EnviroPulse: Providing Feedback about the Expected Affective Valence of the Environment

Deltcho Valtchanov, Mark Hancock
University of Waterloo
Waterloo, Ontario, Canada
{dvaltchanov,mark.hancock}@uwaterloo.ca

ABSTRACT
Interacting with nature is beneficial to a person's mental-state, but it can sometimes be difficult to find environments that will induce positive affect (e.g., when planning a run). In this paper, we describe *EnviroPulse*—a system for automatically determining and communicating the expected affective valence (EAV) of environments to individuals. We describe a prototype that allows this to be used in real-time on a smartphone, but *EnviroPulse* could easily be incorporated into GPS systems, mapping services, or image-based systems. Our work differs from existing work in affective computing in that, rather than detecting a user's affect directly, we automatically determine the EAV of the environment through visual analysis. We present results that suggest our system can determine the EAV of environments. We also introduce real-time affective visual feedback of the calculated EAV of images, and present results from an informal study suggesting that real-time visual feedback can be used for induction of affect.

Author Keywords
Human Factors; Affective Computing; Mobile Interfaces

INTRODUCTION
Over the past forty years, psychologists have demonstrated strong and consistent evidence that interacting with nature is highly important for psychological and physical well-being [10,26]. This principle, known as *biophilia,* is based on empirical evidence demonstrating that interactions with the natural world improve cognitive abilities (e.g., the ability to sustain attention, vigilance, etc.) and reduce stress and its associated negative effects (e.g., high blood pressure, increased stress hormones, and diminished autoimmune response) [26,27]. In the absence of this interaction with the natural world, the chronic stress leads to increases in mental disorders (e.g., anxiety disorders, attention deficit disorders, etc.), stronger negative emotions such as fear and anger, and diminished life satisfaction [10]. Unfortunately, while many people (61%) recognize that stress is a serious problem, only a small number (35%) know how to effectively reduce it [2]. As a result, 78 percent say their stress has increased or stayed the same in the past five years [2].

CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04…$15.00
http://dx.doi.org/10.1145/2702123.2702510

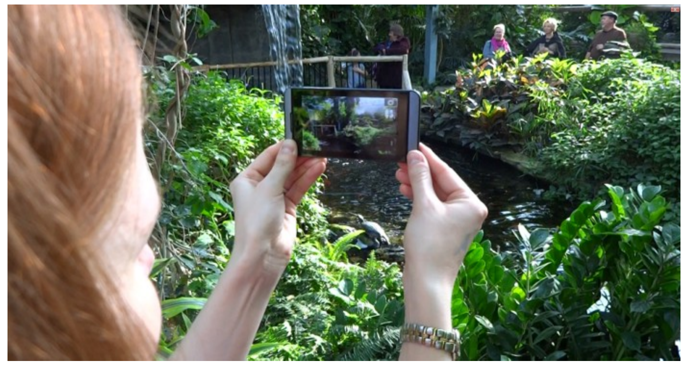


Figure 1. An individual viewing her environment through her phone's camera before she takes a picture.

In human-computer interaction (HCI), the importance of identifying and alleviating stress has been recognized, and various different sensing and mobile intervention technologies have been developed to combat the "stress epidemic" [19]. The field of affective computing has already demonstrated that technology can be used to detect and respond to emotional states [12]. People's emotions can be detected through facial features [12], gestures [4], and physiological measures [20]. Computational methods have also been described for determining the aesthetics of an image specifically [8]. However, the majority of this research has focused on responding to current emotional state, rather than on providing feedback about the likely impact of the environment on that emotion.

In this paper we present *EnviroPulse,* a technology derived from research on the psychological impact of environments, images, and videos to support a novel kind of affective computing. Rather than focusing on real-time detection of a person's affect [12], *EnviroPulse* automatically calculates an expected affective valence (EAV) score for the environment through visual analysis of image and video input in real-time. It can be used to direct people toward places and scenery that promote better psychological well-being (rather than acting as a barrier to them). For example, it can be used to crawl through Google Street View data to develop affective valence maps of the world. Using this, map-

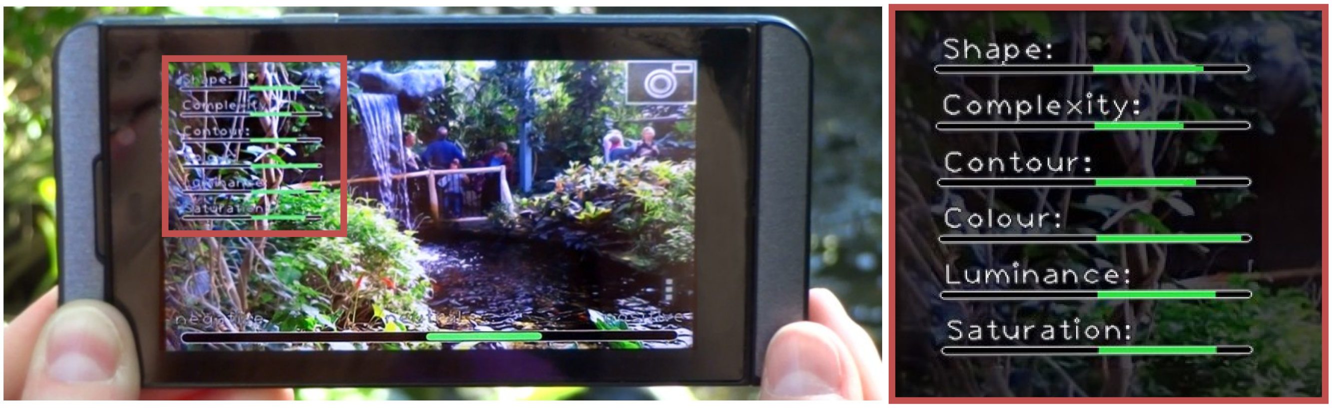


Figure 2. (left) A person using our prototype smartphone camera app to get real-time feedback about the psychological impact of the environment. (right) A zoomed in view (darkened background) of the detailed visual feedback provided with multiple bars.

ping services, such as Google and Bing maps can be augmented with an "affective valence" layer (much like the traffic layers) to show users which locations and routes are most psychologically positive. *EnviroPulse* can also be incorporated into gaming, camera, GPS, and augmented reality smartphone applications to provide real-time analysis of the environment using the smartphone's camera (Figure 2). In order to test *EnviroPulse*, both "in the wild" and in the lab, we describe a prototype and how it is integrated into a smartphone camera app to allow for real-time feedback about a photo or video being captured. Our prototype provides information about the expected affective valence (EAV) of visual data, allowing the smartphone to be used as a "smart window" into the environment (Figure 2) by panning the camera around to find environments that are psychologically positive, and potentially photo-worthy. In order to communicate the EAV score of the environment to the user, we implemented four types of visual feedback of affective valence: two that present factual information via bar graph and data glyph, and two that *induce* emotion and provide individuals with immediate emotional feedback.

