Visual angle modulates affective responses to audiovisual stimuli

Accepted manuscript

Dominik Gall*, Marc Erich Latoschik

Human-Computer Interaction, University of Würzburg, Am Hubland, 97074 Würzburg, Germany

Abstract

What we see influences our emotions. Technology often mediates the visual content we perceive. Visual angle is an essential parameter of how we see such content. It operationalizes visible properties of human–computer interfaces. However, we know little about the content-independent effect of visual angle on emotional responses to audiovisual stimuli. We show that visual angle alone affects emotional responses to audiovisual features, independent of object perception. We conducted a $2 \times 2 \times 3$ factorial repeated-measures experiment with 143 undergraduate students. We simultaneously presented monochrome rectangles with pure tones and assessed valence, arousal, and dominance. In the high visual angle condition, arousal increased, valence and dominance decreased, and lightness modulated arousal. In the low visual angle condition, pitch modulated arousal, and lightness affected valence. Visual angle weights the affective relevance of perception modalities independent of spatial representations. Visual angle serves as an early-stage perceptual feature for organizing emotional responses. Control of this presentation layer allows for provoking or avoiding emotional response where intended.

Keywords: visual angle, field of view, screen size, emotion, immersion, multimodal perception

1. Introduction

What we see influences and evokes emotions. This content-based evocation works with single pictures, e.g., displaying a crying person or a tortured animal, as well as by stories told as, e.g., in movies. Notably, this effect does not only occur in response to iconic or tangible objects or events. For example, color-to-emotion associations elicit affective responses even to very simple stimuli (Kaya and Epps, 2004).

Today, what we see is often mediated or even determined by technology. Visual information is central for graphical displays prominent in current Human-Computer Interaction paradigms, from 2D Graphical User Interfaces to Virtual and Augmented Reality. Inevitably, the graphical information presented in such displays impacts the affective response of users. For the design of graphical user interfaces, it is, therefore, crucial to understand the effect of visual displays on emotions. Such knowledge allows for preventing unintended or eliciting intended emotions. The content layer surely has affective power (e.g., Lang et al. (1999)); however, we know little about the affective power of the presentation layer.

The presentation layer of graphical displays has several properties, including display size, resolution, maximum brightness, color resolution, contrast, and dynamic range, and field of view. The combination of size, resolution, and field of view determines the overall maximum quantity of visual information displayed simultaneously. The user-to-display distance then determines the perceivable information. The most general measure to operationalize the resulting quantity of visual information is the visual angle.

The goal of this study is to identify how the visual angle modulates the emotional responses to audiovisual content. Previous studies on emotional responses to visual angle focus on the content layer. Larger naturalistic stimuli induced increased arousal and dominance ratings compared to smaller

*Corresponding author

Email addresses: dominik.gall@mytum.de (Dominik Gall), marc.latoschik@uni-wuerzburg.de (Marc Erich Latoschik)

stimuli (Detenber and Reeves, 1996); they also induced lower heart rates and higher skin conductance (Reeves et al., 1999); as well as increased arousal ratings and skin conductance (Codispoti and De Cesarei, 2007). Naturalistic stimuli are more engaging compared to simple stimuli and hence provide suitable means to study emotional responses (see De Cesarei et al. (2017) for a review). Natural scenes, however, induce mechanisms of object and scene recognition. They induce effects that might confound the analyses of visual angle: notably, object size variations and spatial information destiny. Moreover, general attentional processes differ for naturalistic and simple stimuli: Observers, for example, categorized natural scenes in the near absence of spatial attention while they failed to distinguish simple stimuli (Li et al., 2002; Fei-Fei et al., 2005). In this study, we address the emotional effect of visual angle without using photorealistic images.

Size variations of naturalistic stimuli confound the perceived distance of the observer to the stimulus. Size variations also confound the ability of the observer to differentiate the content of the stimuli. Both perceived distance and content discrimination, in turn, affect emotional responses to stimuli. Apart from these high-level processes, low-level perceptual information already starts to initiate emotional responses. This low-level information is independent of presentation properties like color, brightness, spatial information density, or complexity (Junghöfer et al., 2001). Early stages of natural scene recognition are independent of high-level object recognition (Oliva and Torralba, 2006; Torralba et al., 2006). Regularities in the appearance of scenes, however, modulate object recognition even in early processing stages (see De Cesarei et al. (2017) for a review). The use of such regularities limits the impact of visual noise, information density, or object size. The processing of regularities in scenes correlates to neural activities that associate with categorization tasks (De Cesarei et al., 2013, 2015). The visual angle is independent of scene characteristics. We, hence, assume that the visual angle of a percept constitutes a visual feature that is available independent of object recognition processes. In this study, we investigate the impact of visual angle while minimizing the impact of object recognition effect, which might confound emotional responses.

Visual and auditory perception is interdependent (McGurk and MacDonald, 1976). This interdependence affects the processing of emotional stimuli. Visual information influences the perception of emotional auditory content: Videos of musical performance, for example, affect emotional response to musical stimuli (Chapados and Levitin, 2008; Vines et al., 2006, 2011; Thompson et al., 2008). Furthermore, visual information intensifies emotional ratings of congruent auditory content (Cox, 2008). Vice-versa, auditory-induced emotions affect the emotional perception of visual content (Logeswaran and Bhattacharya, 2009; Marin et al., 2012). Hence we hypothesize that visual angle not only affects the emotional responses to visual but also to auditory information. In the current study, we address how the visual angle interacts with auditory content perception to organize emotional responses.

