors and behavioral science: Textons, the fundamental elements in preattentive vision and perception of textures," *Bell Syst. Tech. J.*, vol. 62, no. 6, pp. 1619–1645, Jul. 1983.

[32] P. Verghese and K. Nakayama, "Stimulus discriminability in visual search," *Vis. Res.*, vol. 34, no. 18, pp. 2453–2467, Sep. 1994.

[33] D. E. Berlyne, *Aesthetics and Psychobiology*. New York, NY, USA: Appleton-Century-Crofts, 1971.

[34] J. Friedenberg and B. Liby, "Perceived beauty of random texture patterns: A preference for complexity," *Acta Psychologica*, vol. 168, pp. 41–49, Jul. 2016.

[35] D. H. Saklofske, "Visual aesthetic complexity, attractiveness and diversive exploration," *Perceptual Motor Skills*, vol. 41, no. 3, pp. 813–814, Dec. 1975.

[36] R. P. Taylor, B. Spehar, J. A. Wise, C. W. Clifford, B. R. Newell, C. M. Hagerhall, T. Purcell, and T. P. Martin, "Perceptual and physiological responses to the visual complexity of fractal patterns," *Nonlinear Dyn. Psychol. Life. Sci*, vol. 9, pp. 89–114, Jan. 2005.

[37] C. Stickel, M. Ebner, and A. Holzinger, "The XAOS metric-understanding visual complexity as measure of usability," in *Proc. Symp. Austrian HCI Usability Eng. Group*, 2010, pp. 278–290.

[38] R. Sekuler, *Perception*, 4th ed. Boston, MA, USA: McGraw-Hil, 2002.

[39] M. A. Georgeson and G. D. Sullivan, "Contrast constancy: Deblurring in human vision by spatial frequency channels.," *J. Physiol.*, vol. 252, no. 3, pp. 627–656, Nov. 1975.

[40] J. G. Robson, "Spatial and temporal contrast-sensitivity functions of the visual system," *J. Opt. Soc. Amer.*, vol. 56, no. 8, pp. 1141–1142, Aug. 1966.

[41] E. L. Cameron, C. P. Talgar, and M. Carrasco, "Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size," *Spatial Vis.*, vol. 15, no. 1, pp. 61–75, 2001.

[42] F. W. Campbell and J. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol.*, vol. 197, no. 3, p. 551, 1968.

[43] L. Maffei and A. Fiorentini, "The visual cortex as a spatial frequency analyser," *Vis. Res.*, vol. 13, no. 7, pp. 1255–1267, Jul. 1973.

[44] M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision," *J. Opt. Soc. Amer.*, vol. 61, no. 9, pp. 1176–1186, 1971.

[45] K. K. De Valois, "Spatial frequency adaptation can enhance contrast sensitivity," *Vis. Res.*, vol. 17, no. 9, pp. 1057–1065, Jan. 1977.

[46] H. R. Wilson and R. Humanski, "Spatial frequency adaptation and contrast gain control," *Vis. Res.*, vol. 33, no. 8, pp. 1133–1149, May 1993.

[47] E. Conlon, W. Lovegrove, S. Barker, and E. Chekaluk, "Visual discomfort: The influence of spatial frequency," *Perception*, vol. 30, no. 5, pp. 571–581, May 2001.

[48] E. G. Conlon, W. J. Lovegrove, E. Chekaluk, and P. E. Pattison, "Measuring visual discomfort," *Vis. Cognition*, vol. 6, no. 6, pp. 637–663, Dec. 1999.

[49] A. B. Watson, "Visual detection of spatial contrast patterns: Evaluation of five simple models," *Opt. Express*, vol. 6, no. 1, pp. 12–33, 2000.

[50] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[51] D. L. Ruderman, "The statistics of natural images," *Netw., Comput. Neural Syst.*, vol. 5, no. 4, pp. 517–548, 1994.

[52] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S.-C. Zhu, "On advances in statistical modeling of natural images," *J. Math. Imag. Vis.*, vol. 18, no. 1, pp. 17–33, 2003.

[53] P. J. Bex and W. Makous, "Spatial frequency, phase, and the contrast of natural images," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 19, no. 6, pp. 1096–1106, 2002.

[54] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 12, pp. 2379–2394, 1987.

[55] D. J. Tolhurst, Y. Tadmor, and T. Chao, "Amplitude spectra of natural images," *Ophthalmic Physiological Opt.*, vol. 12, no. 2, pp. 229–232, Dec. 2007.

[56] I. Juricevic, L. Land, A. Wilkins, and M. A. Webster, "Visual discomfort and natural image statistics," *Perception*, vol. 39, no. 7, pp. 884–899, Jul. 2010.

[57] M. M. Marin and H. Leder, "Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music," *PLoS ONE*, vol. 8, no. 8, Aug. 2013, Art. no. e72412.

[58] A. Cavalcante, A. Mansouri, L. K. E. Mansouri, B. Kardec, Y. Takeuchi, N. Matsumoto, and N. Ohnishi, "On the relation between contrast and the perception of visual complexity," in *Proc. 21st Annu. Conf. Jpn. Neural Netw. Soc.*, 2011, p. 2011.