We therefore provide three contributions: (1) we designed and developed *EnviroPulse*, a prototype algorithm and interface for the real-time automated analysis of the expected affective valence of images and video, optimized to run on smartphone hardware; (2) we present empirical results from a validation study which suggests that *EnviroPulse* achieves moderate accuracy for determining the affective valence of environments, and offers greater and more reliable accuracy than the average independent human observer; and (3) we describe the design of four versions of real-time visual feedback of EAV, two traditional and two based on psychological literature on human shape and contour preference, to display the affective analysis from our algorithm (i.e., whether something is stressful or calming/pleasant), which we tested in an informal study.

RELATED WORK

The areas of research most pertinent to our work are the HCI literature on *mobile augmented reality* and *affective computing,* and the *psychology of scene and colour perception and preference*.

Mobile Augmented Reality

Using mobile devices, such as smartphones, to create augmented reality (AR) or mixed reality is a practice that is growing in popularity and complexity as mobile computing power continues to improve [28]. Augmented reality in smartphones can take many forms, including reading barcodes or QR codes from the camera view to give details about a product, labeling nearby landmarks in the camera view, or overlaying interactive visualizations on top of the live video of the real environment [21].

Mobile augmented reality typically uses the smartphone's camera and wireless connection combined with other sensors such as the GPS, accelerometer, and compass to compare the individual's location and image data to a database and pull relevant information (e.g., nearby coffee shops) [21,24]. For example, the smartphone app *Layar* overlays information on top of the video feed coming from the smartphone's camera when a barcode or template image is recognized by the software [21]. The development of augmented reality apps has not been limited to simply displaying information. By overlaying information on top of the video feed from the smartphone's camera, mobile augmented reality apps transform smartphones into windows through which individuals can view an augmented version of their environment. In this paper, we use this idea of smartphones as windows to an augmented version of the environment, and display information that may not be immediately obvious: the expected affective impact of the environment on the individual.

Affective Computing

In the field of affective computing, an effort has been made to detect and respond to emotional states of users [12]. Such methods have been used to categorize media using the viewer as a source of information for valence and arousal by monitoring the viewer's facial expressions [12]. Some work has also focused on applying psychological models of valence and arousal to categorize media along the two-dimensional space of valence (positive or negative) versus arousal (high or low) [11,31]. Similarly, computational methods have also been described for determining the aesthetics of images [8].

The field of affective computing has also explored how this affective data may be visualized and communicated to the user via affective visualizations. For example, recent research by Zhang et al. [31] has shown evidence that affective states prompted by movies may be presented in a simple "affective visualization" which uses four colours to depict the four quadrants of high vs. low arousal and positive vs. negative valence. However, the authors explicitly note that affective visualizations ought to be more intuitive and informative than they currently are.

We propose an alternative to these methods for determining and representing affect in technology. Our *EnviroPulse* prototype builds on this work by using an algorithm that can estimate the affect prompted by an environment, rather than sensing it from the user [11,12, 31]. We also attempt to create informative affective visual feedback that can communicate the expected affective impact of the environment.

Psychology of Scene Perception and Preference

Our work combines the wealth of research on the human scene preference and the affective (emotional) influence of individuals' surroundings (i.e., where they live, work, go to relax, etc.) with modern advancements in smartphone technology. We take into account a diverse set of literature looking at (1) visual reward circuits in the human brain that preferentially respond to some types of scenes [5], (2) how the human brain processes visual characteristics of scenes [27], and (3) the way that images of the environment may be analyzed and related to the activation of neural circuits responsible for positive and negative emotions [14,17,27].

Previous research has demonstrated that humans prefer environments (and images) that contain curved shapes and contours, over those that contain jagged contours or pointed shapes [1,3]. This research has found that jagged objects are perceived as more threatening, activating brain areas associated with fear and threat processing [3].

Similarly, research on scene perception has found that urban environments, such as New York City traffic, prompt higher stress levels compared to their natural counterparts (e.g., Central Park) [26,27]. Research on how images of cities influence the viewer's emotional state suggests that urban environments are more stressful than natural environments because they contain more jagged patterns and shapes [26]. Urban environments which contain smooth contours and round shapes are perceived as less threatening, and may even be perceived as similarly pleasant to some natural environments [27]. This line of research has also found that, as visual complexity (density of shapes and contours) in images increases beyond an ideal threshold, the emotional response to the images become more negative and stressful [27]. This finding is consistent with research on the dose-response curve for the impact of tree cover density on stress reduction [13]. It is important to note that these perceptual and neural mechanisms are believed to be "automatic" and consistent across individuals and populations in the absence of disorders [1,3,5,14,17,26,27].

Psychology of Colour Perception and Preference

To get a more complete picture of human preference for scenes and shapes, we also consulted literature on human colour perception and colour preference. Studies investigating the effects of colour on emotions have discovered several factors that influence how humans respond to colours.

Humans have strong preferences for colour hues ranging from blue to green, a strong to moderate aversion for colour hues ranging from green to yellow and yellow to red, and a moderate preference for colours ranging from red to blue (Figure 3, left) [18,25].