This article reports novel findings on the effect of visual angle on affective responses to audiovisual content features. We investigated the impact of visual angle on affective responses to visual and auditory stimulus features independent of object recognition processes. Visual content features modulated affective perception only when the visual angle was high. Auditory content features modulated affective perception only when the visual angle was low. Our results indicate that visual information availability modulates emotional responses on early processing steps. We suggest that this processing step is independent of object recognition processes.

1.1. Theoretical background

The importance of a percept for the intentions of the perceiver shapes her emotional responses (Vuilleumier, 2005). The visual angle is one of the first perceived properties of such a percept. Previous studies confound the emotional effect of visual angle with variations in perceived proximity and information density.

The visual angle is a function of the physical size of the percept and its distance to the observer. Both parameters affect the evaluation of the importance of the percept (Teghtsoonian and Frost, 1982). The cognitive representation of a percept is size-invariant (Ittelson, 1960; Kolers et al., 1985; Biederman and Cooper, 1992). Hence the size of a percept relates to the physical distance of the observer to the object (Loftus and Harley, 2005). Perceivers link changes in percept size to approaching and receding movements. Approaching movements intensify affective responses compared to receding movements (Mühlberger et al., 2008). This effect even holds if participants just imagined changes in stimulus size (Davis et al., 2011). Participants also represented distal events in a more abstract, schematic way (Fujita et al., 2006; Henderson et al., 2006; Trope and Liberman, 2010). Increasing the perceived physical distance to aversive stimuli reduces their perceived threat (Mobbs et al., 2007; Williams and Bargh, 2008; Blanchard et al., 2004). In line with this reasoning, increasing the size of photorealistic images intensifies subjective arousal and dominance ratings (Codispoti and De Cesarei, 2007; Reeves et al., 1999). Hence this effect of size variation might be due to perceived proximity instead of visual angle. In this study, we analyze the effect of visual angle independent of perceived proximity.

Furthermore, spatial information density may confound the emotional effects of stimulus size variation. Previous studies about the effect of stimulus sizes use naturalistic stimuli (e.g., the International Affective Picture System (Lang et al., 1999)). Size reduction of such naturalistic stimuli increases spatial information density: less space displays the same information. This densification increases the difficulty of discriminating content (Loftus and Harley, 2005). Discrimination of content, in turn, is necessary for a content-specific emotional response. The variation of spatial information density might explain the variation of emotional responses to varying stimulus sizes: De Cesarei and Codispoti (2008) reports that reducing fine-grained details in constant-sized images modulates emotional response in the same way as size variation does. Moreover, a temporal variation of spatial information density might affect emotional responses, though the empirical results are incoherent (see De Cesarei and Codispoti (2013) for a review). Hence we assume that information density confounds emotional responses to the size variation of naturalistic stimuli.

A fixed visual angle defines two states: inside and outside the percept. These separated areas inevitably define a boundary. To operationalize visual angle, we used monochrome rectangles comprising the minimal set of perceptual features: a spatial boundary defined by contrasting colors. We assume that the perception of monochrome rectangles of varying size dissociates the effects of spatial information density from the visual angle. Monochrome rectangles reduce the perceptual features that facilitate object recognition to a minimum (object shape, surface details, three-dimensional shading, texture, and object coloring) (Rossion and Pourtois, 2004; Tanaka et al., 2001). We assumed that monochrome rectangles aggravate a three-dimensional object representation and, consequently, their positioning in space. Color information seems independent of object perception. The recognition of the emotional scenery information is not affected by color information (Codispoti et al., 2012; Bradley et al., 2001, 2003; Sabatinelli et al., 2007).

1.2. Contribution

We tested the effect of visual angle on the emotional perception of audiovisual features. Monochrome color rectangles operationalized visual angle. We varied visual content features among different lightness values of the stimulus color. We operationalized auditory content features as different pitches of pure tones. Participants perceived these tones simultaneously with the visual stimuli. Participants then rated their subjective valence, arousal, and dominance response. We assumed that the visual angle determines emotional responses, independent of object recognition processes. Furthermore, we assumed that the visual angle moderates the emotional effect of visual and auditory content features. For the high visual angle condition, we observed that increased arousal decreased valence and dominance perception, and that visual content features modulated arousal. For the low visual angle condition, we found that pitch modulated arousal and that lightness affected valence. These results indicate that the visual angle weights the emotional relevance of perception modalities. A higher visual angle increases responsiveness to visual features and decreases responsiveness to non-visual features. These findings inform the design of new human–machine interaction techniques that incorporate visual information transfer and emotional processing.

2. Method

2.1. Participants

Undergraduate students (143, 118 women) from the University of Würzburg volunteered to participate in the experiment. All participants provided written informed consent before participation. They received course credit for participation. All participants reported normal or corrected to normal vision and normal hearing. We excluded 7 participants from the analysis for whom technical problems prevented a correct stimulus presentation. The final sample size consisted of $N = 136$ participants (113 women), with age ranging from 18 to 28 years

