Amsterdam University Press

COMPUTATIONAL COMMUNICATION RESEARCH 4.1 (2022) 323-349
HTTPS://DOI.ORG/10.5117/CCR2022.1.009.PENG

ARTICLE

Athec

A Python Library for Computational Aesthetic Analysis of Visual Media in Social Science Research

Yilang Peng
University of Georgia
yilang.peng@uga.edu

Abstract
Visual aesthetics are related to a broad range of communication and psychological outcomes, yet the tools of computational aesthetic analysis are not widely available in the social science community. In this article, I address this gap and provide a tutorial on measuring hand-crafted aesthetic attributes, such as colorfulness and visual complexity. I introduce Athec, a Python library for computational aesthetic analysis in social science research. Furthermore, a case study applies Athec to compare the visual aesthetics of Instagram posts from the two candidates in the 2016 U.S. presidential election, Hillary Clinton and Donald Trump, indicating how amateurishness and authenticity are reflected in politicians' visual messages. With computational aesthetic analysis tools, communication researchers can better understand the antecedents and outcomes of visual aesthetics beyond visual media content.

Keywords: computational aesthetics, computer vision, aesthetic analysis, image feature, visual complexity, authenticity

Introduction

Visual analysis in social science research considers both the content and aesthetics of visual media. Prior scholarship has demonstrated the theoretical significance of visual aesthetics, indicating that aesthetic attributes are associated with a broad range of communication and psychological outcomes, such as viewers' emotional responses to visual stimuli, images' aesthetic appeal, social media content's popularity, and visual messages' persuasiveness (Bakhshi & Gilbert, 2015; Labrecque & Milne, 2012; Lazard & Mackert, 2014; Matz et al., 2019; Morgan et al., 2003; Peng, 2017; Sutton et al., 2019; Valdez & Mehrabian, 1994). Meanwhile, the rapid development of computer vision has given social scientists techniques of analyzing large-scale visual content, for example, identifying protests in social media images or politicians' emotional displays (Joo et al., 2019; Peng, 2018). A few tutorials targeted at social scientists have provided detailed documentation on the implementation of computer vision tools in analyzing content themes (Joo & Steinert-Threlkeld, 2018; Williams et al., 2020). Still, in addition to recognizing images' content, computer vision tools can also help scholars investigate visual media aesthetics (Dhar et al., 2011; Ke et al., 2006).

This tutorial article aims to address this gap by providing an overview of computational aesthetic analysis tools that could benefit visual analysis in social science research. This paper focuses on introducing hand-crafted attributes that can be measured using computer vision tools, such as colorfulness and visual complexity (Liu et al., 2016; Peng & Jemmott, 2018). These attributes do not reflect images' content directly, but they likely impact viewers' reactions to images. These visual attributes are linked to a variety of psychological and communication variables, for example, social media users' emotional states and personality, images' aesthetic appeal and online virality, and websites and app interfaces' attractiveness (Dhar et al., 2011; Machajdik & Hanbury, 2010; Matz et al., 2019; Miniukovich & De Angeli, 2016; Peng & Jemmott, 2018; Reinecke et al., 2013).

Furthermore, this article introduces Athec (a shortened form of aesthetic), a Python library for conducting computational aesthetic analysis of visual media.[1] Some examples of popular computer vision libraries include OpenCV and scikit-image (Bradski, 2000; Van der Walt et al., 2014), but these libraries tend to focus on performing computer vision tasks and are not dedicated to measuring image aesthetics per se. Furthermore, while a growing body of scholarship has used computational aesthetic analysis, the codes to measure aesthetics may not have been made publicly available or are published in various sources or programming languages. This paper contributes to this body of research by preparing a toolkit written in Python, a programming language popular among computational social scientists. With a case study that compares the visual aesthetics of Instagram posts published by Hillary Clinton and Donald Trump, the two candidates in the 2016 U.S. presidential election, I further demonstrate how scholars can investigate visual aesthetics to answer meaningful social science questions.

Computational Aesthetic Analysis in Social Science Research

The aesthetics of visual media have attracted scholarly attention from multiple fields, including psychology, communication, and computer science (Bo et al., 2018; Brachmann & Redies, 2017). Computational aesthetics, a branch of computer vision, has been described as "the research of computational methods that can make applicable aesthetic decisions in a similar fashion as humans can" (Hoenig, 2005, p. 16). One major task in computational aesthetics is "to simulate the human visual system and brain to measure and quantify aesthetics" (Bo et al., 2018, p. 2). Below, I provide an overview of several emerging directions to demonstrate computational aesthetics tools' utility in social science research.

One line of research has investigated how visual aesthetics influence various kinds of digital behavioral data related to visual media (Miniukovich & De Angeli, 2016; Totti et al., 2014). Extensive research has examined visual content's virality on social media platforms (Bakhshi & Gilbert, 2015; Deza & Parikh, 2015; Khosla et al., 2014; Totti et al., 2014). For example, Bakhshi and Gilbert (2015) analyzed Pinterest images' popularity and demonstrated that the percentages of red, purple, and pink were related positively to virality, whereas green, blue, black, and yellow had negative associations. Other kinds of digital behaviors also have been examined. Miniukovich and De Angeli (2016) extracted computational measures of the visual saliency and complexity of app icons in Google's app store and linked them to installs and ratings. Such findings should be particularly relevant to communication research, as they provided theoretical insights into what makes messages or innovations attract attention and propagate, as well as practical implications for crafting viral communication campaigns.

Certain psychological outcomes also are associated with visual aesthetics. Visual media's aesthetic appeal is one major outcome of interest in the field of computational aesthetics. Scholars have used various approaches to examine what factors contribute to images' visual appeal, including photos, artwork, and webpages (Bo et al., 2018; Datta et al., 2006; Matz et al., 2019). Matz et al. (2019) examined how a list of computationally calculated features predicted marketing images' general appeal, as well as appeal to people with different personality traits. Sentiments or emotions associated with images are also an important area of investigation (Machajdik & Hanbury, 2010). For example, Lu et al. (2012) investigated how different shapes in images, such as lines and ellipses, predicted viewers' emotional responses. Scholars also have investigated other psychological outcomes associated with visual media, including complexity, memorability, interestingness, and novelty (Constantin et al., 2019; Isola et al., 2011). These bodies of inquiry potentially could help us better understand visual media's psychological effects, particularly regarding specific visual aesthetics' role.

Furthermore, extant research has investigated how visual media's aesthetic features are associated with people and organizations who produce or share them. One line of research has examined how visual aesthetics reflect content producers' communicative intent and bias. Messing et al. (2016) found that negative political campaigns often feature darkened images of political candidates and that such an aesthetic style can activate negative stereotypes. In another body of scholarship, researchers collected social media users' self-reported traits, along with their social media posts, and found that the aesthetics of visual media that users posted predicted their characteristics, such as gender, personality, and depression status (Guntuku et al., 2019; Kim & Kim, 2018; Liu et al., 2016; Reece & Danforth, 2017). Reece and Danforth's (2017) analysis of Instagram users found that depressed users tended to post darker, less saturated, and bluer images than users with healthy mental states. These studies may help us better understand how and why users share visual media, which potentially could advance computer-mediated communication theories.