When analyzing the effects of colour brightness and colour and saturation on perceived pleasure, researchers have found that brighter (non-white) and more saturated colours are more calming and pleasurable (Figure 3, right). In contrast, darker colours were more likely to elicit feelings of anger, hostility, or aggression. We integrate these findings on scene and colour preference into a prototype for real-time detection of affect induced by an image or video. It should be noted that researchers exploring the effects of colour on emotion highlight that their findings are relevant only in situations in which colours are reasonable (e.g., blue meat would not be preferred even though blue is a highly preferred colour) [18,25].

ENVIROPULSE: ANALYSING AFFECT IN IMAGES

We developed *EnviroPulse*, an algorithm for determining the affective response that is likely to be induced by an environment. In this section, we describe the algorithm that could be incorporated into any system that involves images of the environment, such as mapping services or GPS systems, and then describe the design of an interface for a smartphone camera app that can sense and communicate the affect of photographs and videos.

Implementation

We conceptualized our *EnviroPulse* algorithm using research on the automatic and psychophysiological factors that shape human scene perception and affective responses to visual stimuli [1,3,5,14,25,26,27]. The algorithm is designed to automatically determine the expected affective valence of images and videos across populations. In this section, we highlight some of the challenges in aggregating this research and developing a computational algorithm for real-time visual analysis. We then describe the details of our algorithm.

Emotional Response Factors

We aggregated psychological findings from a variety of research on human scene and colour preference to create six composite curves representing the six factors believed to be responsible for automatic emotional responses to scenes and environments: *shape*, *contour*, *complexity*, *colour hue*,

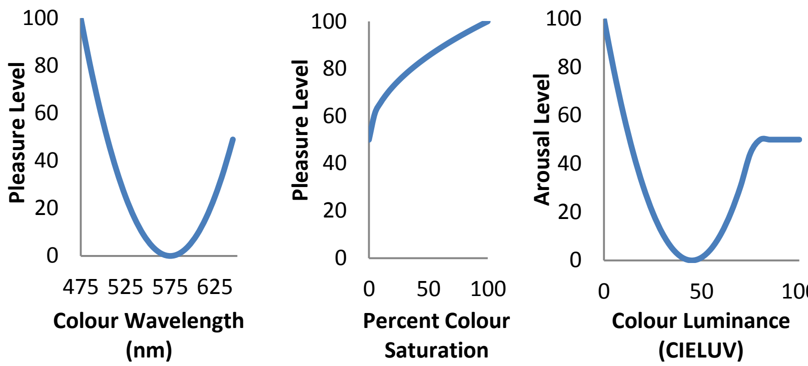


Figure 3. Curves of mappings used in our algorithm from colour space to pleasure and arousal levels. Previous work [25] shows that people prefer (left) blue-green (~475-500 nm) and red (~650 nm), and dislike yellow (~570 nm) and (middle) high colour saturation and that (right) luminance can be used to estimate arousal levels.

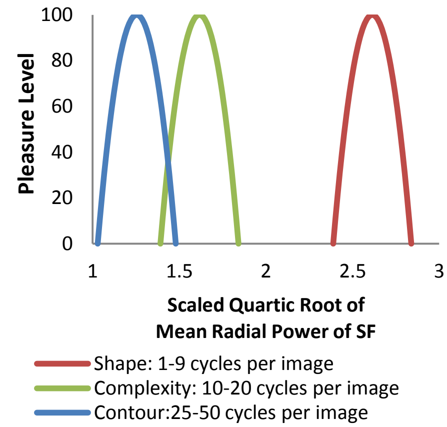


Figure 4. Curves informed by psychology research relating affective responses to shape, complexity and contour with the spatial frequencies of an image.

colour saturation, and *colour brightness* (Figure 3-4) [1,3,25,27]. The composite curves (Figure 3-4) represent both the aggregation of past findings from different sources [1,3,25,27], and our interpretation of these findings, resulting in a novel representation of emotional response factors. This was in part because the psychology literature on human preference presents its findings based on relative subjective categorical scales (e.g., 1 = not pleasant, 4 = moderately pleasant, 7 = highly pleasant). Since the measurement scales differed across researchers and experiments, and were categorical in nature, we had to convert such results into an absolute metric that would allow for cross-research aggregation. We did this by mapping results from previous literature on human preference onto quadratic, and square-root functions, which have been previously shown to represent human responses to a variety of cognitive and emotional situations and stimuli [13,25,27].

Expected Affective Valence (EAV) Score

Our *EnviroPulse* prototype captures the pleasure and arousal associated with changes in the six psychological factors using individual scores for each factor. Each of the scores is computed in parallel and is then combined into one expected affective valence (EAV) score. Each score is calculated using a quadratic function or its inverse, and is normalized between 0 and 1. The equations were extrapolated from previous work [9,14,25,27] and are of the form:

$$y = \left(\frac{x-a}{b}\right)^2,$$

where the score can be either y (quadratic), x (root), or 1-y (inverted quadratic).

Based on Valdez and Mehrabian's theory of colour's contribution to affective response [25], whose findings have been recently been replicated by Palmer and Schloss [18], colour wavelength (quadratic, a = 575, b = 100) and saturation (root, a = 0.5, b = 0.05) are computed for each pixel and averaged to create a colour pleasure score, which is "calmed" using a colour arousal score computed using luminance (inverted quadratic, a = 45, b = 45; or the constant 0.5, neutral, if luminance is above 80, white), again averaged across all pixels. This combined equation becomes:

$$color\ score = color\ pleasure \cdot color\ arousal$$

To compute theoretical pleasure levels for the content in the scene, a scaled discrete Fourier transform is used on a grayscale version of the scene to extract the average radial power of each of the three ranges of spatial frequencies (Figure 3, [9,14,16,27]). A score for each of *shape* (inverted quadratic, a = 2.614, b = 0.225), *complexity* (inverted quadratic, a = 1.61627, b = 0.225), and *contour* (inverted quadratic, a = 1.255, b = 0.225) is calculated, and their average is used to provide an overall *content score*.