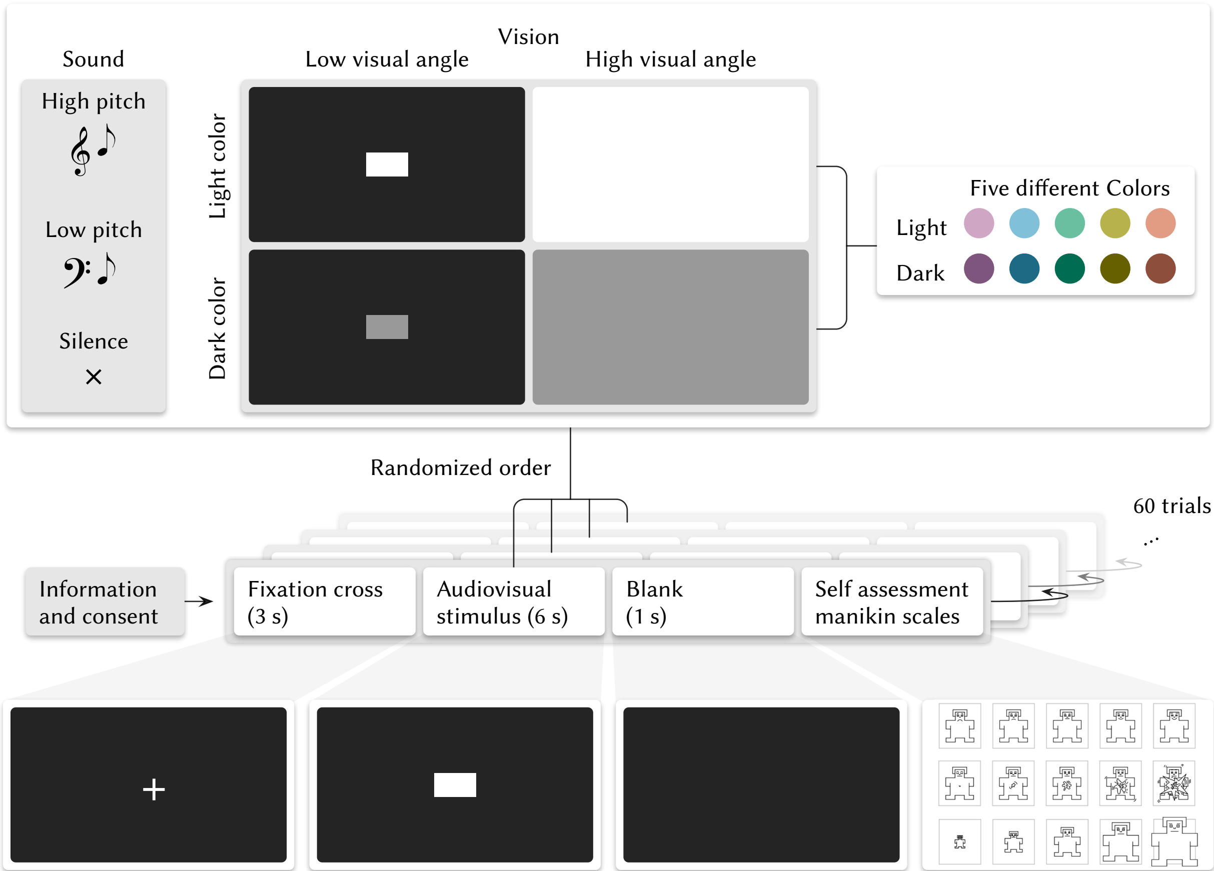

Figure 1: **Procedure and stimuli**. Participants completed 60 trials in a repeated measures 2 × 2 × 3 factorial design. Each trial began with the presentation of a fixation cross. Then participants saw a monochrome rectangle and heard a sound. The rectangle was either small or filled the whole screen. The rectangle had one of five colors, either light or dark. Participants heard either a pure tone in a high or low pitch or no tone at all. We corrected pitches for equal-loudness. Afterward, participants reported valence, arousal, and dominance on self-assessment manikin scales.

($M = 20.85$, $SD = 1.86$). Participants were naive regarding the hypotheses of the experiment. This study received ethical approval from the institutional ethics committee.

2.2. *Design and procedure*

The experiment followed a repeated-measures 2 × 2 × 3 factorial design with the factors visual angle (low vs high), lightness (low vs high), and pitch (none vs low vs high). We repeated each of these 12 conditions with five hues. Each participant completed 60 trials in a balanced, randomized order. Trials began with the presentation of a fixation cross (3 s). Participants then simultaneously saw a monochrome rectangle and heard no or a pure tone for 6 s. The rectangle varied in lightness, visual angle, and hue. The pure tone varied in pitch. After a black display (1 s), participants self-assessed valence, arousal, and dominance on self-assessment manikin scales. Participants selected responses with a mouse and proceeded to the next trial by clicking a button. We instructed participants to report their emotional responses to the audio-visual stimulus. Figure 1 illustrates the experimental design.

2.3. *Experimental manipulation*

Stimuli consisted of the simultaneous presentation of a visual and an auditory cue for 6 s. Visual cues consisted of monochrome rectangles that varied in size, lightness, and hue. We used an RGB approximation of the Munsell color space to specify the presented colors. The Munsell color space aims to represent color in a psychophysically plausible way (Kuehni, 2002), consisting of three dimensions: value (light to dark), chroma (gray to colored) and hue (circular scale). We used the five principal hues red (10R), yellow (10Y), green (10G), blue (10B), and purple (10P). For the factor lightness, we used the Munsell value 7 for the high lightness and 4 for low lightness. Chroma was constant at 6. We approximated the resulting Munsell colors (hue value/chroma) with the following RGB values (in brackets): 10R 7/6: (225,156,134), 10Y 7/6: (184,177,84), 10G 7/6: (108,191,161), 10B 7/6: (130,192,216), 10P 7/6: (207,167,196), 10R 4/6: (141,78,62), 10Y 4/6: (102,95,6), 10G 4/6: (4,108,83), 10B 4/6: (34,106,131), and 10P 4/6: (125,86,123). In the low visual angle condition, the rectangle was 139 mm long and 78 mm high. For the focal viewing distance of 550 mm, this corresponds to 15° horizontal and 8° vertical visual angle. In the high visual angle condition, the rectangle comprised 930 mm × 523 mm (80° × 51°). In the high pitch condition, we presented a pure tone at 523.251 Hz (musical note C5) and 48 dB. In the low pitch condition, we presented a pure tone at 130.813 Hz (musical note C3) and 65 dB. We corrected the loudness of the high pitch tone for equal-loudness (Suzuki and Takeshima, 2004) to the low pitch tone by -17 dB. We obtained the correction value from a small pre-study ($N = 3$). It lies within the expected range (Suzuki and Takeshima, 2004). We assume that the perceived loudness of the two tones was approximately equal. We sampled both tones with 48 kHz.