Approaches to Computational Aesthetic Analysis

Prior research has developed at least three major computational approaches to aesthetic analysis (Brachmann & Redies, 2017). One involves developing hand-crafted visual features that capture visual attributes that should relate to aesthetics, which typically are proposed based on art theories, psychology, and photographic principles (Brachmann & Redies, 2017; Datta et al., 2006; Ke et al., 2006; Machajdik & Hanbury, 2010; Miniukovich & De Angeli, 2016; Zhao et al., 2014). These features range from basic ones – such as brightness and saturation, which can be calculated easily by averaging pixel values – to more complicated attributes, such as the rule of thirds and visual complexity, which require more complex calculations and image processing techniques. These hand-crafted features then are used to predict image outcomes, for example, aesthetic appeal or sentiment (Datta et al., 2006; Machajdik & Hanbury, 2010).

Alternatively, researchers can use generic image features, such as scale-invariant feature transform (SIFT) and speeded-up robust features (SURF), to predict image outcomes. In image processing, a feature carries a piece of information about an image region (Umbaugh, 2010). Features can be visual structures – for example, points, corners, edges, and shapes – or an array of numerical values from a feature detection algorithm. These generic features often are not designed to predict aesthetics, but rather to fulfill other computer vision tasks, such as image matching and object classification (Brachmann & Redies, 2017; Lowe, 1999; Umbaugh, 2010).[2] These features can be used to predict image outcomes, such as aesthetic appeal, sentiment, and higher-level aesthetic attributes, such as the depth of field or the rule of thirds (Dhar et al., 2011; Isola et al., 2011; Mai et al., 2011).

Scholars also have applied deep learning models, such as Convolutional Neural Networks (CNN), to predict image aesthetics (Talebi & Milanfar, 2018). Previous research has used deep learning models to recognize and classify content themes in visual media (Joo et al., 2019; Joo & Steinert-Threlkeld, 2018; Peng, 2021; Zhang & Peng, forthcoming). Similarly, researchers can use deep learning models to predict visual media's aesthetic appeal (Talebi & Milanfar, 2018) or high-level aesthetic features, such as color harmony and the rule of thirds (Kong et al., 2016). Scholars can also adopt a transfer learning approach, for example, extracting features with a pre-trained model and incorporating them into a prediction model (Ha et al., 2020; Khosla et al., 2014).

Each approach has advantages and limitations. One advantage of using hand-crafted features is that they generally are easier to interpret, so researchers can determine which visual characteristics predict image outcomes (Brachmann & Redies, 2017; Zhao et al., 2014). In addition, communication scholars are particularly interested in what message properties contribute to communication outcomes. Hand-crafted features can be calculated in a standardized way and serve as measures of message properties, and their effects can be compared across different studies. For example, visual complexity has been conceptualized as a message property that influences many psychological outcomes and also can be captured easily using computer algorithms (King et al., 2020; Pieters et al., 2010).

By comparison, researchers often achieve better prediction performance by using generic features or deep learning models, although it is often not easy to interpret the factors that contribute to predictions (Brachmann & Redies, 2017). Computer vision scholars are continually developing methods to capture images' intermediate features and representations that explain predictions in deep learning (Joo & Steinert-Threlkeld, 2018). For example, scholars can use activation maps to locate image regions that are important in making a CNN categorize an image (Joo & Steinert-Threlkeld, 2018). Still, at this stage, the interpretability of generic image features and deep learning techniques might be relatively limited (Brachmann & Redies, 2017).

Table 1 Description of Visual Attributes in Athec

Attribute	#	Description
Size	7	File size, width, height, aspect ratio, diagonal length, image size, and file size divided by image size.
Color model statistics	30	Summary statistics (i.e., mean, median, standard deviation, minimum, maximum, the first/third quartile, skewness, kurtosis, and entropy) of the RGB channels.
	32	Summary statistics of the HSV channels. For H, circular mean and circular standard deviation also are included.
	32	Summary statistics of the HSL channels. For H, circular mean and circular standard deviation also are included.
	30	Summary statistics of the XYZ channels.
	30	Summary statistics of the *L*a*b* channels.
	10	Summary statistics of the grayscale channel.
Contrast	1	A range that covers a certain percentage of the brightness histogram (Ke et al., 2006).
	2	The number of peaks detected on the brightness histogram and the largest gap between peaks.
Colorfulness	1	Colorfulness formula in Hasler and Suesstrunk (2003).
	1	Colorfulness based on the distance between two color distributions (Datta et al., 2006).
Color percentages	11	The percentages of pixels categorized as red, orange, yellow, green, blue, pink, purple, black, white, gray, and brown (Van De Weijer et al., 2007).
Color variety	2	Shannon index and Simpson index of the percentages of categorized colors, excluding black, gray, and white.
	1	Hue count formula in Ke et al. (2006).
Visual complexity	1	The percentage of image area occupied by edges after edge detection (Matz et al., 2019).
	1	The average distance among all the pairs of edge points (Peng & Jemmott, 2018).
	1	The relative size of a minimal bounding box that contains a certain percentage of edge points (Ke et al., 2006).
	1	The number of segments after image segmentation (Machajdik & Hanbury, 2010; Matz et al., 2019).
	N	The relative sizes of the *N* largest segments.
	1	Total saliency values after saliency detection, normalized by image size.
	1	The number of image blocks that add to a certain percentage of saliency values after the image is partitioned into $n \times n$ blocks (Mai et al., 2012).
	1	The relative size of a minimal bounding box that contains a certain percentage of saliency values (Mai et al., 2011).
	1	The consistency between two saliency maps, measured as the number of image blocks that are both selected as "saliency blocks" in the two images (Mai et al., 2012).

Rule of thirds	10	The coordinate of the center of the mass (CoM) based on a saliency map and its distances to the four thirds lines and the four intersections.
	8	The percentages of saliency values inside the four strips around the four thirds lines and the four rectangles that cover the four intersections (Mai et al., 2012).
Visual balance	2	The distances from the CoM based on a saliency map to the middle vertical and horizontal lines.
Sharpness	1	Standard deviation of the Laplacian (Liu et al., 2016).
	1	Sharpness based on the number of frequencies in the Fourier transform domain that pass a threshold (Ke et al., 2006).
	1	Standard deviation of maximum local variations (Bahrami & Kot, 2014).
Depth of field	1	The average sharpness measure of the four inner blocks divided by the average sharpness of all the blocks after an image is partitioned into 4 × 4 blocks (Datta et al., 2006).
Line dynamics	4	The number of total lines detected and the numbers of lines categorized as horizontal, vertical, and slanting lines (Ke et al., 2006; Machajdik & Hanbury, 2010).

Athec: A Python Library for Computational Aesthetic Analysis

To facilitate the computational analysis of visual aesthetics among social scientists, I introduce Athec, a Python library for computational aesthetic analysis that focuses on measuring pre-defined, hand-crafted features. Table 1 presents all the visual attributes covered in this tutorial.

Size

An image can be represented as a $w \times h$ matrix of pixels, with w and h being the image's width and height. Pixels are the image's building blocks, and they store information about color in different parts of the image (Rosebrock, 2017). Athec provides the image's width, height, size, aspect ratio, diagonal length, and file size. File size can indicate visual complexity (Pieters et al., 2010).

Color Model Statistics

Athec provides various statistics (i.e., mean, median, standard deviation, minimum, maximum, the first/third quartile, skewness, kurtosis, and entropy) to summarize each channel in common color models, including grayscale, RGB, HSV, HSL, $L^*a^*b^*$, and XYZ. In a grayscale image, each pixel is a single number that represents different shades of gray, typically ranging from 0 (black) to 255 (white) (Figure 1a) (Rosebrock, 2017). In a color image, each pixel represents color as a tuple of values in a color model, typically three numbers. For example, the RGB color model uses different combinations of red (R), green (G), and blue (B) to reproduce a broad range of colors. Each pixel uses three numbers that range from 0 to 255 to indicate red, green, and blue (Figure 1b).