The colour score and content score are then combined into the EAV score using weights inferred from the variance and regression analysis in Valdez and Mehrabian [25], and Valtchanov [27], as follows:

$$EAV\ Score = 0.3 \cdot colour\ score + 0.7 \cdot content\ score$$

Scores of 40 or below represent a negative affective impact, while scores from 40–60 represent a neutral affective impact, and scores of 60 or greater represent a positive affective impact. We used this procedure for calculating the emotional impact of incoming image data, to label scenes as psychologically positive, neutral, or negative, which we then make available in real-time through visual feedback.

Incorporating EnviroPulse into Image-Based Software

The EAV score can be calculated for any image, and therefore can be incorporated into any image- or camera-based technology. For instance, the EAV score could be calculated for all images on Google Street View, allowing Google Maps to provide feedback about the likely affective impact of a walking route. *EnviroPulse* could also be incorporated into wearable head-mounted technology, such as

Google Glass, warning of danger when a person is spending too much time in dense city environments. We next describe a specific design that incorporates *EnviroPulse* into a smartphone camera application to provide real-time feedback about the expected average emotional response to the environment, and the images/video currently being captured.

AFFECTIVE VISUAL FEEDBACK IN A CAMERA APP

To demonstrate how *EnviroPulse* can be used in an interface, we implemented a smartphone camera application that uses EAV score to measure the emotional valence of the image/video feed coming from the smartphone's camera. This required that the result of the analysis be presented in a rapid, comprehensive, and intelligible manner to the individual using a smartphone. We wanted to avoid simply presenting the overall EAV score in a numerical format, since we expected that it would encourage users to always try to achieve the highest rating. Instead, we made it possible to recognize how moving the smartphone camera around (i.e., changing in the visible scene) results in more positive or negative scores along the six emotional response factors. Visualizing the multiple dimensions that determine EAV will not only help explain the score, but allow more freedom to establish photography or videography that captures a wider spectrum of emotional response, rather than assuming that what this prototype rates as "positive" relates to the desired outcome.

In order to accomplish this, we consulted past attempts at affective visualizations [6,31] and the psychological literature related to affect induction [26,29], as well as literature in multi-dimensional visualizations [15], to inform the design of four versions of real-time visual feedback. We designed four examples of visual feedback of the calculated EAV score because we wanted to investigate whether factual or affective visual feedback could better communicate the estimated affective valence of an environment. Our first two versions use a traditional approach of communicating the information by simply presenting the facts in bar or glyph forms, and our second two versions make use of our algorithmic knowledge of affect to *induce* the affect more abstractly through waves and shapes.

Factual Visual Feedback

To communicate the information contained within our analysis prototype, we first considered the six scores as separate dimensions in a multi-dimensional visualization, and used traditional information visualization techniques for displaying these multiple dimensions: *bars* and *glyphs*.

Multiple Bars

We used *multiple bars* (Figure 5, top-left), similar to a bar chart, which projects from a centred neutral label toward either a negative or positive label anchor while altering its colour along a red to white to green gradient to indicate the negative, neutral, or positive valence of the environment. We used both a single bar to indicate the overall score at the bottom of the display, and a series of six bars to represent the multiple dimensions used to determine the score (shape, contour, complexity, colour hue, colour saturation, and colour brightness). The colour scheme was chosen based on the commonly used social symbols of red indicating threat (e.g., red traffic lights, red stop sign, warning labels on drugs, etc.) and green indicating something positive (e.g., green traffic lights, nature/vegetation, etc.) since it is a highly pleasant colour [25].

Glyphs

We also designed visual feedback based on star *glyphs* [15] (Figure 5, bottom-left), which displays scores of the environment along six axes representing the measured dimensions used to calculate affective valence. Much like *multiple bars*, it gradually shifts between red, white, and green to represent the overall negative, neutral or positive valence.

Affective Visual Feedback

Since the foundation of our prototype is based on the principle that a person's affect can be estimated from knowledge of several essential features of an image, our second approach to visualizing this information was to communicate our affective analysis using the affective channel, rather than the more traditional approach of simply displaying the facts. We compared the representations of valence used in the Self-Assessment Manikin (SAM) [6] with previous literature in psychology on affective responses to shapes and contours, and found that the SAM exploits well-documented relationships between affect and shapes and contours [1,3,22,27]. Similar to the SAM graphic, we used the relationship between shapes, contours, colours and affect to generate two affective versions.

We recognize that there is some redundancy in displaying affect about an image by inducing affect through visual feedback (this affect will be induced by the environment or image itself). However, it is worth noting that participants in the studies on human affective responses to scenes [26] took up to 10 minutes to fully feel the effect of the environment. The techniques used in our affective visual feedback were shown to induce these responses in seconds.

Affective Waves

We presented visual feedback through a *wave* (Figure 5, top-right), which alters its contour type (smooth vs. jagged), amplitude (low vs. high), pulse rate (short vs. long), and colour (red to white to green representing negative to neutral to positive overall valence). This design was inspired by Poffenberger and Barrows' [22] research on the affective responses to waves, which demonstrated that jagged waves were found to be agitating, while smooth waves were found to be calming; waves with higher amplitudes were found to be more arousing compared to those with lower amplitudes; and waves with short periods were found to be more unnerving than those with longer periods. *Affective wave* was coded so that stressful negative environments produced jagged, tall, rapid red waves while calming positive environments produced smooth, short, wide green waves.

