



Change Blindness Is Influenced by Both Contrast Energy and Subjective Importance within Local Regions of the Image

Wietske Zuiderbaan^{1*}, Jonathan van Leeuwen^{1,2} and Serge O. Dumoulin^{1,2,3}

¹ Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands, ² Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, ³ Spinoza Centre for Neuroimaging, Amsterdam, Netherlands

OPEN ACCESS

Edited by:

Britt Anderson,
University of Waterloo, Canada

Reviewed by:

Iris I. A. Groen,
National Institutes of Health,
United States
Meera Mary Sunny,
Indian Institute of Technology
Gandhinagar, India

*Correspondence:

Wietske Zuiderbaan
wietske.zuiderbaan@gmail.com

Specialty section:

This article was submitted to
Perception Science,
a section of the journal
Frontiers in Psychology

Received: 23 January 2017

Accepted: 19 September 2017

Published: 04 October 2017

Citation:

Zuiderbaan W, van Leeuwen J and
Dumoulin SO (2017) Change
Blindness Is Influenced by Both
Contrast Energy and Subjective
Importance within Local Regions
of the Image. *Front. Psychol.* 8:1718.
doi: 10.3389/fpsyg.2017.01718

Our visual system receives an enormous amount of information, but not all information is retained. This is exemplified by the fact that subjects fail to detect large changes in a visual scene, i.e., change-blindness. Current theories propose that our ability to detect these changes is influenced by the gist or interpretation of an image. On the other hand, stimulus-driven image features such as contrast energy dominate the representation in early visual cortex (De Valois and De Valois, 1988; Boynton et al., 1999; Olman et al., 2004; Mante and Carandini, 2005; Dumoulin et al., 2008). Here we investigated whether contrast energy contributes to our ability to detect changes within a visual scene. We compared the ability to detect changes in contrast energy together with changes to a measure of the interpretation of an image. We used subjective important aspects of the image as a measure of the interpretation of an image. We measured reaction times while manipulating the contrast energy and subjective important properties using the change blindness paradigm. Our results suggest that our ability to detect changes in a visual scene is not only influenced by the subjective importance, but also by contrast energy. Also, we find that contrast energy and subjective importance interact. We speculate that contrast energy and subjective important properties are not independently represented in the visual system. Thus, our results suggest that the information that is retained of a visual scene is both influenced by stimulus-driven information as well as the interpretation of a scene.

Keywords: change detection, contrast energy, subjective importance, natural images, scene perception

INTRODUCTION

Intuitively, our visual representation of the outside world appears to be highly detailed, however, the change blindness paradigm highlights limitations of this visual representation (Rensink, 2000; Martin et al., 2001). The change blindness paradigm reveals our inability to see changes in two sequentially presented visual scenes when separated by a disruption like a saccade or a flicker (for review see: Rensink, 2002; Simons and Ambinder, 2005). Without the disruption, the change is often easy to detect. The disruption highlights the limits of our ability to retain and compare information from one visual scene to the other (Simons and Rensink, 2005).

Some changes in the visual scene are easier to detect than others. Many change blindness studies focus on the notion that changes made in parts of a scene that represent the general interpretation of the image are detected faster (Rensink et al., 1997; O'Regan et al., 1999; Shore and Klein, 2000; Stirk and Underwood, 2007; Sampanes et al., 2008). Some studies refer to the general interpretation as regions of high interest (Rensink et al., 1997; O'Regan et al., 1999; Shore and Klein, 2000; Verma and McOwan, 2010), where others refer to a semantic summary of the scene (gist) that is related to our knowledge of the world (Stirk and Underwood, 2007; Sampanes et al., 2008).

What causes these changes to be detected faster? In change blindness, subjects have to compare two image representations. Working memory is supposed to play an important role in this comparison process. One interpretation is that parts of a scene related to the general interpretation receive more attention, and are therefore more likely to be encoded and compared (Simons and Rensink, 2005). On the other hand, we know that the early visual system, i.e., primary visual cortex, encodes stimulus-driven image properties such as contrast energy (De Valois and De Valois, 1988; Boynton et al., 1999; Olman et al., 2004; Mante and Carandini, 2005; Dumoulin et al., 2008).

Some studies controlled for stimulus-driven image properties, such as stimulus, size, brightness, color, and saliency (Rensink et al., 1997; Kelley et al., 2003). These studies demonstrate that the effect of change blindness survives matching stimulus-driven properties. Yet, none of these studies considered contrast energy, the most well known computation of the early visual system. Contrast energy is not the same as saliency. Whereas contrast energy is a component of saliency, saliency is influenced by many other stimulus-driven properties. Previous change blindness studies did not always account for differences in stimulus-driven image properties (Verma and McOwan, 2010) and stimulus-driven properties might contribute to change detection, such as saliency or size of the change (Landman et al., 2003; Verma and McOwan, 2010; Spotorno et al., 2013). We predict that if change detection is founded on computations of the early visual system, contrast energy would contribute to our ability to detect changes.

Here, we ask whether contrast energy contributes to our ability to detect changes in a visual scene. We focus on contrast energy because this is one of the most well-described and most fundamental computations in early visual cortex that lies at the heart of visual computations (De Valois and De Valois, 1988). Furthermore, we compared change detection for these stimulus-driven image properties with the subjective image interpretation. Last, we investigated whether our subjective image interpretation interacts with the stimulus-driven image property contrast energy in our ability to detect changes to a visual scene.