Alternatively, color can be described as a combination of three properties: hue, lightness, and chroma (Fairchild, 2013). Hue represents the specific color tone (e.g., red). Lightness or brightness describes the color's appearance in terms of how much light is emitted. Chroma, saturation, or colorfulness describes color intensity or vividness (Fairchild, 2013).[3] All three properties have been found to influence communication outcomes and psychological functions, such as emotional responses, cognitive tasks, impression formation, and reactions to visual messages, such as ads and food packaging (Elliot & Maier, 2014; Labrecque & Milne, 2012; Valdez & Mehrabian, 1994; Vermeir & Roose, 2020).

The HSV (Figure 1d) or HSL (Figure 1e) color model uses hue (H), saturation (S), and value (the V in HSV) or lightness (the L in HSL) to represent each color. In the HSV or HSL color model, hue is represented computationally as a circular continuum that starts with red and orange, then gradually transitions to yellow, green, and blue, and ends with purple and red (Figure 1c).

Athec includes a few other frequently used color models, including XYZ and $L^*a^*b^*$. In the XYZ color model, the Y dimension was designed to imitate human perceptions of luminance (Ibraheem et al., 2012). In the $L^*a^*b^*$ color model, L^* represents perceptual lightness, a^* represents the green–red continuum, and b^* represents the blue–yellow continuum (Ibraheem et al., 2012) (Figure 1f). In social science research, the $L^*a^*b^*$ color model is frequently used to quantify skin tone, with L^*, a^*, and b^* indicating skin lightness, redness, and yellowness, respectively (Stephen et al., 2009).

Brightness

Athec transforms an image into grayscale (Figure 2b). In a grayscale image, the average brightness value can measure brightness. Other channels in different color models also reflect brightness, for example, V in HSV, L in HSL, and Y in XYZ.

Contrast

Contrast reflects the difference between light and dark. An image with high contrast has bright and dark regions that are visually distinct, whereas

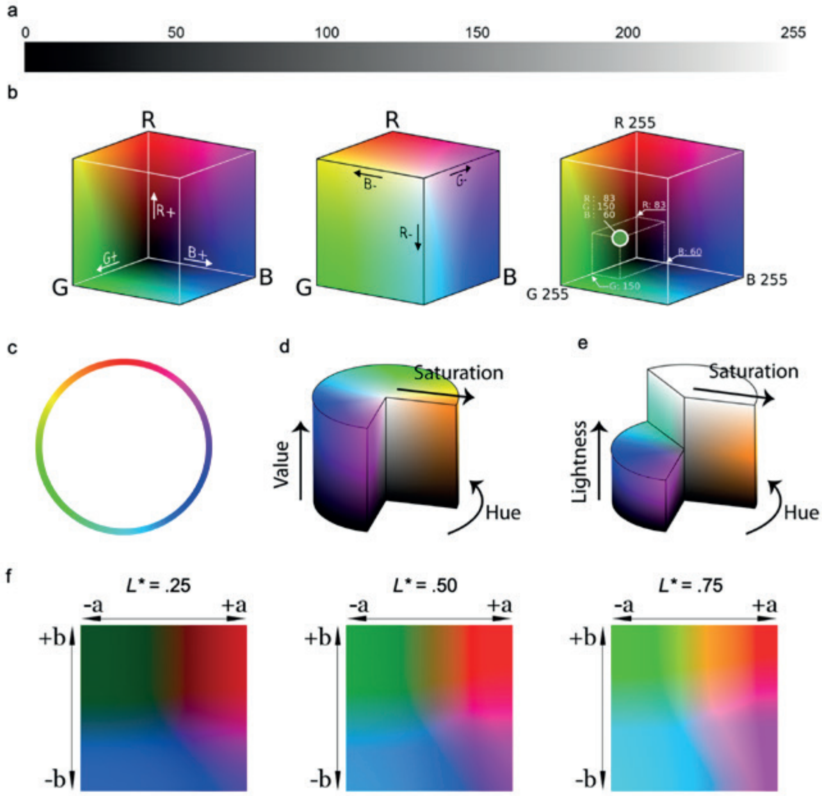

Figure 1 Illustrations of different color models: (a) grayscale; (b) RGB; (c) hue circle; (d) HSV; (e) HSL; (f) *L*a*b**. Credits: b from Horst Frank; d and e from SharkD; f from JakobVoss.

an image with low contrast has a narrow range of brightness and looks washed out. Athec converts the original image into grayscale and creates a brightness histogram that maps the percentages of pixels that fall into a particular brightness value. The range that covers a certain percentage of the histogram's mass can indicate contrast (Ke et al., 2006) (Figure 2c), as well as other measures of spread in brightness, such as standard deviation. Athec also detects peaks on a smoothed curve of the brightness histogram and provides the number of peaks and the distance between the highest and lowest peaks as indicators of contrast (Figure 2d).

Saturation and Colorfulness

The S value in the HSV color space is a measure of saturation. Other colorfulness measures have been designed to be more in line with

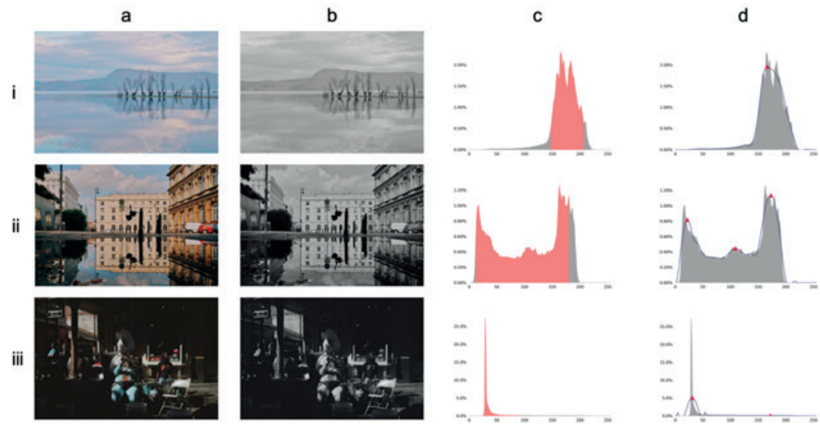

Figure 2 Measures of contrast: (a) original image; (b) grayscale; (c) brightness histogram, with the red area representing a range that contains at least 90% of the histogram; (d) peak detection on the brightness histogram, with red triangles representing the detected peaks on a smoothed curve of the histogram. An image with high contrast (ii, iii) should have a wider range, more peaks detected, or a larger distance between peaks in the brightness histogram than an image with low contrast (i).

human perceptions of colorfulness (Hasler & Suesstrunk, 2003). Athec includes the colorfulness formula introduced in Hasler and Suesstrunk (2003). This formula maps RGB values of an image onto an opponent color space and uses pixel distribution to calculate colorfulness. The second measure of colorfulness divides the RGB color space into 64 bins (four bins per channel) and computes dissimilarity (i.e., using the Earth Mover's Distance) between the original image and a hypothetical monochrome image in which the pixels are distributed evenly in the 64 bins (Datta et al., 2006).

Specific Colors

Different colors can elicit different emotional reactions, cognitive associations, and cultural meanings (Elliot & Maier, 2014; Labrecque & Milne, 2012; Valdez & Mehrabian, 1994). Athec calculates the percentages of 11 basic colors in an image: red, orange, yellow, green, blue, pink, purple, brown, black, white, and gray. This algorithm relies on Van De Weijer et al.'s (2007) data set, which divides the RGB color model into 32,768 bins (32 bins per color channel) and assigns each bin to one of the 11 colors (Figure 3b).