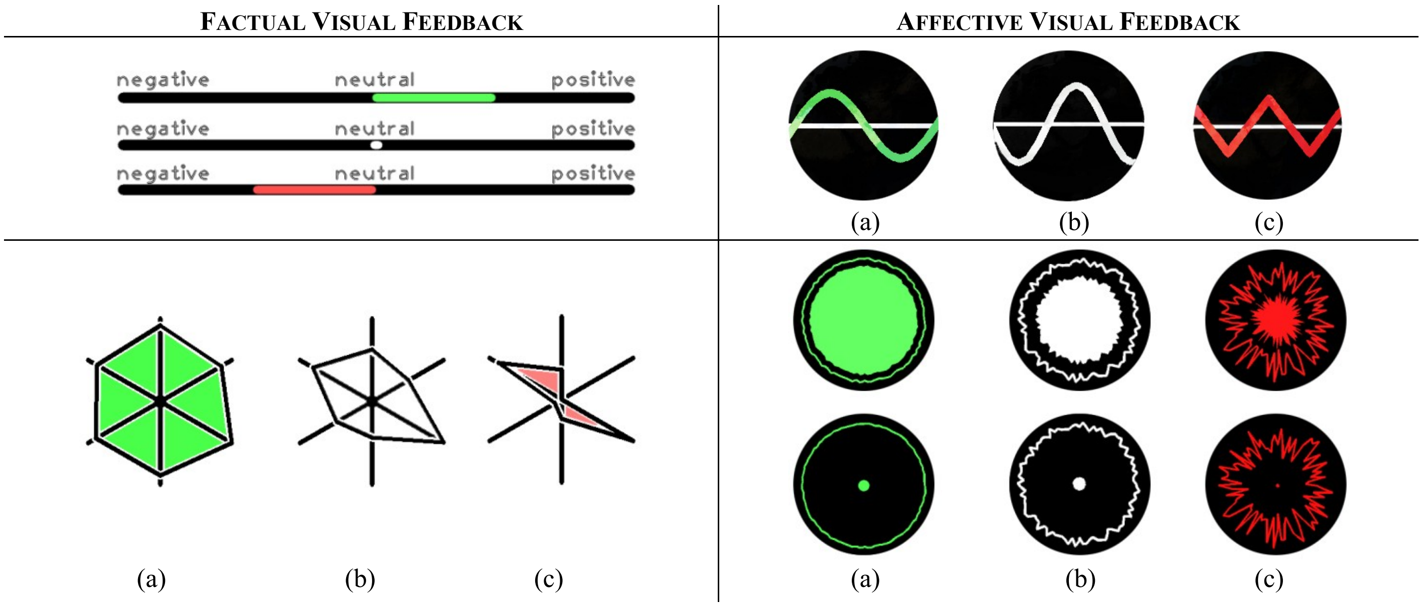


Figure 5. (top-left) bar, (top-right) wave, (bottom-left) glyph, and (bottom-right) pulsing amoeba. For all versions, (a) is moderately positive, (b) is neutral, and (c) is moderately negative and colour changes along a gradient from green, to white, to red, respectively.

Pulsing Amoeba

We also present visual feedback through a *pulsing amoeba* (Figure 5, bottom-right), which alters its contour type (smooth/jagged), its pulse rate (short/long), and its colour (red to white to green, representing negative to neutral to positive valence). For negative stressful environments the amoeba's shape tenses up and shifts toward being jagged and spiked, while the radiating pulse (its "heart rate") dramatically increases and changes to bright red, resembling a warning light. For positive calming environments, the amoeba relaxes and its shape shifts toward a smooth circular contour while its pulse rate slows significantly and its colour changes to green. The pulse occurs from the centre toward the outside, and then retracts back toward the centre.

EVALUATION OF SCENE ANALYSIS ALGORITHM

To determine if our *EnviroPulse* prototype gives estimates of affective valence that are consistent with human responses to environments, we conducted a validation study: our prototype's estimations of the affective valence of images were compared to previous empirical data from human participants. In a previous set of studies by Valtchanov [27], participants viewed photographs of low and high density cities, rural villages, parks, savannah landscapes, forests, beaches, mountains, and lakes while physiological measurements were taken. Participants in these studies also answered questionnaires which assessed the affective impact of viewing each photograph. For our validation study, we used *EnviroPulse* to analyze the photographs used in this past research and compared these scores to those given by participants in the study.

It is important to note that while the research by Valtchanov [27] was partially used to inform the design of *EnviroPulse,* our algorithm uses a composite of several other studies [1,3,9,15,25] to create a new metric for scene analysis: We use six response factors that we have inferred from research [1,3,9,14,25,27], compared to the single response factor that is described by Valtchanov [27]. Since this is the first time these factors have been combined into a single equation for determining EAV of visual stimuli, it required validation.

Validation Experiment Setup

Sixty-five photographs of environments (e.g., cities, parks, forests, mountains, etc.) that were previously used in studies by Valtchanov [27], were projected onto a screen, one at a time, for 3 seconds each. An Android smartphone running our custom camera application with *EnviroPulse* was mounted onto a tripod in front of the screen, such that the smartphone's camera had a complete view of the projected image (Figure 6). Photographs were presented for 3 seconds to allow the smartphone's camera to adjust to changes in luminance and focus between images. Our camera app analyzed each photograph and assigned it an affective score.

EnviroPulse versus Human Perception

We compared the affective scores of the 65 photographs given by our algorithm to the averaged affective scores given by 53 participants (26 male, 27 female) for the same images in Valtchanov [27] using an independent samples t-test. Scores given by our algorithm were on a scale from 0 (highly negative) to 50 (neutral) to 100 (highly positive), while scores from studies by Valtchanov [27] were on a scale from -6 (highly negative) to 0 (neutral) to +6 (highly positive). To allow for comparison, both scores were standardized to the same scale of 0 to 100. The test revealed that the affective scores given by our algorithm were not significantly different from those given by human participants ($t(128) = 0.38$, $p = .70$, $SE = 2.70$). This promising result suggested that our algorithm's analysis was giving similar affective scores to human participants.

In order to measure if *EnviroPulse* correctly categorized the photographs based on valence, we computed the agreement between the valence category assigned by our prototype and that given by the average of the participants in Valtchanov [27] for each of the photographs using Cohen's Kappa (K), a statistic that measures the proportion of agreement above chance [30]. There was moderate agreement ($K = 0.51$, $p < .001$) between the valence category assigned by our algorithm and that assigned by human participants for each image (Figure 7). *EnviroPulse's* categorization was identical to the categorization given by the average participant in 72.8% of cases (81.3% of the negative images, 62.2% of the neutral images, and 75.0% of the positive images). *EnviroPulse's* accuracy was similar to other, more complex, automated algorithms that extract up to 15 image features when determining image aesthetics [8], validating our approach. The moderate accuracy of our algorithm, and past attempts [8,23], reflects the difficulty of automating image analysis to determine subjective preference [8,23].