To this aim, we measured reaction times (RTs) in a flicker-task using the change-blindness paradigm (Rensink et al., 1997); subjects indicated when and where they detected the change between the images. We used images from the Berkeley Segmentation Dataset and Benchmark database (Martin et al., 2001). In this dataset, Martin et al. (2001) asked human observers to identify the most important aspects of the image (**Figure 1B**). We used these manually labeled aspects of the scene to define

and quantify our measure for the subjective image interpretation. Using the combination of the natural images together with their manually labeled images, we were able to measure the amount of change that a manipulation to an image brings both to the stimulus-driven and to the subjective image interpretation. We manipulated the amount of change in subjective importance (manually labeled aspects of the scene) and in the change of the stimulus-driven image property contrast energy.

We found both significantly shorter RTs for manipulations dominated by a change in subjective importance, as well as those dominated by changes in contrast energy. Thus both subjective importance and contrast energy contribute to the speed of change detection. Furthermore, these two properties interacted: shortest RTs were found when the changes both contained a large change in contrast energy and in subjective importance. We show that this interaction effect is not explained by statistical facilitation alone, i.e., the faster RTs in the high-contrast/high-subjective importance condition are not explained by the added benefit of the presence of two independent properties (both contrast energy and subjective importance). These results show that our ability to detect changes in a visual scene is not only influenced by the high-level image interpretation of the image, but is also influenced by stimulus-driven image statistics such as contrast energy. Finally, our results that the interaction effect cannot be explained by statistical facilitation alone, suggest that the stimulus-driven and subjectively important image properties are not independently processed in the visual system, but interact with each other.

MATERIALS AND METHODS

Participants

In total, 60 subjects participated in the experiment (30 female, age range = [18–39], mean age = 23.9, $SD = 4.0$). The total number of 60 subjects was based on a power analysis informed by previous literature (Verma and McOwan, 2010). All subjects had normal or corrected-to-normal visual acuity. The study was approved by the local Ethics Committee of the Utrecht University and the experiments were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All experiments were performed with the informed written consent of the subjects. The duration of the experiment was approximately 30 min. The subjects participated either for course credits or a monetary reward.

Apparatus

The experiment was programmed in MatLab (MathWorks, United States), using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a 21-inch CRT monitor type LaCie-C22BW711 (60 Hz, 1024 × 768) using a Mac Pro 4.1 computer. The monitor was calibrated with a light meter type Gossen Mavo-Monitor USB. The viewing distance was 57 cm, which was maintained using a chin and forehead rest.

Stimuli

The images were taken from the 'Berkeley Segmentation Dataset and Benchmark' database (Martin et al., 2001). We used a

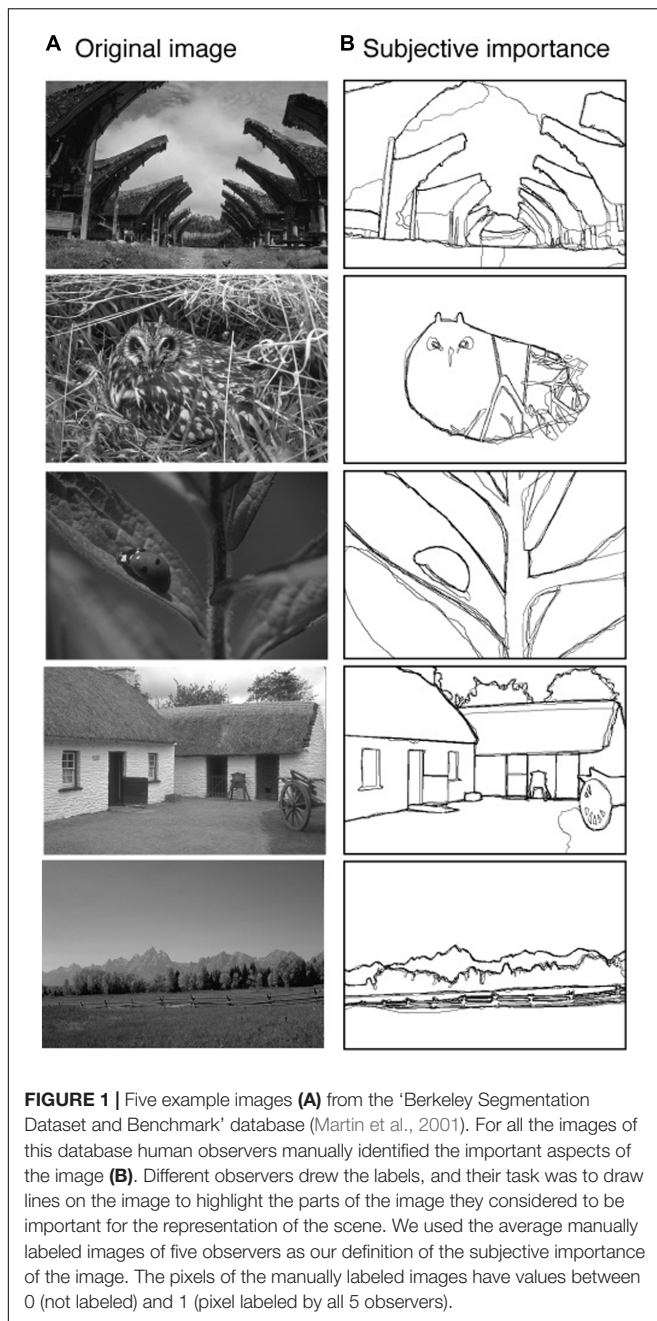


FIGURE 1 | Five example images (A) from the 'Berkeley Segmentation Dataset and Benchmark' database (Martin et al., 2001). For all the images of this database human observers manually identified the important aspects of the image (B). Different observers drew the labels, and their task was to draw lines on the image to highlight the parts of the image they considered to be important for the representation of the scene. We used the average manually labeled images of five observers as our definition of the subjective importance of the image. The pixels of the manually labeled images have values between 0 (not labeled) and 1 (pixel labeled by all 5 observers).