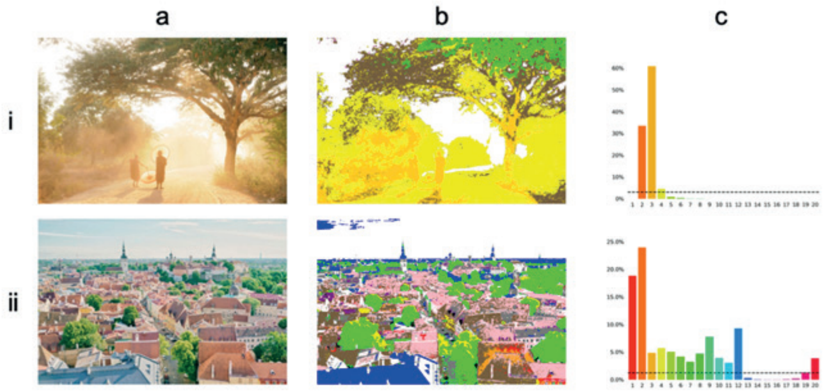

Figure 3 Measures of color-related attributes: (a) original image; (b) color categorization; (c) hue count histogram, with perceptually colorful pixels assigned into 20 hue bins and the black line denoting the threshold. An image with high color variety (ii) should have more diverse color categories and more hue bins above the threshold than an image with low color variety (i).

Color Variety

Color variety reflects the diversity or variation of colors (specifically hues) in an image. Athec provides two methods. The first is based on color categorization, as an image of high color variety should have more evenly distributed color categories (Zhao et al., 2014) (Figure 3b). This method excludes black, white, and gray and calculates the variation in the percentages of the remaining eight colors using two diversity measures: the Shannon index and the Simpson index. The second method uses Ke et al.'s (2006) hue count formula. This method first selects perceptually colorful pixels that are somewhat saturated and not too dark or bright (e.g., $S > 0.2$ and $0.15 < V < 0.95$ in the HSV color model), then sorts these pixels into 20 hue bins, and counts the number of bins with enough pixels that pass a certain threshold (Ke et al., 2006) (Figure 3c).

Visual Complexity

Visual complexity reflects the amount of variation in a visual stimulus (Pieters et al., 2010). Visual complexity is an important variable in persuasion and visual media effects because it influences a broad range of psychological outcomes, such as processing fluency, perceived informativeness, and emotional responses (King et al., 2020; Pieters et al., 2010). A JPEG image's file size can indicate visual complexity (Pieters et al., 2010). Athec also provides the following visual complexity measures, which are based on various image processing techniques.

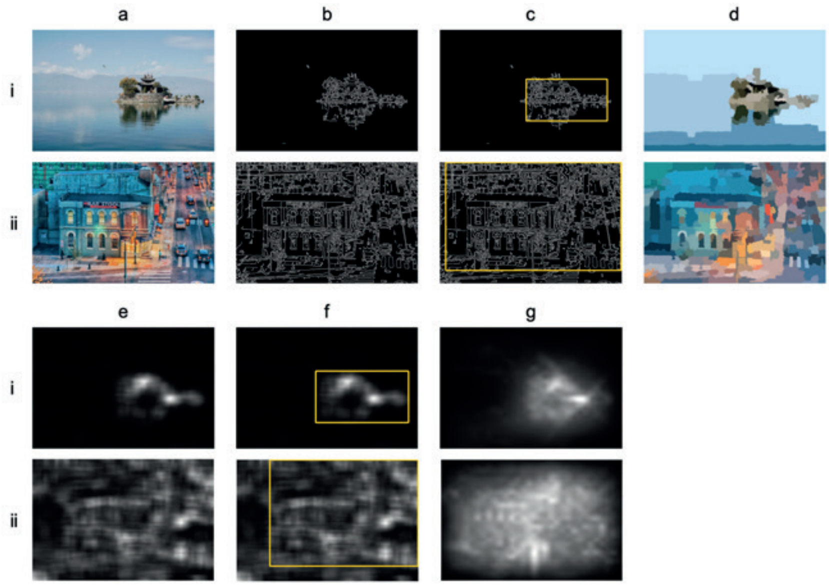

Figure 4 Measures of visual complexity: (a) original image; (b) edge detection; (c) a bounding box that contains at least 90% of the edges; (d) segmentation; (e) saliency detection using spectral residual; (f) a bounding box that contains at least 80% of the saliency values; (g) saliency detection using graph-based visual saliency. Compared with a visually simple image (i), a complex image (ii) has denser and more evenly distributed edges, larger bounding boxes, more and smaller segments, and less concentrated and less consistent saliency maps.

Edge Detection

Edge detection finds pixel points in an image in which certain image characteristics, such as brightness or color, dramatically change (Umbaugh, 2010). Edges usually represent objects or textures' contours (Umbaugh, 2010) (Figure 4b). A complex image that has many objects or perceptual details should have many edges, so the percentage of image area occupied by edge points can be a measure of feature complexity (Peng & Jemmott, 2018). Furthermore, a simple image with a clean background should have edges that are concentrated in a small area of an image, whereas a complex image should have edges that spread across the whole frame, far away from each other. Therefore, the average distance among edge points and the size of a bounding box that contains a certain percentage of the edge points can quantify visual complexity (Ke et al., 2006; Peng & Jemmott, 2018) (Figure 4c).

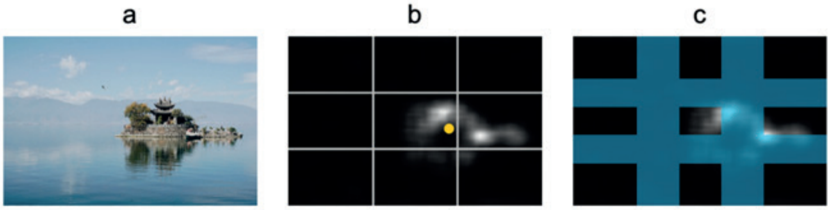

Figure 5 Measures of the rule of thirds: (a) original image; (b) the four thirds lines and the center of mass calculated from saliency; and (c) the four strips and four intersections. An image following the rule of thirds should have its center of mass closer to the four lines/four intersections and have a higher percentage of saliency values in the four strips/rectangles.

Segmentation

Image segmentation is a computer vision task that partitions pixels in an image into segments so that pixels in the same segment share similar characteristics. Segments roughly correspond to objects in the image (Datta et al., 2006) (Figure 4d). The number of segments should reflect visual complexity (Machajdik & Hanbury, 2010; Peng & Jemmott, 2018). Furthermore, an image with low complexity often has a clean, uniform background that should be partitioned into one large segment (Mai et al., 2012). Athec provides the sizes of the *N* largest segments, so one can use the size of the largest segment or the sum of the top *N* segments to capture visual complexity.

Saliency Detection

Saliency detection is a computer vision task that identifies salient objects that stand out from the image. A saliency map can be represented as a grayscale image in which each pixel's brightness stores the saliency value (Figure 4e, 4g). One measure of visual complexity is to find a minimal bounding box that contains at least a certain percentage of the total saliency value (Figure 4f). Another method for measuring complexity uses the compactness of the subject of interest (Mai et al., 2012), which divides the saliency map into $n \times n$ blocks and finds the smallest number of blocks that add to a certain percentage of the image's total saliency (Mai et al., 2012). Finally, visual complexity is reflected in the consistency in results from different saliency detection methods (Mai et al., 2012). For a compositionally simple image that has subjects that stand out from the background, different saliency detection techniques should generate largely similar results (Figures 4e, 4g). This method defines the top image blocks (e.g., 60%) with the highest saliency values in a saliency map as “salient blocks.” The consistency between two saliency maps based on the same image can be measured as the percentage of overlapping salient blocks between them, which can then indicate visual complexity (Mai et al., 2012).