To gain perspective on the meaning of our achieved level of agreement between human data and EnviroPulse, we investigated how accurately the participants in the study could estimate other participants' affect scores. This was done by computing K values between participants from Valtchanov [27] and the average across all participants. On average, independent human observers were worse at estimating the response of the overall population ($K = 0.36$) compared to our *EnviroPulse* prototype ($K = 0.51$). This promising result suggests that individuals may be unaware of the impact the environment has on them, and that many individuals could benefit from *EnviroPulse's* analysis of the environment.

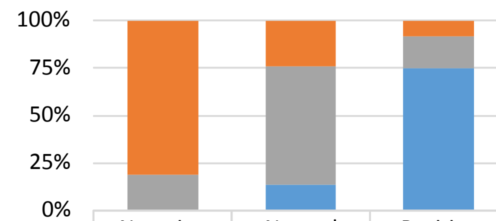
INFORMAL STUDY ON AFFECTIVE VISUAL FEEDBACK

The validation study provided evidence that our scene analysis algorithm was able to discern the affective valence of environments using a smartphone camera. This indicated that we had succeeded in implementing past psychological findings on scene, shape, contour, and colour preference into a prototype that could calculate the valence of a scene with moderate reliability. We next examined whether our four versions of visual feedback could successfully communicate the outcome of this calculation to individuals in a simple, rapid, and meaningful manner. We used the Android SDK (version 2.3 or higher) to develop the animated visual feedback and implemented them into our Android camera application. A detailed view of the six emotional response factors used to calculate the overall affective valence of the scene was optionally displayed (using *multiple bars*) for all four types of visual feedback.

Participants

We deployed an *EnviroPulse* camera app to a group of 9 individuals (7 male, 2 female) of mixed academic expertise. Four had a background in engineering and human-computer interaction, and the other 5 had a background in psychology. All participants installed our *EnviroPulse* camera app onto their Android smartphones (including Galaxy S3,

Figure 6. Smartphone mounted on tripod analyzing projected photographs of various environments.



Participant Rating of Environment:	Negative [0,40]	Neutral (40,60)	Positive [60,100]
	EnviroPulse Rating of Environment:		
Negative [-6.0,-1.2]	**81.3%**	24.3%	8.3%
Neutral (-1.2, 1.2)	18.8%	**62.2%**	16.6%
Positive [1.2,6.0]	0.0%	13.5%	**75.0%**

Figure 7. Agreement between expected affective valence score from *EnviroPulse* and ratings from participant data [27].

Nexus 4, Galaxy Nexus, Galaxy Note) and were asked to casually use it over the course of two weeks.

Interview Responses

Participants were individually interviewed about their experience with our augmented camera app after they had used it for two weeks. During the interview, they were shown a video of the app analyzing different environments using the four types of visual feedback, and then asked about the different states of affective valence (e.g., jagged rapidly pulsing red amoeba vs. smooth slowly pulsing green amoeba).

Participant Impressions on Ease of Use

Participants were asked about which version they found easiest to understand. All participants unanimously agreed that it was the dynamically coloured bar. Participants reported that it was easy to understand because it projected from neutral (centre of the screen) to negative (left) or positive (right), combined with an "intuitive" colour (red for negative valence, and green for positive valence). Participants reported that this made it easy to quickly understand whether their environment was negative or positive based on the direction, size, and colour of the bar.

Participants were asked to rank the four types of feedback in terms of their intuitiveness. Overall participants rated them as follows: 1) *bar,* 2) *pulsing amoeba*, 3) *wave*, 4) *glyph.* This ranking appeared to follow the amount of information presented. The *bar* and *pulsing amoeba* presented only the overall affective valence of the scene, while the *wave* presented two (arousal and affect), and the *glyph* presented the scores for the six emotional response factors (i.e., shape, contour, complexity, colour hue, colour saturation, and luminance). While it may be a coincidence, it is possible that the complexity of the displayed data impacted how "intuitive" participants found the visual feedback to be. It is possible that feedback with more information may become preferred over time.

Participant Impressions on Aesthetics

Participants were asked which visual feedback they found to be most aesthetically interesting. Eight out of nine participants described the *pulsing amoeba* as being the most "captivating". Participants commented on how they liked the continuous pulse of the amoeba and how it changed with the environment. Participants described feeling like the *pulsing amoeba* was responding to the environment.

When viewing the video showing the different states of the *pulsing amoeba*, four of the participants spontaneously noted how the jagged and rapidly pulsing red-coloured state of the amoeba (i.e., the form it takes when the environment would induce a negative response) felt "threatening and unpleasant". Similarly, three of the participants spontaneously commented that the smooth, slowly pulsing, green state of the amoeba was "soothing and aesthetically pleasant to look at". These were exciting comments, since they indicated that we were successfully eliciting affect by manipulating the shape, contours, and colour of the visual feedback in accordance with previous research [3,22]. When participants were asked to rank the feedback in terms of aesthetics, they were ranked as follows: 1) *pulsing amoeba*, 2) *wave*, 3) *bar*, 4) *glyph.* These results were interesting since they indicated that the two versions employing principles from the shape preference literature were rated higher in aesthetic appeal. While further studies would be required to empirically validate these findings, these early results suggest that it may be possible to create affective visual feedback using design rules based on psychological research [3,22].

Participant Impressions on Accuracy of EnviroPulse

Participants were asked about whether they felt the score given by the camera app matched how they felt about the environment. None of the participants reported feeling like the tool was incorrectly assessing the affective valence of their environment (e.g., when walking, their office, the local park, etc.). All participants were informed that they were testing an early version of the tool that may show inaccurate analyses, in an attempt to prompt mistrust in its accuracy, but this did not change their minds. Based on *EnviroPulse's* moderate accuracy for assessing a wide range of environments, it is possible that participants did not encounter a situation where they would disagree with it.