selection of the grayscale images and we only used the images that were in landscape orientation. Also, we did not include any images where a human was present in the image. Using Adobe Photoshop CS6 (Version 13.0, United States), we manipulated the image content. We made a large set (502) of different changes to the images, of which we selected 108 different changes, giving 108 independent image pairs in total to be used in the experiment. To balance the different conditions, these 108 image pairs were selected from the larger set of image pairs. These 108 images were selected so every condition contained the same number of images and the changes were not made in the central region of the image; no changes were made to pixels with an eccentricity

smaller than 1.6° of visual angle. Furthermore, we selected the images so there were no significant differences between the four categories in terms of eccentricity of the changes, changes in size, distance from the center, mean contrast, mean luminance change and spatial frequency change of the manipulated area. Also, multiple changes were made to the same image, but every subject only saw one change associated with one image. Changes to the images were for instance deletions of (part of) objects, or changes only in textures. **Figure 2** shows four example image pairs. The arrows indicate where the manipulation to the image was made. All stimuli were presented on a gray background. The size of the images was 481 pixels \times 321 pixels, extending $18.6^\circ \times 12.4^\circ$ of visual angle.

Measure of Change in Contrast Energy and Subjective Importance

The aim of the study was to investigate the role of the stimulus-driven image property contrast energy and the subjective image interpretation on our ability to detect changes in a visual scene. We simultaneously measured the influence of both; for every image manipulation we calculated how big the change in contrast energy was and its corresponding change in subjective importance. Where we investigated the individual role of both properties, this also allowed us to measure the interaction effect of contrast energy and subjective importance on change detection.

Calculation of Change in the Stimulus-Driven Image Property Contrast

To compute the amount of contrast energy change that the manipulation to an image brings, the first step was to calculate the local contrast energy values, both of the original image and of the manipulated image. For every pixel of the image we calculated the local contrast energy inside a spatial window of neighboring pixels. This spatial window is defined by a Gaussian weighting function:

$$w_i = \exp\left(-\frac{(x_i - x_c)^2 + (y_i - y_c)^2}{2(\sigma)^2}\right) \quad (1)$$

where (x_i, y_i) is the location of the i th pixel and (x_c, y_c) is the location of the pixel at the center of the image patch. σ is the standard deviation of the Gaussian window and defines the size of the spatial window. For the size of σ , we used 19.4 pixels, which corresponds to 0.75° of visual angle.

The local contrast energy value is based on the Root-Mean-Squared (RMS)-contrast (Pelli, 1997; Bex and Makous, 2002) which is defined as the standard deviation of the luminance intensities relative to the mean. The RMS-contrast is weighted by the Gaussian function to obtain the local contrast energy value per pixel:

$$\text{local_contrast_energy} = \sqrt{\frac{1}{\sum_{i=1}^N w_i} \sum_{i=1}^N w_i \frac{(L_i - L)^2}{L^2}} \quad (2)$$

Where w_i is the Gaussian weighting function. N is the number of pixels in the spatial window. L is the mean luminance from the