Rule of Thirds

According to the rule of thirds, a picture can be partitioned into nine equal parts using two horizontal and two vertical lines (thirds line). To create an aesthetically pleasing image, the main subject or the focus should be placed along the third lines or on their intersections (Amirshahi et al., 2014; Bo et al., 2018; Dhar et al., 2011; Mai et al., 2011) (Figure 5a). Athec provides several measures of the rule of thirds based on saliency maps (Mai et al., 2011). The first method finds the center of mass (CoM) of saliency values (Dhar et al., 2011; Hübner & Fillinger, 2016): Pixels with a saliency value can be viewed like point particles in physics that carry a ‘visual’ weight, so we can calculate the CoM on a saliency map (Figure 5b). Athec calculates the distances from the CoM to the four thirds lines and the four intersections. The second method creates four strips (e.g., with 1/6 of the width/height of the image) around each thirds line. Athec measures the saliency values in the four strips and four rectangles at these strips’ intersections (Figure 5c). Finally, Athec also partitions the image into an $n \times n$ grid and measures the average saliency value in each block (Mai et al., 2011). Although these raw saliency values do not directly reflect the rule of thirds, they can be used to build a prediction model of an image’s composition (Mai et al., 2011).

Visual Balance

Visual balance also refers to the arrangement of visual elements in an image. If an image only has one subject placed at the center, the image is balanced perfectly (Hübner & Fillinger, 2016). If the image contains multiple elements, the image is balanced if these elements’ perceptual weights compensate for each other (Hübner & Fillinger, 2016). To measure balance, Athec uses the CoM calculated from a saliency map, mentioned above, and measures its distances to the middle vertical and horizontal lines (Hübner & Fillinger, 2016; Thömmes & Hübner, 2018).

Sharpness and Depth of Field

Sharpness measures whether an image is blurry or sharp. The effective use of sharpness and blur often indicates technical quality and sophisticated compositional skill in photography (Datta et al., 2006; Ke et al., 2006).

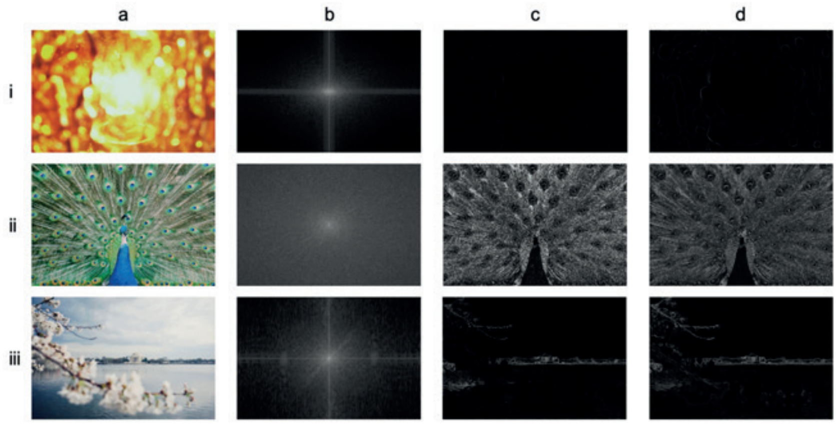

Figure 6 Measures of sharpness: (a) original image; (b) magnitude spectrum in the Fourier transform; (c) Laplacian; (d) maximal local variation. A sharp image (ii) has more high frequencies and more varied Laplacian and maximal local variation than a blurry image (i). An image with a low depth of field (iii) has both sharp and blurry regions.

Athec provides three measures of sharpness. The first method uses the Fast Fourier transform to convert the original image in the spatial domain to the Fourier domain. Figure 6b visualizes the magnitude spectrum, with the DC value (i.e., the image mean) shifted to the center (Fisher et al., 2003). This image is the same size as the original. Each point corresponds to a frequency in the original image, and its brightness denotes the corresponding frequency's magnitude (log-transformed). The further away a point is from the center, the higher its frequency. With the Fourier transform, high frequencies correspond to edges and fine details, and low frequencies correspond to smoother, more blurry regions in the original image (Fisher et al., 2003). As Figure 6b shows, blurry images generally have darker borders than sharp images, indicating that these images lack high frequencies. Athec measures sharpness as the number of frequencies with a magnitude above a certain threshold, divided by image size (Ke et al., 2006).

The second method uses the standard deviation of an image's Laplacian (Liu et al., 2016). The Laplacian highlights image areas that contain rapid intensity changes and is often used to detect edges (Fisher et al., 2003). The Laplacian of a sharp, in-focus should be more varied (Figure 6c).

The final method uses the standard deviation of the maximum local variation (MLV) of all pixels. For each pixel, MLV is the maximal intensity difference between the pixel and its eight neighboring pixels (Bahrami &

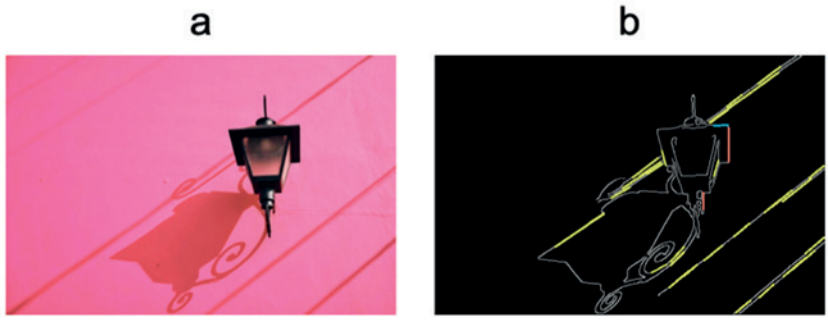

Figure 7 Measures of line dynamics: (a) original image; (b) line detection based on an edge map, with blue, red, and yellow lines representing horizontal, vertical, and slanting lines, respectively.

Kot, 2014) (Figure 6d). Sharp images generally have more spread-out MLV values (Bahrami & Kot, 2014).

In addition, Athec provides a measure of the depth of field. In an image with a shallow depth of field, some compositional elements are in sharp focus, whereas other objects, the background, or the foreground is blurred (Figure 6iii). Such an aesthetic can draw viewers' attention toward the subject of interest and away from irrelevant elements (Dhar et al., 2011). Following previous research (Datta et al., 2006), Athec divides an image into 4 × 4 blocks and measures the depth of field as the average sharpness estimates in the inner four blocks of the image, divided by the average sharpness of all the blocks.

Line Dynamics

The lines in an image can provoke emotional responses (Lu et al., 2012; Machajdik & Hanbury, 2010; Thömmes & Hübner, 2018). Athec uses the Hough transform to detect lines in an image, along with each line's length and orientation ($-90° < \theta < 90°$) (Figure 7c). These lines also can be grouped into horizontal (e.g., $-10° < \theta < 10°$), vertical (e.g., $\theta < 80°$ or $\theta > 80°$), and slanting lines (Lu et al., 2012; Machajdik & Hanbury, 2010).