DISCUSSION

In this paper, we explored the possibility of developing technology that can automatically determine the expected affective valence of the environment using the psychological correlates of scene, shape, contour, and colour perception [1,3,22,25,27] in order to enable novel interaction methods. We also investigated the use of affective visual feedback, which can potentially induce affect, as a method for communicating the expected affective valence of the environment. We believe that by using design principles inspired by research on human scene, shape, and contour perception [1,3,22,27], it is possible to create technology that is effective and efficient in affective analysis and visualization. With a camera application that incorporates *EnviroPulse*, such as the one we developed and deployed to participants, any individual with a smartphone can use their phone's camera to get a real-time affective analysis of their immediate environment and then see the results on their screen. However, there are many more possible applications of *EnviroPulse* outside of integration into camera apps.

Applications of EnviroPulse

Unlike the current reactive technologies which attempt to measure a person's affect and respond to it, *EnviroPulse* is a technology which determines the expected affective valence of an image or video, and can be incorporated into existing technologies in a variety of novel ways:

EnviroPulse can be used to crawl through street-view data from mapping services, such as Google and Bing maps, to create a "affective valence" layer (much like the traffic layers) to show users which locations and routes are most psychologically beneficial. This could be integrated into existing computerized mental-health interventions developed by the HCI community [19] to help reduce chronic stress by pre-emptively directing users through psychologically positive environments.

The affective layers for maps could be integrated into virtual assistants (e.g., Siri and Google Now), allowing them to suggest pleasant nearby areas, or display stored photos from scenic locations when given emotional input (e.g., "I'm feeling stressed", "I want to find a happy place").

Using *EnviroPulse,* photo, video, and augmented reality applications could also begin to sense the affective valence of the surrounding environment, creating awareness and potentially prompting engagement with the environment in a positive stress-reducing manner.

Lastly, mobile games could be augmented with *EnviroPulse* to guide players toward psychologically beneficial environments. For example, a "virtual pet" can be made to desire environments that are psychologically positive, becoming happier when the player shows it such an environment via the smartphone camera. *EnviroPulse* is the core technology necessary in any of these systems to sense the expected affective valence of real-world environments.

Limitations and Improvements of EnviroPulse

Our validation study indicates that *EnviroPulse* may be better at estimating the valence of an environment compared to the average human observer, and that it is comparable to previous attempts to automatically determine subjective visual aesthetics of images [8]. The moderate accuracy of our algorithm reflects the difficult problem of automatically computing subjective preference without individualized user data such as age, gender, et cetera [8,23].

Thus, in order to increase accuracy when determining the affective valence of environments using our prototype or similar automated approaches [8], additional individualized data must be considered. Past literature indicates that people use a combination of bottom-up perceptual mechanisms (similar to those used by *EnviroPulse*) and top-down cognitive mechanisms (such as past experience and expectations) when responding to an environment [27,28]. Our focus on implementing mathematical interpretations of the automatic bottom-up perceptual mechanisms for affective responses has allowed us to automatically analyze and categorize the affective valence of scenes in real-time using a smartphone. However, it is evident determining subjective preference with higher accuracy will require that the system is seeded with user statistics over time.

We believe that *EnviroPulse*, and other similar systems, should learn from the individual and dynamically adjust calculations for preference based on user input. By doing this, *EnviroPulse* could reweight the six emotional response factors to provide a more personalized experience. For example, if someone has a negative association with a particular type of environment, the characteristics of the environment could be associated with negative affect: A sandy beach on a sunny day may be a perceptually pleasant sight for most individuals [27], but someone who gets sunburned easily may experience negative affect (i.e., fear of sunburn/pain). In this case, the combination of blue water and yellow/white sand co-occurring could be associated with a personalized assessment of all beach environments by the system, without influencing positive assessments of other environments that contain blue water and other terrain.

Investigation of Affective Visual Feedback

We have presented an informal study describing the opinions of a small group of individuals on four types of visual feedback, two of which presented factual information (via bars and glyph), and two of which communicated affective information through psychological phenomena. While future work would benefit from a more in-depth study with a higher number of participants, there are several informative insights that can be gleaned from the current results. The first, and possibly most important, is that individuals from varied backgrounds are open to new styles of visual feedback when it comes to presenting affective data. Participants in our informal study were most excited and interested in the *pulsing amoeba* and *wave* visual feedback, even though they were not the easiest to interpret. This brings forth an interesting problem: affective visual feedback may be more engaging, but they may not necessarily be the most informative nor easiest to understand, suggesting that there may be a trade-off between affective visual feedback and simpler or more common feedback (e.g., bar graphs).

The affective feedback (e.g., *pulsing amoeba* and *wave)* used in our informal study attempted to incorporate design principles inspired by psychology. Participants suggested that we should make opposite ends of the spectrum (i.e., positive vs. negative) more exaggerated for the affective visual feedback. For example, for the *pulsing amoeba*, participants suggested that it should pulse slower (0.2 Hz) than it does currently (0.5 Hz) for a psychologically positive environment and quicker (2 Hz) than it currently does (1 Hz) for a psychologically negative environment. There are currently no guidelines for developing optimal affective visual feedback that are guaranteed to induce the desired affect in participants. This means that it is very difficult to develop feedback that induces affect reliably and accurately. We have taken early steps into developing affective feedback by consulting psychological literature on affective responses to shape and contours [3,22]. We believe that affective visualizations require more rigorous investigation and refinement before they can be effectively used to communicate emotional data.

CONCLUSION AND FUTURE WORK

In this paper, we presented our design and development of *EnviroPulse*, a prototype technology and interface for the analysis of the affective valence of environments. We presented empirical results which suggest that *EnviroPulse* achieves moderate accuracy for determining the affective valence of environments, and offers greater and more reliable accuracy than the average independent human observer. The accuracy of our technology is similar to other technologies for determining semi-subjective phenomena, such as aesthetics [8]. We also investigated the use of factual versus affective visual feedback to communicate the affective analysis from *EnviroPulse* (i.e., whether something is stressful or calming/pleasant), through an informal study.