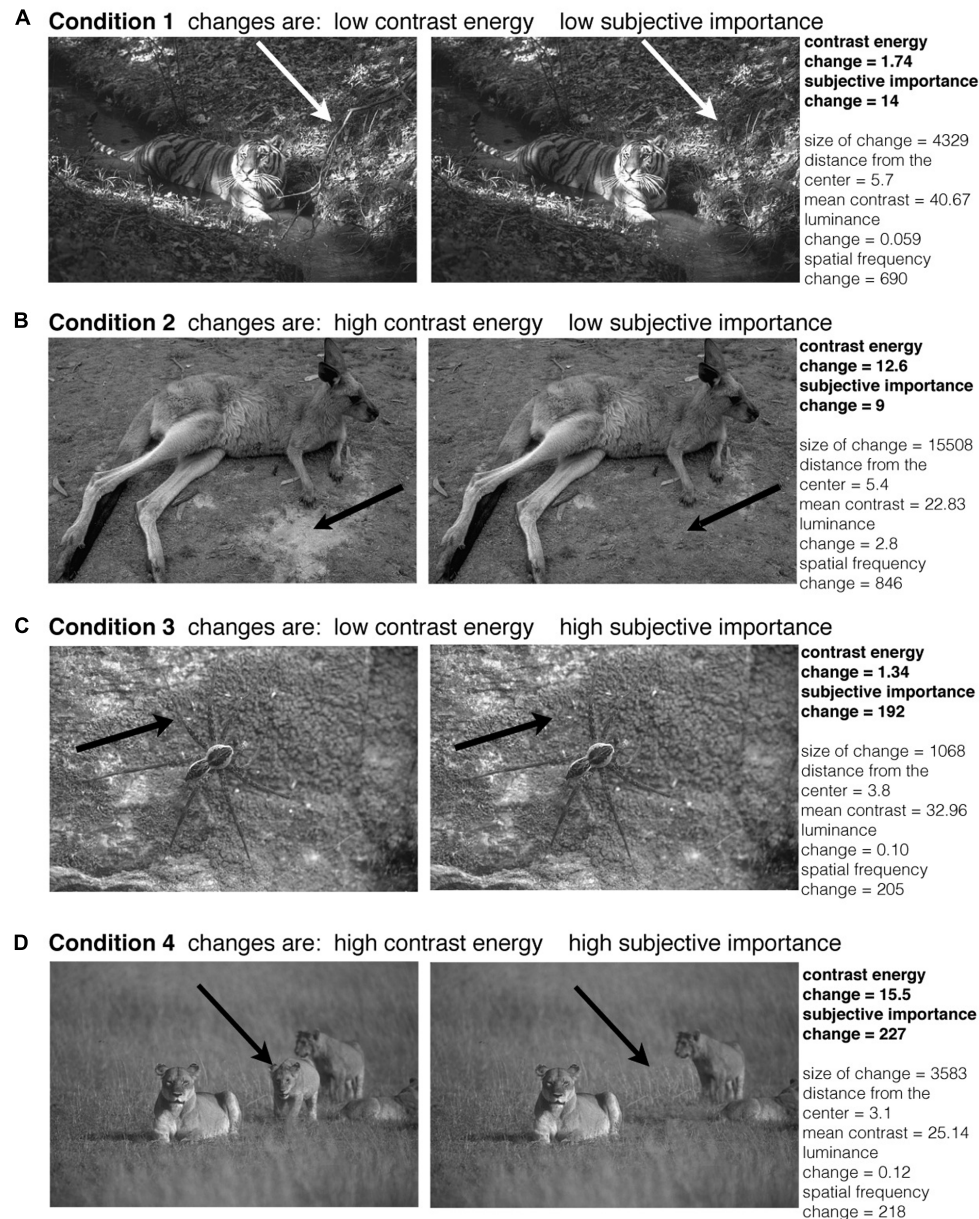


FIGURE 2 | Examples of image pairs that were used in the experiment. The original images were taken from the ‘Berkeley Segmentation Dataset and Benchmark’ database (Martin et al., 2001). The arrows indicate where the manipulation in the image was made (the arrows were not present in the actual experiment). For every condition we show one example image pair (A–D). The conditions are based on the amount of change in contrast energy and subjective importance. The conditions were balanced for changes in size, distance from the center, contrast, luminance and spatial frequency.

pixels inside the spatial window, and L_i is the luminance of the i th pixel.

Using the local contrast energy values per pixel, we computed for every image pair the change in contrast energy between the original image and the manipulated image. Figure 3 gives an illustration of this procedure. To compute the amount of local contrast energy change, we used the local contrast energy values of the pixels that were changed in the image; the red line represents this region. The difference between the mean local contrast energy values from the original image and the mean local

contrast energy values from the manipulated image defined our measure of contrast energy change.

Calculation of Change in Subjective Importance of the Scene

For every image from the ‘Berkeley Segmentation Dataset and Benchmark’ database (Martin et al., 2001), there are manually labeled images available. In each manually labeled image, a human observer drew lines on the image to highlight the parts of the image they considered to be important for the representation