Implementing Athec Measures

Preparing and Preprocessing Images

Athec is available on GitHub and pypi.org. Athec uses functions from other libraries, including Pillow, NumPy, OpenCV, scikit-image, Matplotlib, SciPy, and PyEMD. Images often come in a variety of formats and sizes. To facilitate image analysis in Athec, one needs to transform the images into either JPEG or PNG formats with three RGB channels. Several algorithms might take up a significant amount of computer power, so researchers can resize the images' dimensions to speed up calculations. The Python codes below demonstrate how one can resize an image so that its both sides do not exceed 400 pixels.

```
    import os
    from athec import misc

imgname = "example1.jpg"
    img_path = os.path.join("img original", imgname)
    resize_path = os.path.join("img resize", imgname)
    misc.tf_resize(img_path, resize_path, max_side = 400)
```

Calculating Visual Aesthetics

Researchers then calculate visual aesthetics using different Athec functions, which can take multiple inputs. Usually, Athec functions can use an image's file path as the input and returns the measures.

```
    import os
    from athec import misc, color

imgname = "example1.jpg"
    imgpath = os.path.join("img resize", imgname)
    result = color.attr_colorful(imgpath)
```

Alternatively, some Athec functions can take the image directly, represented as a NumPy array, as the input.

```
    imgarray = misc.read_img_rgb(imgpath)
    result = color.attr_colorful(imgarray)
```

For a few visual features that involve image processing techniques, Athec first transforms the image into an intermediate representation (e.g., an edge map) and uses this representation to calculate attributes. Furthermore, the interpretability of results is desirable in social science research. Along with the calculations, Athec can save intermediate representations of images as PNG files.

```
    import os
    from athec import misc, edge

imgname = "example1.jpg"
    imgpath = os.path.join("img resize", imgname)
    edgepath = os.path.join("img transform", "edge canny", imgname)
    edges = edge.tf_edge_canny(imgpath, edgepath)
    result = edge.attr_complexity_edge(edges)
```

Alternatively, for several visual aesthetics, Athec can read data directly from an intermediate representation and perform the calculations. This option is available because several different image processing algorithms are often available. Athec uses popular algorithms incorporated into existing computer vision libraries, such as Canny for edge detection. However, researchers can also use other image processing algorithms available elsewhere. In the following codes, researchers already obtained edge detection results using another method (e.g., holistically nested edge detection) and saved the edges as a PNG image. Researchers then extracted visual features with Athec directly from the edge image.

```
    import os
    from athec import misc, edge

imgname = "example1.jpg"
    edgepath = os.path.join("img transform", "edge HED", imgname + ".png")
    result = edge.attr_complexity_edge(edgepath)
```

Choosing Parameters for Athec Functions

Some Athec functions need researchers to choose parameters, and these decisions might affect final results. For example, for edge detection, researchers can decide whether they should use a Gaussian filter to remove noises from the image. Furthermore, the Canny edge detection algorithm also requires an upper and lower threshold. Athec uses a method based on Otsu thresholding and asks researchers to provide the ratio between the two thresholds. Figure 8 shows how the edge detection results vary by different parameters: As the size of the Gaussian filter increases, and as the ratio between thresholds increases, fewer edges are detected.

Under such circumstances, researchers may experiment with different sets of parameters to decide the optimal one by examining intermediate representations. Nevertheless, while these parameters likely change aesthetic

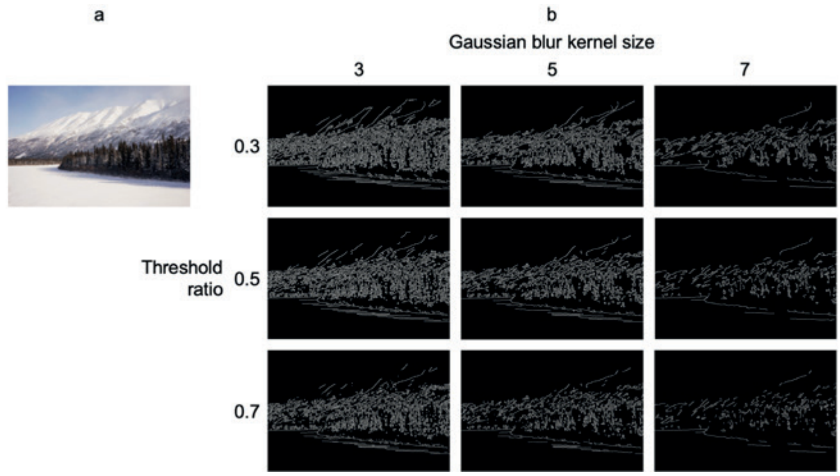

Figure 8 Edge detection results vary by the kernel size used for Gaussian blur and the ratio between the lower and upper thresholds.

measures' absolute values, these measures that use different parameters often are highly correlated with each other. Therefore, the variations might have relatively limited effects if researchers mostly are interested in the associations among aesthetic features and the links between aesthetic features and image outcomes. Researchers may also repeatedly run analyses with different parameters to ensure that the conclusions are robust across different specifications.

Case Study: Visual Aesthetics of Instagram Posts from Political Candidates

Next, this paper introduces a case study to demonstrate how scholars can apply computational aesthetic analysis to understand communication phenomena. Scholars have observed that populist politicians may publish media content with certain aesthetics to cultivate an image of authenticity and amateurishness, and to present themselves as "authentic outsiders" not tied to traditional politics (Enli, 2017). An analysis of linguistic styles in tweets by candidates Donald Trump and Hillary Clinton in the 2016 U.S. presidential election campaign indicated that Trump's campaign was less traditional and professional, and more amateurish than Clinton's (Enli, 2017). We should observe a similar pattern in the aesthetics of campaign visuals.

Table 2 Comparison of Visual Aesthetics between Clinton's and Trump's posts

Attribute	Clinton *M*	Trump *M*	*t*	df	*p*	Cohen's *d*
Brightness	112.03	87.26	9.51	825.56	<.001	0.66
Contrast	172.39	174.07	−0.58	922.26	0.56	−0.04
Colorfulness	49.76	50.91	−0.59	732.73	0.55	−0.04
Color variety	4.96	6.38	−7.80	854.38	<.001	−0.53
Visual complexity (file size)	0.40	0.47	−9.47	951.88	<.001	−0.61
(edge density)	0.10	0.13	−7.60	1001.44	<.001	−0.48
(edge distance)	0.32	0.34	−8.85	710.58	<.001	−0.66
(edge box size)	0.70	0.76	−8.23	710.12	<.001	−0.62
Sharpness	37.51	50.02	−12.84	902.08	<.001	−0.86
Depth of field	1.30	1.11	10.73	815.27	<.001	0.75

Unequal variances *t*-tests are conducted for all the attributes.

This case study compared the visual aesthetics of the two candidates' image posts in 2016. The data set contained 470 images from Clinton (@hillaryclinton) and 856 images from Trump (@realdonaldtrump). The analysis focused on a few visual attributes that might signal professional quality and aesthetic appeal, including brightness, contrast, colorfulness, visual complexity, and sharpness. With Athec, brightness and contrast were measured using average brightness in the grayscale image and the minimal range that covered 90% of the brightness histogram, respectively. Colorfulness and color variety were measured based on Hasler and Suesstrunk (2003) and Ke et al. (2006). Visual complexity was measured using the image's weighted file size, edge density, the average distance among edge points, and the size of a bounding box that contained 90% of the edges. Sharpness was measured as the standard deviation of Laplacian. Depth of field was measured as the inner four blocks' average sharpness divided by the average sharpness of all the blocks after an image was partitioned into 2 × 2 blocks.