[59] L. Kacha, N. Matsumoto, A. Mansouri, and A. Cavalcante, "Predicting perceived complexity using local contrast statistics and fractal information," Courrier du Savoir, Biskra, Algeria, 2013, pp. 89–97.

[60] P. Machado, J. Romero, M. Nadal, A. Santos, J. Correia, and A. Carballal, "Computerized measures of visual complexity," *Acta Psychologica*, vol. 160, pp. 43–57, Sep. 2015.

[61] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc., Ser. B, Methodol.*, vol. 57, no. 1, pp. 289–300, Jan. 1995.

[62] M. Strathern, "'Improving ratings': Audit in the British University system," *Eur. Rev.*, vol. 5, no. 3, pp. 305–321, 1997.

[63] D. Bradley and G. Roth, "Adaptive thresholding using the integral image," *J. Graph. Tools*, vol. 12, no. 2, pp. 13–21, Jan. 2007.

[64] D. Durmus and W. Davis, "Blur perception and visual clarity in light projection systems," *Opt. Express*, vol. 27, no. 4, p. A216, 2019.

[65] N. Chater, "Reconciling simplicity and likelihood principles in perceptual organization," *Psychol. Rev.*, vol. 103, no. 3, p. 566, 1996.

[66] R. K. Garg, "The influence of positive and negative wording and issue involvement on responses to likert scales in marketing research," *Market Res. Soc., J.*, vol. 38, no. 3, pp. 1–11, May 1996.

[67] R. L. Solso, *Cognition and the Visual Arts*. Cambridge, MA, USA: MIT Press, 1996.

[68] A. J. Wilkins and B. J. W. Evans, "Visual stress, its treatment with spectral filters, and its relationship to visually induced motion sickness," *Appl. Ergonom.*, vol. 41, no. 4, pp. 509–515, Jul. 2010.

[69] M. V. Mclean, "Brightness contrast, color contrast, and legibility," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 7, no. 6, pp. 521–527, Dec. 1965.

[70] S. Werner and J. Diedrichsen, "The time course of spatial memory distortions," *Memory Cognition*, vol. 30, no. 5, pp. 718–730, Jul. 2002.

[71] L. N. Thibos and A. Bradley, "The limits of performance in central and peripheral vision," in *SID Dig. Tech. Papers*, vol. 22, 1991, pp. 301–303.

[72] W. Hu and W. Davis, "Dimming curve based on the detectability and acceptability of illuminance differences," *Opt. Express*, vol. 24, no. 10, p. A885, 2016.

[73] D. Durmus and W. Davis, "Object color naturalness and attractiveness with spectrally optimized illumination," *Opt. Express*, vol. 25, no. 11, pp. 12839–12850, 2017.

[74] P. A. Jay, "Lighting and visual perception," *Lighting Res. Technol.*, vol. 3, no. 2, pp. 133–146, Jun. 1971.

[75] K. Chiu and P. Shirley, "Rendering, complexity, and perception," in *Photorealistic Rendering Techniques*. Berlin, Germany: Springer, 1995, pp. 21–33.

[76] J.-S. Botero-Valencia, J. Valencia-Aguirre, D. Durmus, and W. Davis, "Multi-channel low-cost light spectrum measurement using a multilayer perceptron," *Energy Buildings*, vol. 199, pp. 579–587, Sep. 2019.

[77] A. Jacobs, "High dynamic range imaging and its application in building research," *Adv. Building Energy Res.*, vol. 1, no. 1, pp. 177–202, Jan. 2007.

[78] J. M. Wolfe, "Visual search in continuous, naturalistic stimuli," *Vis. Res.*, vol. 34, no. 9, pp. 1187–1195, May 1994.

[79] D. Vázquez, A. A. Fernández-Balbuena, H. Canabal, C. Muro, D. Durmus, W. Davis, A. Benítez, and S. Mayorga, "Energy optimization of a light projection system for buildings that virtually restores artworks," *Digit. Appl. Archaeology Cultural Heritage*, vol. 16, Mar. 2020, Art. no. e00128.

[80] D. Durmus and W. Davis, "Optimising light source spectrum for object reflectance," *Opt. Express*, vol. 23, no. 11, p. A456, 2015.

DORUKALP DURMUS received the B.S. degree in electrical engineering from Istanbul Technical University, in 2006, the M.S. degree in architectural lighting design from the KTH Royal Institute of Technology, in 2009, and the Ph.D. degree in architectural sciences from The University of Sydney, in 2018.

He completed his postdoctoral training with the Advanced Lighting Team, Pacific Northwest National Laboratory, USA. Since 2020, he has been an Assistant Professor with the Department of Architectural Engineering, Pennsylvania State University, PA, USA. His research interests include the spectral optimization of light sources, colorimetry, visual perception, intelligent lighting systems, and energy efficiency.

Dr. Durmus is a Fellow of the Higher Education Academy (FHEA). He was a recipient of the HJ Cowan Architectural Science Scholarship. He received the Australian Postgraduate Award (APA).

• • •