Measure of change in contrast energy and subjective importance

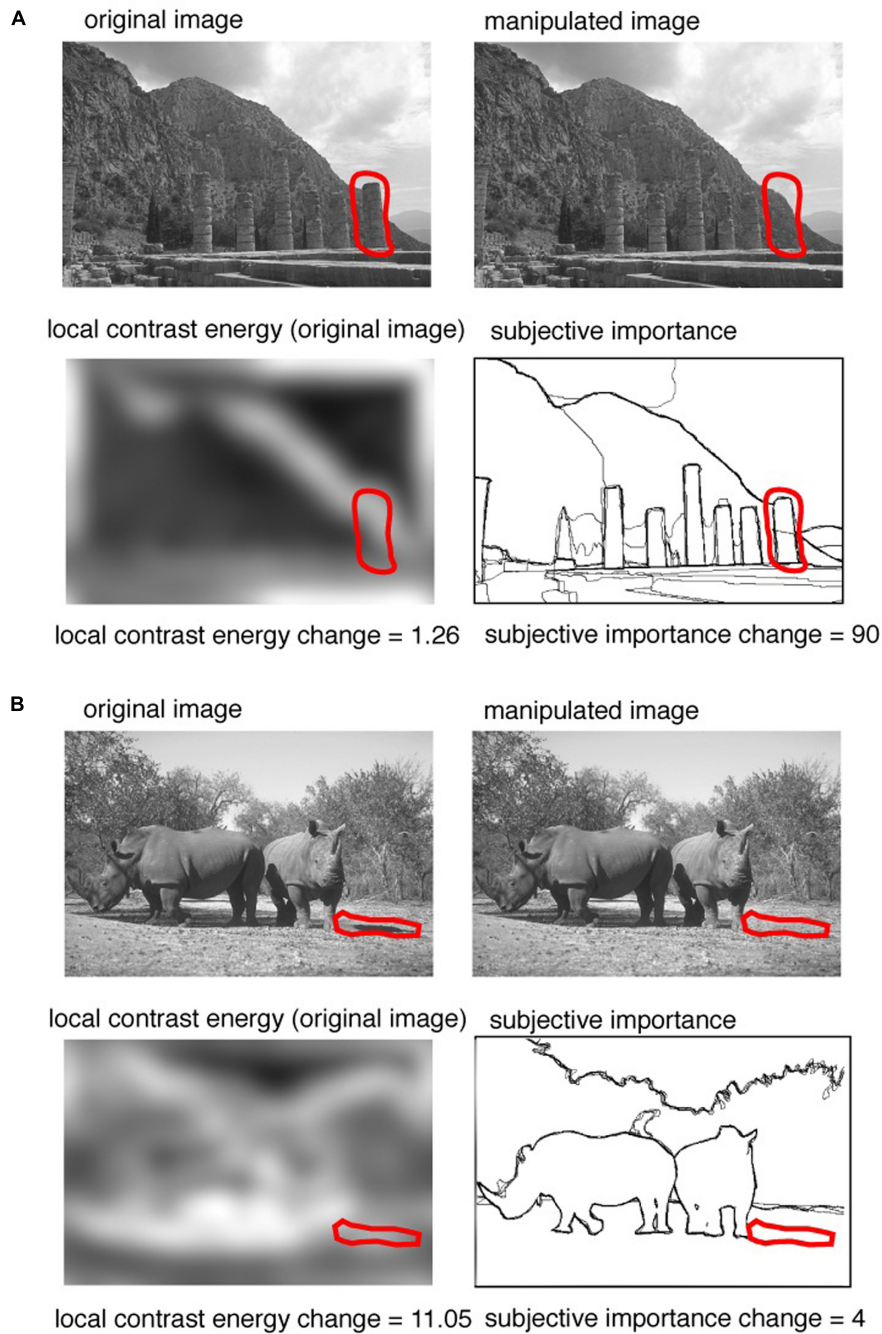


FIGURE 3 | An illustration of the procedure of calculating the change in contrast energy and subjective importance. Two example images [taken from the 'Berkeley Segmentation Dataset and Benchmark' database (Martin et al., 2001)] are shown with changes predominant in subjective importance (**A**) and contrast energy (**B**). The red line represents the region manipulated to alter the image. From this region we calculated the difference in local contrast energy as well as the amount of subjective importance.

of the scene. We used the average manual labeled image of 5 observers as our definition of subjective importance. In **Figure 1** five examples of natural images (**Figure 1A**) are shown together with the averaged manually labeled images (**Figure 1B**).

We used the averaged manually labeled images to calculate the amount of change in subjective importance that a manipulation brings to an image. The pixels of the manually labeled images have values between 0 (not labeled) and 1 (pixel labeled by all

5 observers). The amount of change in subjective importance was calculated by summing the values of the pixels in the averaged manually labeled images that were manipulated. The computation is illustrated in **Figure 3**. The red line in **Figure 3** represents the changed region. From this region we computed how big the change in the manually labeled aspects of the image was and this gives us our measure of change in subjective importance from this region we also computed the amount of change in contrast energy. Two example images are shown: one in which the change is dominated by subjective importance (**Figure 3A**) and one in which the change is dominated by contrast energy (**Figure 3B**). The region was controlled for changes in size, distance from the center, mean luminance and mean contrast of the changed area and spatial frequency (see below).

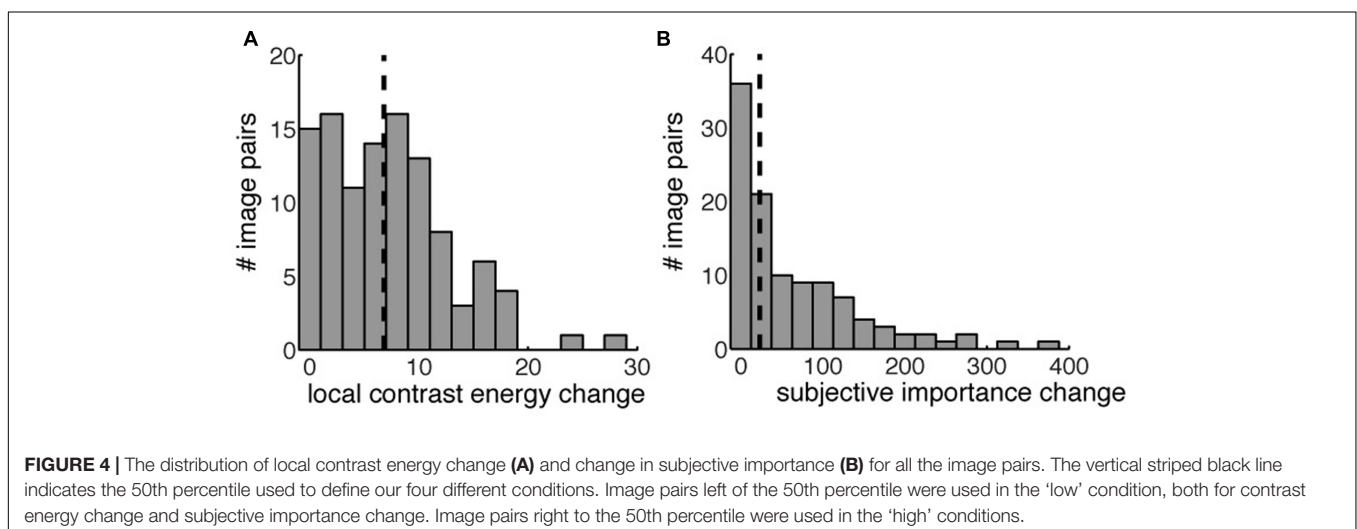
Conditions

Based on the changes in contrast energy and subjective importance we defined four conditions based on the image-pairs that are 'low' and 'high' in their differences for contrast and subjective importance. The 108 image-pairs that we used in the experiment were selected from a larger subset of images to equally balance the four conditions. We balanced the conditions for (i) the number of images per condition and (ii) changes in size, distance from the center, mean contrast, mean luminance change and spatial frequency change of the manipulated area. We made a within subject design that compared high versus low contrast energy and high versus low subjective importance. Furthermore we also modeled the interaction between contrast energy and subjective importance. The definition of 'low' and 'high' was based on the 50th percentile of the values for contrast energy change and subjective importance change (median split). We took the first 50th percentile of the changes for the condition 'low' and the second 50th percentile for the condition 'high,' both for contrast energy as for subjective importance. **Figure 4** shows the histograms of the changes in contrast energy (**Figure 4A**) and subjective importance (**Figure 4B**); the vertical black striped line represents the 50th percentile. All the images below this

boundary-line are placed in the 'low' condition. The images above the 50th percentile are placed in the 'high' condition. **Figure 2** shows for every condition an example of an image pair.