Table 2 presents the results from multiple *t*-tests that compared the aesthetics of both politicians' posts. The images from Clinton's campaign generally were brighter, had fewer colors, and were visually simpler than

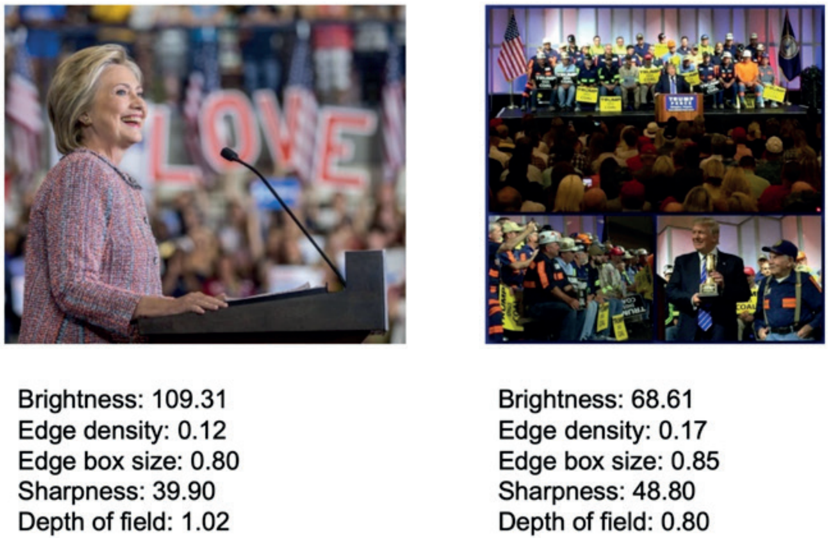

Figure 9 Examples from Clinton's and Trump's Instagram accounts

Trump's posts. Also, Clinton's posts used more blur and shallow depth of field than Trump's (Figure 9). The two politicians' posts did not differ significantly regarding contrast and colorfulness. In photography and graphic design, an aesthetic that uses simple composition, shallow depth of field, or a blurred background often can highlight the subject of interest in images and reflect sophisticated composition skills. The differences in visual aesthetics suggest that Clinton's posts might demonstrate a higher level of professional quality and aesthetic appeal, whereas Trump's visuals looked more casual and amateurish. Such findings echo previous observations that populist politicians often cultivate a sense of amateurishness to appear more authentic and to contrast themselves with traditional politicians. Together, this case study presented a scenario in which communication scholars can use Athec for computational aesthetic analysis and investigate meaningful patterns in visual media related to social science theories.

Future Directions

With computational aesthetic analysis tools, communication researchers can better understand visual aesthetics' antecedents and outcomes beyond visual media content. Such visual media can be diverse, such as photos, videos, ads, logos, data visualizations, websites, and app interfaces (e.g., Labrecque

& Milne, 2012; Lazard & Mackert, 2014; Matz et al., 2019; Miniukovich & De Angeli, 2016; Morgan et al., 2003; Murashka et al., 2021; Peng, 2021; Pieters et al., 2010). As noted earlier in the literature review, a growing body of scholarship is using computational aesthetic analysis to answer questions related to social science. Prior research has applied computational aesthetics to investigate both the production of visual messages and their effects, such as virality, aesthetic appeal, and emotion (Constantin et al., 2019; Dhar et al., 2011; Machajdik & Hanbury, 2010; Matz et al., 2019). Meanwhile, several less examined research directions remain. Visual aesthetics likely affect a broad range of other communication outcomes, including visual messages' perceived effectiveness, comprehensibility, informativeness, and sensation value (Lazard & Mackert, 2014; Morgan et al., 2003; Pieters et al., 2010; Sutton et al., 2019). Future research can extend these bodies of inquiry and apply computational methods to investigate how visual aesthetics affect various communication outcomes.

One future direction is to study the link between computationally coded visual attributes and human perceptions of these attributes. Some research has found correspondence between computational measures and human perceptions, or that computational measures can predict human judgment of certain visual aesthetics (Hasler & Suesstrunk, 2003; Pieters et al., 2010). How well computationally calculated aesthetic attributes reflect perceived aesthetics in different contexts and for different data sets should be a focus of future research. However, it is notable that computational measures do not need to match human perceptions perfectly to be useful. They can serve as intrinsic message properties while human perceptions serve as mediators between message properties and message outcomes (O'Keefe, 2003).

Finally, although Athec covers many popular aesthetic attributes in previous research, many measures still are not incorporated into the library. Athec will be available on GitHub, and others can contribute to the library by adding new measures of aesthetic attributes, which potentially will be incorporated into future versions.

Notes

1. The package is available at https://github.com/yilangpeng/Athec
2. Some studies refer to these generic features as "hand-crafted features" as opposed to features learned from deep learning models (Nanni et al., 2017). Nevertheless, this paper follows Brachmann and Redies (2017) and only refers to hand-crafted features as attributes designed to reflect image

aesthetics or other aesthetics-related properties such as sentiment and interestingness.

3. There are some differences between brightness and lightness, and between chroma, colorfulness, and saturation, which are not covered in this tutorial but can be found in other sources (Fairchild, 2013).

References

Amirshahi, S. A., Hayn-Leichsenring, G. U., Denzler, J., & Redies, C. (2014). Evaluatingthe rule of thirds in photographs and paintings. *Art & Perception*, 2(1–2), 163–182.

Bahrami, K., & Kot, A. C. (2014). A fast approach for no-reference image sharpnessassessment based on maximum local variation. *IEEE Signal Processing Letters*, 21(6), 751–755.

Bakhshi, S., & Gilbert, E. (2015). Red, purple and pink: The colors of diffusion on pinterest. *PloS one*, 10(2), e0117148.

Bo, Y., Yu, J., & Zhang, K. (2018). Computational aesthetics and applications. *Visual Computing for Industry, Biomedicine, and Art*, 1(1), 6.

Brachmann, A., & Redies, C. (2017). Computational and experimental approaches to visual aesthetics. *Frontiers in Computational Neuroscience*, 11, 102.

Bradski, G. (2000). The OpenCV library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11), 120–123.

Constantin, M. G., Redi, M., Zen, G., & Ionescu, B. (2019). Computational understanding of visual interestingness beyond semantics: Literature survey and analysis of covariates. *ACM Computing Surveys*, 52(2), 1–37.

Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. In *European conference on computer vision* (pp. 288–301).

Deza, A., & Parikh, D. (2015). Understanding image virality. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1818–1826).

Dhar, S., Ordonez, V., & Berg, T. L. (2011). High level describable attributes for predicting aesthetics and interestingness. In *IEEE computer society conference on computer vision and pattern recognition* (pp. 1657–1664).

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, *65*, 95–120.

Enli, G. (2017). Twitter as arena for the authentic outsider: Exploring the social media campaigns of trump and clinton in the 2016 us presidential election. *European Journal of Communication*, 32(1), 50–61.

Fairchild, M. D. (2013). *Color appearance models*. John Wiley & Sons.

Fisher, R., Perkins, S., Walker, A., & Wolfart, E. (2003). *The hypermedia image processing reference.* https://homepages.inf.ed.ac.uk/rbf/HIPR2/hipr_top.htm

Guntuku, S. C., Buffone, A., Jaidka, K., Eichstaedt, J. C., & Ungar, L. H. (2019). Understanding and measuring psychological stress using social media. In *Proceedings of the international AAAI conference on web and social media* (Vol. 13, pp. 214–225).

Ha, Y., Park, K., Kim, S. J., Joo, J., & Cha, M. (2020). Automatically detecting image–text mismatch on Instagram with deep learning. *Journal of Advertising*, 50(1), 52–62.