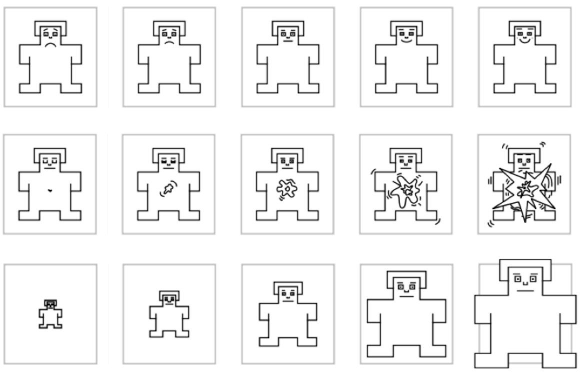

Figure 2: **Self-assessment manikin scales**.We assessed self-reported valence (top), arousal (center), and dominance (bottom) immediately after stimulus presentation. We labeled the self-assessment manikin scales with a 9-point Likert scale from 1 (*low/negative*) to 9 (*high/positive*).

2.4. Measures

We used the self-assessment manikin scales (Bradley and Lang, 1994) as primary outcome measures. Figure 2 depicts the self-assessment manikin scales. Self-assessment manikin scales allow non-verbal pictorial assessment of self-reported affective experience immediately after stimulus presentation. We used self-assessment manikin scales with five pictures, labeled with a 9-point Likert scale from 1 (*low/negative*) to 9 (*high/positive*). This measure assumes the conceptualization of emotion as three independent dimensional bipolar factors: valence, arousal, and dominance. Osgood (1952) and Russell and Mehrabian (1977) Valence conceptualizes approach or avoidance tendencies. Arousal conceptualizes the perceived level of physiological activity. Dominance conceptualizes the perceived level of control. Before the experiment, we described the self-assessment manikin scales to the participants, as proposed by Lang et al. (1999) Dimensional self-reports about affective experiences that are made directly after an emotion-eliciting event have reasonable validity (Mauss and Robinson, 2009). The validity and reliability of the self-assessment manikin scales are reasonable (Bradley and Lang, 1994). In general, dominance is considered the least sensitive scale among the three and seems to correlate positively with valence (Bradley and Lang, 1994; Warriner et al., 2013; Russell, 1979). We collected additional questionnaire measures after the experiment, which we did not include in this report.

2.5. Apparatus

We used a standard PC for stimulus presentation and response registration. A 42 in (106.68 cm) LCD screen (NEC MultiSync V422) displayed visual stimuli against a black background in a darkened room. Before the experiment, we color-calibrated the screen with a colorimeter (Datacolor Spyder 5 ELITE). We presented auditory stimuli with headphones (Sennheiser HD 201). During the experiment, a chin rest supported the head of the participants. The distance between the eyes of the participants and the monitor was 55 cm. Participants used a mouse to select responses. Figure 3 illustrates the apparatus.

2.6. Statistical analysis

Valence, arousal, and dominance ratings consist of Likert scale data, that approximate interval measurement (Boone and Boone, 2012). We aggregated ratings over five hue categories. For each

Figure 3: **Apparatus**. Participants sat in front of a screen at a fixed viewing distance. A chin rest supported their heads. Visual stimuli comprised the full screen in the high visual angle condition. In the low visual angle condition, the visual stimuli comprised only a fraction of the screen. Participants answered questions with a mouse. We presented auditory stimuli in headphones.

dependent variable, we then applied a repeated-measures analysis of variance (ANOVA) to analyze for main and first-order interaction effects of visual angle, lightness, and pitch. When necessary, we Greenhouse–Geisser corrected degrees of freedom. To achieve a global alpha level of 5%, we Bonferroni–Holm adjusted the significance thresholds of the 18 ANOVA tests for multiple comparisons. We report generalized η^2 (η_g^2) as a measure of effect size. Post hoc, we pairwise contrasted the levels for significant interaction effects. We Bonferroni–Holm adjusted the significance thresholds of the comparisons to the 18 a priori tests and the 27 post hoc tests. We used R (R Core Team, 2018) and the afex package (Singmann et al., 2018) to analyze the data. Data and code for all analyses are available at https://doi.org/10.17605/OSF.IO/CTU4G.

3. Results

We conducted a repeated-measures experiment with three factors. In each trial, we presented a monochrome light- or dark-colored rectangle either covering a low or high visual angle. Simultaneously, we presented none, a low, or a high pitch tone (corrected for equal-loudness). Afterward, participants reported subjective valence, dominance, and arousal ratings. We excluded 7 participants from analysis due to technical problems during the experiment. We included all remaining $N = 136$ participants (113 women) into the analysis. Each participant conducted 60 trials. We applied univariate repeated measures ANOVAs. We Bonferroni–Holm adjusted the alpha levels of the 18 tests for main and first-order interaction effects to a global alpha level of 5%. We report generalized η^2 (η_g^2) as a measure of effect size.