The conditions were balanced for changes in size, distance from the center, mean contrast, mean luminance change and spatial frequency change of the manipulated area. The size of the change was calculated as the number of pixels that were changed in the original image. The distance from the center was calculated as the eccentricity where the center of mass of the change was. The mean contrast was calculated as the average contrast of the manipulated area in the original image and the manipulated image. The luminance change was the difference of the mean luminance of the original image and the manipulated image in the area that was changed. The change in spatial frequency was calculated by the median of the subtracted spatial frequency distributions (calculated by fast Fourier analysis of the local image patch defined by the changed area) in the original and the manipulated image. In **Figure 2** the corresponding values for the calculated changes are reported next to the image pairs. No significant changes were found using an Anova to test for group differences [changes in size: $F(3,104) = 0.66$, $p = 0.58$, distance from the center: $F(3,104) = 1.75$, $p = 0.16$, mean contrast: $F(3,104) = 1.17$, $p = 0.33$, maximum contrast of both images $F(3,104) = 1.07$, $p = 0.37$, minimum contrast of both images $F(3,104) = 0.60$, $p = 0.61$, mean luminance difference: $F(3,104) = 1.03$, $p = 0.38$, spatial frequency difference $F(3,104) = 0.31$, $p = 0.81$], or the Kruskal-Wallis test to test for group differences [changes in size: $H(3,104) = 2.76$, $p = 0.43$, distance from the center: $H(3,104) = 5.01$, $p = 0.17$, mean contrast: $H(3,104) = 4.18$, $p = 0.24$, maximum contrast of both images $H(3,104) = 4.51$, $p = 0.21$, minimum contrast of both images $H(3,104) = 1.84$, $p = 0.61$, mean luminance difference: $H(3,104) = 4.19$, $p = 0.24$, spatial frequency difference $H(3,104) = 1.5$, $p = 0.68$].

We selected 108 different image pairs (i.e., 108 different manipulations) from a larger subset of image pairs so that every condition contained 27 of these image pairs. Some of the images were used more than once, so different manipulations to the



images could be made. To make sure that every subject did not see multiple manipulations to the same image, we made 6 different subsets, in which every original image was only inserted once. The subsets contained 36 different image pairs and the subsets were balanced so that every condition contained 9 individual image pairs. Every subset of images was presented to 10 subjects. Apart from these image pairs that were used for the experiment, we also used 4 trial image pairs. Every trial image pair was an example from one of the 4 conditions. These trial image pairs were used before the actual experiment began.

Design

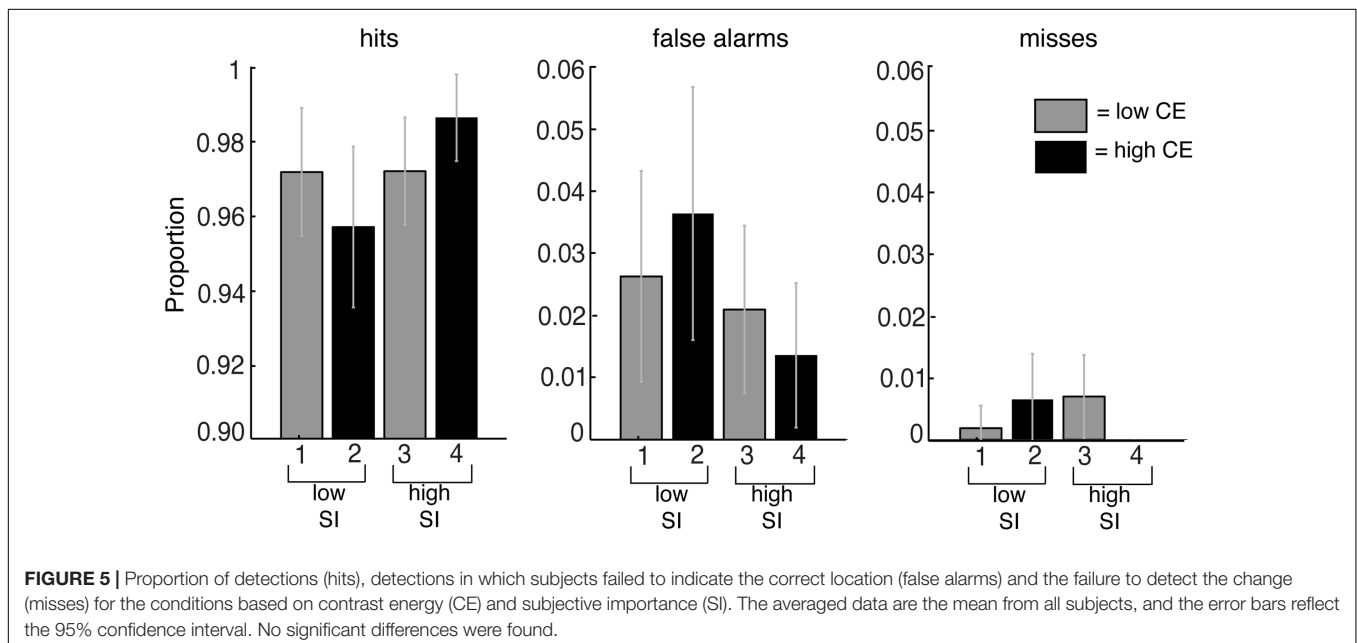
For the change blindness experiment we used a flicker task (Rensink et al., 1997). In the flicker task, an original image was repeatedly alternated with a manipulated image, until the observer noticed the change. The original and the manipulated image were presented for 600 ms and were separated by a gray screen that was presented for 100 ms. The separation with a gray screen of approximately 80 ms prevents the observer from seeing the change due to movement transients caused by the change between the two images (Rensink et al., 1997). The maximum duration of the experiment was 240 s. Subjects were instructed to indicate when they saw the change by pressing the space bar. After this response, the last presented image from the image pair was presented again. At this image, the subjects indicated with the mouse where they had noticed the change in the image. After every trial the subject received feedback whether the change was correctly detected. The answer was classified correctly when the subject pointed within a small region (approximately 0.75° of visual angle) of the change. The 36 different image pairs were presented in a randomized order for every individual experiment. The experiment started with four trial image pairs to make the participants acquainted with the experiment, these trials were excluded from the analysis.