Hasler, D., & Suesstrunk, S. E. (2003, June). Measuring colorfulness in natural images. In *Human vision and electronic imaging VIII* (Vol. 5007, pp. 87–95). International Society for Optics and Photonics.

Hoenig, F. (2005). Defining computational aesthetics. *Computational Aesthetics in Graphics, Visualization and Imaging*, 13–18.

Hübner, R., & Fillinger, M. G. (2016). Comparison of objective measures for predicting perceptual balance and visual aesthetic preference. *Frontiers in Psychology*, 7, 335.

Ibraheem, N. A., Hasan, M. M., Khan, R. Z., & Mishra, P. K. (2012). Understanding color models: A review. *ARPN Journal of Science and Technology*, 2(3), 265–275.

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? In *Proceedings of the 2011 IEEE conference on computer vision and pattern* (pp. 145–152).

Joo, J., Bucy, E. P., & Seidel, C. (2019). Automated coding of televised leader displays: Detecting nonverbal political behavior with computer vision and deep learning. *International Journal of Communication*, 13, 4044–4066.

Joo, J., & Steinert-Threlkeld, Z. C. (2018). *Image as data: Automated visual content analysis for political science*.https://arxiv.org/abs/1810.01544

Ke, Y., Tang, X., & Jing, F. (2006). The design of high-level features for photo quality assessment. In *2006 IEEE computer society conference on computer vision and pattern recognition* (Vol. 1, pp. 419–426).

Khosla, A., Das Sarma, A., & Hamid, R. (2014). What makes an image popular? In *Proceedings of the 23rd international conference on world wide web* (pp. 867–876).

Kim, Y., & Kim, J. H. (2018). Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos: An exploratorystudy. *Information Processing & Management*, 54(6), 1101–1114.

King, A. J., Lazard, A. J., & White, S. R. (2020). The influence of visual complexity on initial user impressions: testing the persuasive model of web design. *Behaviour & Information Technology*, 39(5), 497–510.

Kong, S., Shen, X., Lin, Z., Mech, R., & Fowlkes, C. (2016). Photo aesthetics ranking network with attributes and content adaptation. In *European conference on computer vision* (pp. 662–679).

Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *Journal of the Academy of Marketing Science*, 40(5), 711–727.

Lazard, A., & Mackert, M. (2014). User evaluations of design complexity: The impact of visual perceptions for effective online health communication. *International Journal of Medical Informatics*, 83(10), 726–735.

Liu, L., Preotiuc-Pietro, D., Samani, Z. R., Moghaddam, M. E., & Ungar, L. H. (2016). Analyzing personality through social media profile picture choice. In *Tenth international AAAI conference on web and social media.*(pp. 211–220).

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150–1157).

Lu, X., Suryanarayan, P., Adams Jr, R. B., Li, J., Newman, M. G., & Wang, J. Z. (2012). On shape and the computability of emotions. In *Proceedings of the 20th ACM international conference on multimedia* (pp. 229–238).

Machajdik, J., & Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. In *Proceedings of the 18th ACM international conference on multimedia* (pp. 83–92).

Mai, L., Le, H., Niu, Y., Lai, Y.-C., & Liu, F. (2012). Detecting rule of simplicity from photos. In *Proceedings of the 20th ACM international conference on multimedia* (pp. 1149–1152).

Mai, L., Le, H., Niu, Y., & Liu, F. (2011). Rule of thirds detection from photograph. In *2011 IEEE international symposium on multimedia* (pp. 91–96).

Matz, S. C., Segalin, C., Stillwell, D., Müller, S. R., & Bos, M. W. (2019). Predicting the personal appeal of marketing images using computational methods. *Journal of Consumer Psychology*, 29(3), 370–390.

Messing, S., Jabon, M., & Plaut, E. (2016). Bias in the flesh: Skin complexion and stereotype consistency in political campaigns. *Public Opinion Quarterly*, 80(1), 44–65.

Miniukovich, A., & De Angeli, A. (2016). Pick me! getting noticed on google play. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 4622–4633).

Morgan, S. E., Palmgreen, P., Stephenson, M. T., Hoyle, R. H., & Lorch, E. P. (2003). Associations between message features and subjective evaluations of the sensation value of antidrug public service announcements. *Journal of Communication*, 53(3), 512–526.

Murashka, V., Liu, J., & Peng, Y. (2021). Fitspiration on Instagram: Identifying topic clusters in user comments to posts with objectification features. *Health Communication*, 36(12), 1537–1548.

Nanni, L., Ghidoni, S., & Brahnam, S. (2017). Handcrafted vs. non-handcrafted features for computer vision classification. *Pattern Recognition*, *71*, 158–172.

O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory*, 13(3), 251–274.

Peng, Y. (2017). Time travel with one click: Effects of digital filters on perceptions of photographs. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 6000–6011).

Peng, Y. (2018). Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision. *Journal of Communication*, 68(5), 920–941.

Peng, Y. (2021). What makes politicians' Instagram posts popular? Analyzing social media strategies of candidates and office holders with computer vision. *The International Journal of Press/Politics*, 26(1), 143–166.

Peng, Y., & Jemmott III, J. B. (2018). Feast for the eyes: Effects of food perceptions and computer vision features on food photo popularity. *International Journal of Communication*, 12.

Pieters, R., Wedel, M., & Batra, R. (2010). The stopping power of advertising: Measures and effects of visual complexity. *Journal of Marketing*, 74(5), 48–60.

Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6(1), 1–12.

Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013). Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2049–2058).

Rosebrock, A. (2017). *Deep learning for computer vision with python*. https://pyimagesearch.com/deep-learning-computer-vision-python-book/

Stephen, I. D., Smith, M. J. L., Stirrat, M. R., & Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *International Journal ofPrimatology*, 30(6), 845–857.

Sutton, J. A., Yang, S., & Cappella, J. N. (2019). Perceived effectiveness of objective features of pictorial warning messages. *Tobacco Control*, 28(e1), e24–e30.

Talebi, H., & Milanfar, P. (2018). NIMA: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8), 3998–4011.

Thömmes, K., & Hübner, R. (2018). Instagram likes for architectural photos can be predicted by quantitative balance measures and curvature. *Frontiers in Psychology*, 9, 1050.

Totti, L. C., Costa, F. A., Avila, S., Valle, E., Meira Jr, W., & Almeida, V. (2014). The impact of visual attributes on online image diffusion. In *Proceedings of the 2014 ACM conference on web science* (pp. 42–51).

Umbaugh, S. E. (2010). *Digital image processing and analysis: human and computer vision applications with CVIPtools*. CRC press.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4), 394–409.

Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., … Yu, T. (2014). scikit-image: image processing in python. *PeerJ*, 2, e453.

Van De Weijer, J., Schmid, C., & Verbeek, J. (2007). Learning color names from real-world images. In *2007 IEEE conference on computer vision and pattern recognition* (pp. 1–8).

Vermeir, I., & Roose, G. (2020). Visual design cues impacting food choice: A review and future research agenda. *Foods*, 9(10), 1495.

Williams, N. W., Casas, A., & Wilkerson, J. D. (2020). Images as data for social science research: An introduction to convolutional neural nets for image classification. Cambridge University Press.

Zhang, H., & Peng, Y. (Forthcoming). Image clustering: An unsupervised approach to categorize visual data in social science research. *Sociological Methods & Research*.

Zhao, S., Gao, Y., Jiang, X., Yao, H., Chua, T.-S., & Sun, X. (2014). Exploring principles-of-art features for image emotion recognition. In *Proceedings of the 22nd ACM international conference on multimedia* (pp. 47–56).