In the future, we plan to improve the accuracy of *EnviroPulse* by integrating user-specific profiles that capture user input over time. We are also working on implementing real-time sound analysis into *EnviroPulse* (by using input from the phone's microphone), since there is strong evidence in psychology literature that sound and music can have profound influences on physiological, cognitive and behavioural responses to environments [26,29]. We believe that adding sound analysis and user-specific profiles to our prototype will significantly enhance accuracy and reliability.

Acknowledgements

We would like to thank the Natural Sciences & Engineering Council of Canada (NSERC), the Social Sciences & Humanities Research Council (SSHRC) and its IMMERse network, and the University of Waterloo Games Institute.

REFERENCES

1. Amir, O., Biederman, I., & Hayworth, K.J. The neural basis for shape preferences. *Vision research 51*, 20 (2011), 2198-2206.
2. American Psychological Association, *Stress in America 2013*, http://www.apa.org/news/press/releases/stress/
3. Bar, M., & Neta, M. Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia 45,* 10 (2007), 2191-2200.
4. Baltrusaitis, Tadas, et al. Real-time inference of mental states from facial expressions and upper body gestures. In *Automatic Face & Gesture Recognition and Workshops (FG 2011),* IEEE Press, (2011), 909-914.
5. Biederman, I., & Vessel, E.A. Perceptual pleasure and the brain. *American Scientist 94*, 3 (2006), 249-255.
6. Bradley, M.M., & Lang, P.J. Measuring emotion: the self-assessment manikin and the semantic differential. *J BEHAV THER EXP PSY 25*, 1 (1994), 49-59.
7. Coyne, R. Nature vs. Smartphones. *Interactions, 21,5* (2014), 24-31.
8. Datta, R., Joshi, D., Li, J., & Wang, J. Z. Studying aesthetics in photographic images using a computational approach. In *Computer Vision–ECCV* (2006), 288-301
9. Fernandez, D., & Wilkins, A.J. Uncomfortable images in art and nature. Perception 37, 7 (2008), 1098-1113.
10. Gullone, E. The biophilia hypothesis and life in the 21st century: Increasing mental health or increasing pathology? *Journal of Happiness Studies*, *1*,3 (2000), 293-322.
11. Hanjalic, A. Extracting moods from pictures and sounds: Towards truly personalized TV. *IEE Signal Processing Magazine, 23*, 2 (2006), 90-100.
12. Hernandez, J., Hoque, M.E., and Picard, R.E. Mood Meter: large-scale and long-term smile monitoring system. In *ACM SIGGRAPH 2012 Emerging Technologies*, ACM Press (2012), 1.
13. Jiang, B., Chang, C. Y., & Sullivan, W. C. A dose of nature: Tree cover, stress reduction, and gender differences. *Landscape and Urban Planning*, *132* (2014), 26-36.
14. Juricevic, I., Land, L., Wilkins, A., & Webster., M.A. Visual discomfort and natural image statistics. *Perception 39,* 7 (2010), 884-899.
15. Lie, A. E., Kehrer, J., & Hauser, H. Critical design and realization aspects of glyph-based 3D data visualization. In *Proc. 25th Spring Conference on Computer Graphics*, ACM Press (2009), 19-26.
16. Oliva, A., & Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. Journal of Computer Vision 42*, 3 (2001), 145-175.
17. O'Hare, L., & Hibbard, P. B. Spatial frequency and visual discomfort. *Vision Research 51,* 15, (2011), 1767-1777.
18. Palmer, S. E., & Schloss, K. B. An ecological valence theory of human color preference. In *Proc. National Academy of Sciences*, *107*, 19, (2010), 8877-8882.
19. Paredes, P., Gilad-Bachrach, R., Czerwinski, M., Roseway, A., Rowan, K., & Hernandez, J. PopTherapy: Coping with Stress through Pop-Culture. In Proc *Pervasive Technologies for Healthcare*, (2014).
20. Poh, M.Z., Kim, K., Goessling, A.D., Swenson, N.C., Picard, R.W., *Heartphones: Sensor Earphones and Mobile Application for Non-obtrusive Health Monitoring.* In Proc. *IEEE International Symposium on Wearable Computers*, (2009) 153-154.
21. Pence, H.E. Smartphones, smart objects, and augmented reality. *REF LIB 52*, (2010): 136-145.
22. Poffenberger, A.T., and Barrows, B.E. The Feeling Value of Lines. *Journal of Applied Psychology 8*, 2 (1924), 187-205.
23. Quercia, D., Schifanella, R., & Aiello, L.M. The shortest path to happiness: Recommending beautiful, quiet and happy routes in the city. In *Proc*. *Hypertext and social media*, ACM Press (2014), 116-125.
24. Sandor, C., Dey, A., Cunningham, A., Barbier, S., Eck, U., Urquhart, D., Marner, M.R., Jarvis, G., and Rhee, S. Egocentric space-distorting visualizations for rapid environment exploration in mobile mixed reality. In *Virtual Reality Conference (VR) 2010*, IEEE Press (2010), 47-50.
25. Valdez, P., & Mehrabian, A., Effects of color on emotions. *Journal of Experimental Psychology: General 123*, 4 (1994), 394-409.
26. Valtchanov, D., & Ellard, C. Physiological and affective responses to immersion in virtual reality: Effects of nature and urban settings. *JCR 3*, 4 (2010), 359-373.
27. Valtchanov, D., Exploring the Restorative Effects of Nature: Testing A Proposed Visuospatial Theory. (Doctoral dissertation). U of Waterloo, ON, CA, 2013.
28. van Krevelen, D.W.F., & Poelman, R. A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality 9,* 2 (2010), 1-19.
29. Västfjäll, D. Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae 5*, 1 (2002), 173-211.
30. Viera, A.J., & Garrett, J.M. Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine 37*, 5 (2005), 360-363.
31. Zhang, S., Huang, Q., Jiang, S., Gao, W., & Tian, Q. Affective visualization and retrieval for music video. *Multimedia, IEEE Transactions on Multimedia 12*, 6 (2010), 510-522.