Visual angle had a significant main effect on valence ($F(1,\ 135) = 12.55$, $p < .001$, $\eta_g^2 < .01$), arousal ($F(1, 135) = 112.93$, $p < .001$, $\eta_g^2 = .05$), and dominance ($F(1, 135) = 43.19$, $p < .001$, $\eta_g^2 = .05$). Lightness had a significant main effect on valence ($F(1, 135) = 8.66$, $p = .004$, $\eta_g^2 < .01$) and arousal ($F(1, 135) = 18.60$, $p < .001$, $\eta_g^2 < .01$). Pitch had a significant main effect on valence ($F(2, 270) = 80.22$, $p < .001$, $\eta_g^2 = .12$), arousal ($F(2, 270) = 167.21$, $p < .001$, $\eta_g^2 = .19$), and dominance ($F(1.74, 235.10) = 54.41$, $p < .001$, $\eta_g^2 = .07$). Table 2 summarizes the estimated marginal means of significant main effects.

Visual angle and lightness showed a significant interaction effect on valence ($F(1, 135) = 11.67$, $p < .001$, $\eta_g^2 < .01$) and arousal ($F(1, 135) = 19.77$, $p < .001$, $\eta_g^2 < .01$). Visual angle and pitch also showed a significant interaction effect on arousal ($F(2, 270) = 6.73$, $p = .001$, $\eta_g^2 < .01$). Table 1 summarizes the p and η_g^2-values of the main and interaction effects. Figure 4 shows two-way interaction plots of the significant interactions.

Post hoc, we used pairwise contrasts to compare estimated marginal means for combinations of visual angle with lightness and pitch levels. We Bonferroni–Holm adjusted alpha levels of comparisons to a total number of 45 tests (27 post hoc and 18 ANOVA tests). Estimated marginal means averaged over pitch significantly differed pairwise ($p < .001$) for arousal; except for the pairing low visual angle and high lightness compared to low visual angle and lightness ($p = .027$). Estimated marginal means averaged over pitch significantly differed pairwise ($p < .001$) for valence; except for the following pairings: high visual angle and lightness compared to low lightness with high ($p = .415$) as well as low visual angle ($p = .640$); high visual angle and low lightness compared to low visual angle and lightness ($p = .044$). Estimated marginal means

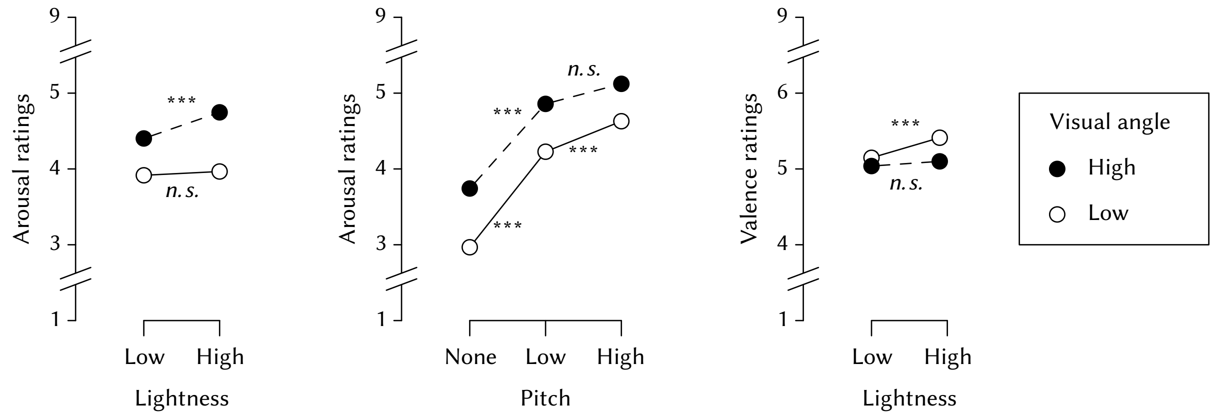

Figure 4: **Interaction plots for significant first-order interaction effects**. Dots indicate estimated marginal means of arousal and valence ratings for low (white) and high visual angle (black) as a function of lightness or pitch levels. Arousal and valence ratings range from 1 (*low/negative*) to 9 (*high/positive*). Along lines, p-values indicate results of pairwise mean comparisons. Abbreviations: *n. s.*: not significant (Bonferroni–Holm adjusted), ***: $p < .001$.

Table 1: **Main and interaction effects**. Summary of p and η_g^2-values for main and first order interaction effects on valence, arousal, and dominance ratings. Abbreviations: *n. s.*: not significant (Bonferroni–Holm adjusted), ***: $p < .001$

	Valence		Arousal		Dominance	
	p	η_g^2	p	η_g^2	p	η_g^2
Visual angle	***	$< .01$	***	.05	***	.05
Lightness	.004	$< .01$	***	$< .01$	.039	*n. s.*
Pitch	***	.12	***	.19	***	.07
Visual angle × lightness	***	$< .01$	***	$< .01$	.253	*n. s.*
Visual angle × pitch	.489	*n. s.*	.001	$< .01$	.235	*n. s.*
Lightness × pitch	.458	*n. s.*	.682	*n. s.*	.519	*n. s.*

averaged over lightness significantly differed pairwise ($p < .001$) for arousal; except for the following pairings: high visual angle and pitch compared to high visual angle and low pitch ($p = .006$); low visual angle and high pitch compared to high visual angle and low pitch ($p = .023$).

4. Discussion

We tested the effects of visual angle on emotional responses to audiovisual stimuli. Stimuli consisted of minimal spatial features to not induce three-dimensional object representations. Higher visual angle increased arousal and decreased valence and dominance during stimulus exposure. Only in the high visual angle conditions did visual content features (lightness) modulate arousal. Only in the low visual angle conditions did auditory features (pitch) modulate arousal, and did lightness affect valence responses. Arousal indicates the strength of emotions in terms of felt physiological activation (Barrett and Russell, 1999). These results indicate that the visual angle weights the emotional relevance of perception modalities: Visual features have higher emotional relevance when the visual angle is high, whereas auditory features have higher emotional relevance when the visual angle is low. This processing step does not require spatial object representations. We conclude that visual angle serves as an early-stage perceptual feature for organizing emotional responses to audiovisual stimuli.