Analysis

We measured the time from the onset of the stimulus to the time at which the subject pressed the spacebar as the RT. Trials in which the subject did not report to see a change (misses) or did not report the change at the accurate location (false alarms) were analyzed using repeated measures ANOVA (Figure 5). There were no significant differences in the proportion of false alarms [$F(3,236) = 1.39, p = 0.25$] and misses [$F(3,236) = 1.62, p = 0.19$] in the different conditions, so these trials were excluded from the analysis (61 out of 2160 trials).

We analyzed the RTs using a general linear model (GLM, Matlab, Mathworks, United States). We used the GLM-approach for the analysis on the RTs, since the distributions of the RTs are skewed (non-normal). The non-normal distribution of the RTs was confirmed using a Kolmogorov–Smirnov test [$D(2098) = 0.87, p << 0.001$]. The GLM-approach allows the use of a non-linear linking function to compensate for the skewed RT-distributions. We used an inverse Gaussian as a linking function in the GLM. The linking function transformed the predicted variables of the GLM-model to the non-normally distribution of the RTs. Using the GLM approach we investigated the relation of the measured RTs with contrast energy change, subjective importance change and the interaction between the two using separate regressors.

We further examined whether the interaction effect can be explained by statistical facilitation alone, i.e., RTs can be faster only because more than one property (contrast energy or subjective importance) was present. For this, we compared the race-model, which is the cumulative distribution functions (CDF) of the high-contrast/low-subjective importance (RT_1) and low-contrast/high-subjective importance (RT_2) conditions with the CDF of the high-contrast/high-subjective importance ($RT_{1\&2}$) condition. The upper bound of statistical facilitation was



calculated using the race model inequality (Raab, 1962; Miller, 1982, 1986; Ulrich et al., 2007):

$$P(RT_{1\&2} < t) \leq P(RT_1 < t) + P(RT_2 < t) \quad (3)$$

The race model inequality reflects the probability (P) that an RT of the two-properties-condition (high-contrast/high-subjective importance) is equal or less to the probability of the cumulative RT distributions of the single-properties-conditions (low-contrast/high-subjective importance and high-contrast/low-subjective importance) at a given time t . Shorter RTs for the condition high-contrast/high-subjective importance compared to the race-model reflect a violation of the race-model and show that the interaction effect cannot be explained by statistical facilitation alone. Note that the race-model only provides an upper bound of statistical facilitation, since the CDF of the race-model sums to 2, whereas the CDFs of the single-properties-conditions only sum to 1. Therefore, only positive violations from the race-model are of interest.

Besides comparing the median RT difference between the high-contrast/high-subjective importance and the race-model, we also showed how the CDFs vary over different quantiles of the RTs (10, 20, and up to 90 percentile). Comparing the RTs for different quantiles showed the absolute difference in RTs between the condition high-contrast/high-subjective importance and the race-model at comparable points of the CDF. We bootstrapped the CDFs over our subjects (using random samples of subjects) to obtain the 95% confidence interval and the p -values to test for a statistical violation of the race model for the different quantiles. We corrected for multiple comparisons (number of tests = 9) using the Bonferroni method (corrected p -value = 0.05/9).

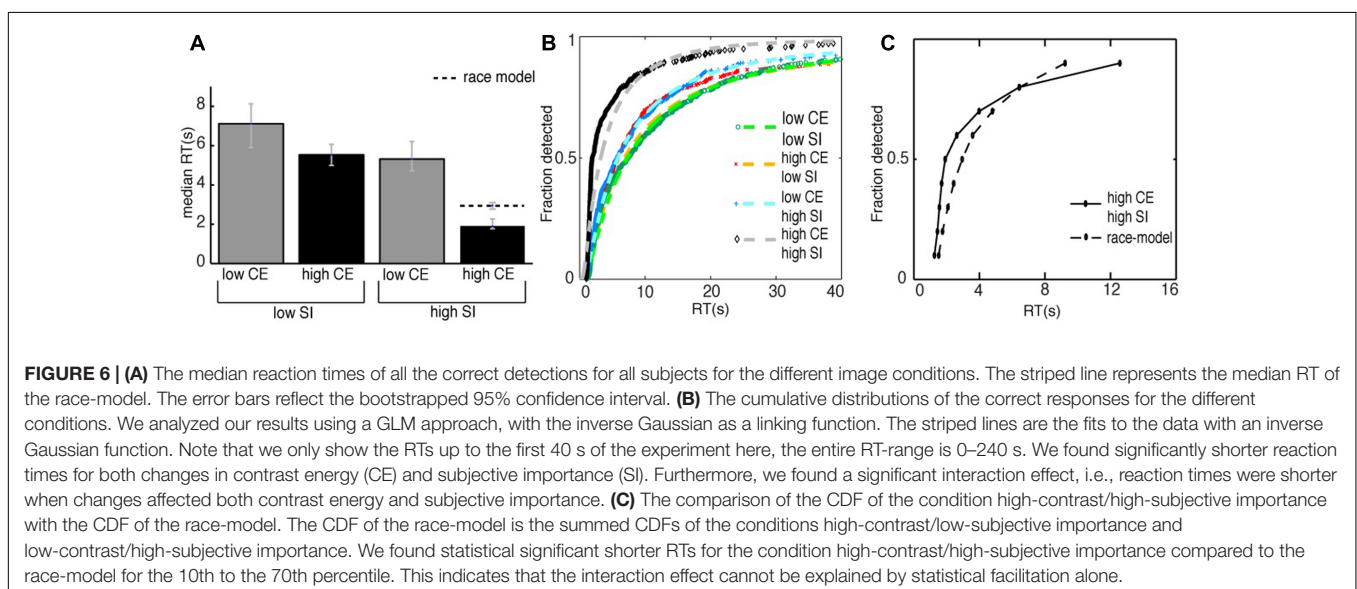
RESULTS

We measured detection rate and RTs for the four different conditions based on changes that are either low or high in

contrast energy (CE) and low or high in subjective importance (SI).