This model extends previous findings that investigate the impact of visual angle in the perception of motivational-relevant spatial objects. Previous studies used photorealistic stimuli. Photorealistic stimuli induce higher emotional response when presented in a large visual angle (Codispoti and De Cesarei, 2007; Reeves et al., 1999). The visual angle of photorealistic stimuli modulates arousal and valence responses to stimuli content (e.g., erotic couples,

Table 2: **Marginal means**. Estimated marginal means and standard errors (SE) of significant main effects. All main effects were significant, except the main effect of lightness on dominance.

		Valence		Arousal		Dominance	
Factor	Level	Mean	SE	Mean	SE	Mean	SE
Visual angle	Low	5.28	0.08	3.94	0.09	5.56	0.10
	High	5.07	0.08	4.57	0.10	4.90	0.10
Lightness	Low	5.09	0.07	4.16	0.09		
	High	5.26	0.08	4.36	0.09		
Pitch	None	5.78	0.08	3.35	0.09	5.77	0.11
	Low	4.92	0.08	4.54	0.11	4.98	0.10
	High	4.83	0.09	4.88	0.11	4.94	0.10

mutilated bodies) (Codispoti and De Cesarei, 2007; Reeves et al., 1999). Photorealistic stimuli induce motivational response patterns (e.g., sexual arousal or threat). Two factors moderate these responses: the perceived distance to the stimulus; and the ability to distinguish the stimuli content. Variation of the visual angle of photorealistic stimuli confounds both. Visual angle determines object size and spatial information density. Our design minimized the effect of these confounders. Hence we attribute the observed gate-keeper effect to the visual angle independent of object recognition processes.

In our study, participants saw the edges of a smaller or larger rectangle. By definition, a visual angle separates two different areas: the area that lies inside the visual angle and the area that lies outside the visual angle. Edges, in turn, constitute at least a two-dimensional object. Hence by definition, it is not possible to fully dissociate visual angle from object recognition. However, the present study decreases the impact of object recognition to a minimum. Two-dimensional geometric forms have little association with characteristics of functionality. The less realistic, three-dimensional, and graspable an object appears to be, the less strong its physical affordance is (Symes et al., 2007). Hence we assume a limited impact of object recognition on our results.

Dominance ratings in our study differ from studies that used photorealistic stimuli. Perceiving the counterpart as dominant in general decreases the own felt dominance. In line with this assumption, increased visual angle decreased dominance ratings in our study. In a previous study, however, participants reported increased dominance for an increased visual angle, when they saw photorealistic stimuli (Detenber and Reeves, 1996). Hence we hypothesize that the presence of objects in the visual percept can invert the effect of visual angle on dominance. This observation supports the assumption that object recognition played a minor role in the presented study. Moreover, it supports the hypothesis that dominance is a cognitive construal of affect states (Barrett and Russell, 1999).

The stimuli in our study, in general, elicited low engagement and neutral feelings (moderate mean ratings for valence, arousal, and dominance). This result reflects the use of minimalistic stimuli (monochrome squares with constant, pure tones) that have minimal motivational relevance. Naturalistic scenes are more engaging (De Cesarei et al., 2017), but also induce more complex responses that could contain confounders. However, the observed effects were present even when contrasting low engaging stimuli. We hypothesize that amplified effect sizes occur in controlled settings with higher engaging stimuli.

4.1. Design implications for human–computer interfaces

In this study, we show that the visual angle modulates the emotional responses to audiovisual information. Future experiments need to assess how this effect generalizes to naturalistic applications. We assume that the observed modulation effect informs the design of human–computer interfaces as follows.

4.1.1. Sensory quantity affects emotional response

Our results support the assumption that the degree of immersion modulates arousal responses to stimuli. Immersion here refers to the amount of sensory information an interface delivers (Slater, 2003). We increased immersion in two ways: first, we increased the visual angle, second we augmented visual with auditive stimuli. Both modifications increased arousal responses. Thus we assume that visual angle and multisensory augmentation provide means to modulate emotional responses to media. This result is in line with previous findings from naturalistic scenarios. For example, users reported higher enjoyment, excitement, and more physical arousal watching movies on large compared to small screens (Lombard et al., 2000; Kim et al., 2011) Users engage in more heuristic, affective processing when watching stimuli on larger screens compared to more systematic, cognitive processing when watching stimuli on small screens (Kim and Sundar, 2016; Kim, 2017). Video games (Thompson et al., 2012) and movies (Rigby et al., 2016) are more engaging on large screens. Movies and pictures with music increase emotional processing compared to stimuli without music (Baumgartner et al., 2006a,b). Our results support the assumption that human–computer interfaces can make use of visual angle and multisensory augmentation to control arousal responses. Emotional responses are, for example, relevant if the interface demands fast decisions or high situational awareness from the users. Decision support systems for medical treatments or the surveillance of critical security systems as in autonomous driving provide such applications. In such use cases, adaptive control of visual angle and augmented audio would provide means to increase arousal responses if appropriate. Controlling arousal, in turn, allows modulating further attention processes, such as short-term memory performance (Maljkovic and Martini, 2005). Innovations in augmented and virtual reality, for example, as in-car displays, increasingly provide means to timely adapt visual angle during security-critical scenarios.

While the visual angle modulates engagement, the visual angle also affects performance on visual tasks. If users can turn their heads towards a target stimuli, an increased visual angle increases navigation and search performance (Arthur and Brooks Jr, 2000). If, however, users have a fixed center of view, their ability to perform tasks on their peripheral visual fields is limited (see Strasburger et al. (2011) for a review): For example, users react slower to peripheral compared to central stimuli. They also have more difficulties in detecting patterns and changes in peripherical compared to central stimuli. Future studies, therefore, need to investigate how altering the visual angle can help balance emotional engagement with task performance.

4.1.2. The visual angle modulates the affective processing of visual features

The arousal response to visual content features (here lightness) only changed significantly if the visual angle was high. This interaction suggests that the emotional processing of visual features requires a sufficiently high visual angle. This finding is consistent with naturalistic studies in which an increase in screen size intensified responses to arousing (Codispoti and De Cesarei, 2007) or pleasant images (Reeves et al., 1999). Human–computer interfaces can use this interaction to reduce or amplify emotional responses to visual content features. For example, a low visual angle might reduce the adverse emotional effects of violent content.

4.1.3. The visual angle modulates the affective processing of auditory features

The arousal response to auditory content features (here pitch) only changed significantly if the visual angle was low. This interaction suggests that a low visual angle facilitates the emotional processing of auditory content features. This finding is in line with findings, that suggest that auditory attention is inversely related to visual engagement: For example, Cate et al. (2009) suggests that auditory cues direct attention to the far peripheral view, away from central visual cues. Furthermore, auditory cues increase performance for interfaces with small buttons (Brewster, 2002). Human–computer interfaces can use this interaction to reduce or amplify emotional responses to auditory content features by modulating the visual angle. For example, medical monitoring tasks can benefit from an efficient balancing of auditory and visual information processing (e.g., Klueber et al. (2019)).

4.2. Conclusion

The visual angle defines the total amount of available visual information. Our results indicate that visual information availability modulates emotional responses on early processing steps. High visual information availability increases emotional response and increases the relevance of visual content information. Low visual information availability increases the relevance of other modalities. We suggest that this effect is predominantly independent of object recognition processes. These findings inform developments of new human–machine interaction techniques that incorporate visual information transfer and emotional processing. It seems critical to be aware of this causal relation for building better interfaces. Control of the presentation layer, i.e., the quantity of perceivable information by visual angle, will help avoid unintended emotional responses and provide a means to provoke emotional effects when desired.

Acknowledgments

The German Federal Ministry of Education and Research supported this work (project 16SV7322). This paper is part of the first author's doctoral thesis.

References

Arthur, K. W. and Brooks Jr, F. P. (2000). *Effects of field of view on performance with head-mounted displays*. PhD thesis, University of North Carolina at Chapel Hill.

Barrett, L. F. and Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current directions in psychological science*, 8(1):10–14.

Baumgartner, T., Esslen, M., and Jäncke, L. (2006a). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International journal of psychophysiology*, 60(1):34–43.

Baumgartner, T., Lutz, K., Schmidt, C. F., and Jäncke, L. (2006b). The emotional power of music: how music enhances the feeling of affective pictures. *Brain research*, 1075(1):151–164.

Biederman, I. and Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1):121.

Blanchard, E. B., Kuhn, E., Rowell, D. L., Hickling, E. J., Wittrock, D., Rogers, R. L., Johnson, M. R., and Steckler, D. C. (2004). Studies of the vicarious traumatization of college students by the september 11th attacks: Effects of proximity, exposure and connectedness. *Behaviour research and therapy*, 42(2):191–205.

Boone, H. N. and Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, 50(2):1–5.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. (2001). Emotion and motivation i: defensive and appetitive reactions in picture processing. *Emotion*, 1(3):276.

Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., and Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral neuroscience*, 117(2):369.

Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing*, 6(3):188–205.

Cate, A. D., Herron, T. J., Yund, E. W., Stecker, G. C., Rinne, T., Kang, X., Petkov, C. I., Disbrow, E. A., and Woods, D. L. (2009). Auditory attention activates peripheral visual cortex. *PloS one*, 4(2):e4645.

Chapados, C. and Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, 108(3):639–651.

Codispoti, M. and De Cesarei, A. (2007). Arousal and attention: Picture size and emotional reactions. *Psychophysiology*, 44(5):680–686.

Codispoti, M., De Cesarei, A., and Ferrari, V. (2012). The influence of color on emotional perception of natural scenes. *Psychophysiology*, 49(1):11–16.

Cox, T. J. (2008). The effect of visual stimuli on the horribleness of awful sounds. *Applied acoustics*, 69(8):691–703.

Davis, J. I., Gross, J. J., and Ochsner, K. N. (2011). Psychological distance and emotional experience: what you see is what you get. *Emotion*, 11(2):438.

De Cesarei, A. and Codispoti, M. (2008). Fuzzy picture processing: Effects of size reduction and blurring on emotional processing. *Emotion*, 8(3):352.

De Cesarei, A. and Codispoti, M. (2013). Spatial frequencies and emotional perception. *Reviews in the Neurosciences*, 24(1):89–104.

De Cesarei, A., Loftus, G. R., Mastria, S., and Codispoti, M. (2017). Understanding natural scenes: Contributions of image statistics. *Neuroscience & Biobehavioral Reviews*, 74:44–57.

De Cesarei, A., Mastria, S., and Codispoti, M. (2013). Early spatial frequency processing of natural images: an erp study. *PloS one*, 8(5):e65103.

De Cesarei, A., Peverato, I. A., Mastria, S., and Codispoti, M. (2015). Modulation of early erps by accurate categorization of objects in scenes. *Journal of vision*, 15(8):14–14.

Detenber, B. H. and Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication*, 46(3):66–84.

Fei-Fei, L., VanRullen, R., Koch, C., and Perona, P. (2005). Why does natural scene categorization require little attention? exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition*, 12(6):893–924.

Fujita, K., Henderson, M. D., Eng, J., Trope, Y., and Liberman, N. (2006). Spatial distance and mental construal of social events. *Psychological Science*, 17(4):278–282.

Henderson, M. D., Fujita, K., Trope, Y., and Liberman, N. (2006). Transcending the" here": the effect of spatial distance on social judgment. *Journal of personality and social psychology*, 91(5):845.

Ittelson, W. H. (1960). *Visual space perception.* Springer, Oxford, England.

Junghöfer, M., Bradley, M. M., Elbert, T. R., and Lang, P. J. (2001). Fleeting images: a new look at early emotion discrimination. *Psychophysiology*, 38(2):175–178.

Kaya, N. and Epps, H. (2004). Relationship between color and emotion: A study of college students. *College Student J*, 38(3):396.

Kim, K. J. (2017). Shape and size matter for smartwatches: effects of screen shape, screen size, and presentation mode in wearable communication. *Journal of Computer-Mediated Communication*, 22(3):124–140.

Kim, K. J. and Sundar, S. S. (2016). Mobile persuasion: Can screen size and presentation mode make a difference to trust? *Human Communication Research*, 42(1):45–70.

Kim, K. J., Sundar, S. S., and Park, E. (2011). The effects of screen-size and communication modality on psychology of mobile device users. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1207–1212. ACM.

Klueber, S., Wolf, E., Grundgeiger, T., Brecknell, B., Mohamed, I., and Sanderson, P. (2019). Supporting multiple patient monitoring with head-worn displays and spearcons. *Applied ergonomics*, 78:86–96.

Kolers, P. A., Duchnicky, R. L., and Sundstroem, G. (1985). Size in the visual processing of faces and words. *Journal of Experimental Psychology: Human Perception and Performance*, 11(6):726.

Kuehni, R. G. (2002). The early development of the munsell system. *Color Research & Application*, 27(1):20–27.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1999). International affective picture system (iaps): Technical manual and affective ratings. *Gainesville, FL: The Center for Research in Psychophysiology, University of Florida*, 2.

Li, F. F., VanRullen, R., Koch, C., and Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, 99(14):9596–9601.

Loftus, G. R. and Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review*, 12(1):43–65.

Logeswaran, N. and Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neuroscience letters*, 455(2):129–133.

Lombard, M., Reich, R. D., Grabe, M. E., Bracken, C. C., and Ditton, T. B. (2000). Presence and television. the role of screen size. *Human Communication Research*, 26(1):75–98.

Maljkovic, V. and Martini, P. (2005). Short-term memory for scenes with affective content. *Journal of vision*, 5(3):6–6.

Marin, M. M., Gingras, B., and Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion*, 12(3):618.

Mauss, I. B. and Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*, 23(2):209–237.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., Dolan, R. J., and Frith, C. D. (2007). When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science*, 317(5841):1079–1083.

Mühlberger, A., Neumann, R., Wieser, M. J., and Pauli, P. (2008). The impact of changes in spatial distance on emotional responses. *Emotion*, 8(2):192.

Oliva, A. and Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in brain research*, 155:23–36.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological bulletin*, 49(3):197.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reeves, B., Lang, A., Kim, E. Y., and Tatar, D. (1999). The effects of screen size and message content on attention and arousal. *Media Psychology*, 1(1):49–67.

Rigby, J. M., Gould, S., Brumby, D., and Cox, A. (2016). Watching movies on netflix: investigating the effect of screen size on viewer immersion. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI 2016*, pages 714–721. ACM.

Rossion, B. and Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2):217–236.

Russell, J. A. (1979). Affective space is bipolar. *Journal of personality and social psychology*, 37(3):345.

Russell, J. A. and Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, 11(3):273–294.

Sabatinelli, D., Bradley, M. M., Lang, P. J., Costa, V. D., and Versace, F. (2007). Pleasure rather than salience activates human nucleus accumbens and medial prefrontal cortex. *Journal of neurophysiology*, 98(3):1374–1379.

Singmann, H., Bolker, B., Westfall, J., and Aust, F. (2018). *afex: Analysis of Factorial Experiments*. R package version 0.20-2.

Slater, M. (2003). A note on presence terminology. *Presence connect*, 3(3):1–5.

Strasburger, H., Rentschler, I., and Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13–13.

Suzuki, Y. and Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, 116(2):918–933.

Symes, E., Ellis, R., and Tucker, M. (2007). Visual object affordances: Object orientation. *Acta psychologica*, 124(2):238–255.

Tanaka, J., Weiskopf, D., and Williams, P. (2001). The role of color in high-level vision. *Trends in cognitive sciences*, 5(5):211–215.

Teghtsoonian, R. and Frost, R. O. (1982). The effects of viewing distance on fear of snakes. *Journal of behavior therapy and experimental psychiatry*, 13(3):181–190.

Thompson, M., Nordin, A. I., and Cairns, P. (2012). Effect of touch-screen size on game immersion. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, pages 280–285. British Computer Society.

Thompson, W. F., Russo, F. A., and Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, 22(8):1457–1470.

Torralba, A., Oliva, A., Castelhano, M. S., and Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological review*, 113(4):766.

Trope, Y. and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review*, 117(2):440.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., and Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, 118(2):157–170.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., and Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1):80–113.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, 9(12):585–594.

Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior Research Methods (Online)*, 45(4):1191.

Williams, L. E. and Bargh, J. A. (2008). Keeping one's distance: The influence of spatial distance cues on affect and evaluation. *Psychological Science*, 19(3):302–308.