The average detection rate for all changes in all image conditions was 97%, i.e., all the changes to these images were detected within the 240 s that the images were presented. **Figure 5** shows the proportion of detections (hits), detections in which subjects failed to indicate the correct location (false alarms) and failures to detect the change (misses). Since there were no significant differences in the proportion of false alarms and misses in the different conditions we excluded these trials from the analysis. Thus only the trials in which the subjects correctly detected the changes to the images were used for further analysis.

We measured the time from the beginning of the stimulus, until the time the subjects noticed the change and pressed the spacebar as the RT. **Figure 6A** shows the median RTs per condition, the error bars reflect the bootstrapped 95% confidence intervals. The condition in which the manipulations to the image were the most difficult to detect was the low-contrast/low-subjective importance condition, i.e., in this condition the RT was longest (median RT = 7.1 s). Comparing this to the high-contrast/low-subjective importance condition, we see that changes in the latter condition were detected faster (median RT = 5.5 s, median decrease RT = 1.6 s). Comparing the first condition to the low-contrast/high-subjective importance condition, we also found a faster detection of the manipulations (median RT = 5.3 s, median decrease RT = 1.8 s). Last, the manipulations of the high-contrast/high-subjective importance condition were detected fastest (median RT = 1.9 s, median decrease RT = 5.2 s). These results indicate an effect both for the amount of subjective importance change and the amount of contrast energy change on the RT. A manipulation to an image with a larger change in either contrast energy or subjective importance to the image led to a faster detection of the manipulation. The results also suggest an interaction effect for the amount of contrast energy change and the amount of subjective



importance change. A manipulation that was both high in its change for contrast energy and subjective importance resulted in the fastest detection of the manipulation.

To quantify these observations, we performed a GLM-analysis to compare the effects of contrast energy change, subjective importance change and the interaction of the two on the RTs. Since the distribution of RTs was skewed, we did not use the canonical linking function that tests for differences in the conditions based on their mean. Instead of the canonical linking function, we used the inverse Gaussian linking function to take into account the non-normal or non-Gaussian distribution of the data. The differences of the distributions are now based on the fit of the inverse Gaussian distribution. **Figure 6B** shows the fits of the inverse Gaussian function to the distribution of the data for the different conditions. All fits had a variance explained $R^2 > 0.97$. Using the GLM analysis, we found a significant effect for the level of contrast energy change: $t(2095) = 9.92, p < 0.001$. Where the number 2095 reflects the total amount of (correct) RTs from all the subjects. Also, we found a significant effect for the level of subjective importance change: $t(2095) = 6.54, p < 0.001$. Furthermore, we found a significant effect for the interaction of the amount of contrast energy change and the amount of subjective importance change: $t(2095) = 7.90, p < 0.001$. Thus, our ability to detect changes in a visual scene is influenced by both the subjective image interpretation and the stimulus-driven image properties, and they interact.

Besides our analysis based on the four different conditions, we repeated the analysis using a GLM with ranked values for contrast energy and subjective importance. Using this analysis we found a significant effect for both contrast energy ($p < 0.001$) and the interaction of contrast energy and subjective importance ($p = 0.013$). However, the effect for subjective importance was not significant ($p = 0.0578$). This analysis assumes a linear relationship between the values for contrast energy and subjective importance, which is not likely the case and we assume this is the reason why we don't find a significant effect for subjective importance using this analysis method. Therefore, we believe that the split analysis is more correct.

To further examine this interaction effect, we investigated whether it can be explained by statistical facilitation alone. We compared the CDF of the high-contrast/high-subjective importance condition to the CDF of the race model (**Figure 6C**). We compared the bootstrapped CDFs at different quantiles, from the 10th to the 90th percentile. For the 10th to the 70th percentile we found a statistically significant violation of the race model ($p < 0.001$ for all significant percentiles, Bonferroni corrected for multiple comparisons), with a maximum of 1 s difference in RT that was unexplained by statistical facilitation.

DISCUSSION

In our study we used manually labeled images to investigate the simultaneous effect of stimulus-driven and subjectively

important image properties on our ability to detect changes in a visual scene. We measured RTs within a change blindness task and compared RTs over four different conditions. The conditions were based on the amount of change in the stimulus-driven image property contrast energy and the amount of change in the subjective image interpretation defined by the manual labels.

Shorter RTs for High Change in Subjective Importance

Analyzing the RTs for the different conditions, first, we found a significant effect for subjective importance. RTs were shorter for manipulations that have a larger change in their subjective image interpretation. This measure of change was based on the manually labeled aspects of the images, where people indicated which parts represented the image most (Martin et al., 2001). Our results are in line with previous change blindness studies, which showed that changes made in parts of a scene that were considered to be important for their high-level image representation were detected earlier (Rensink et al., 1997; O'Regan et al., 1999; Shore and Klein, 2000; Stirk and Underwood, 2007; Sampanes et al., 2008; Verma and McOwan, 2010). However, different notions of the high-level image representation have been used in these studies. Some refer to it as the gist; the general interpretation of a scene (Stirk and Underwood, 2007; Sampanes et al., 2008), where others refer to it as regions of high interest (Rensink et al., 1997; O'Regan et al., 1999; Shore and Klein, 2000; Verma and McOwan, 2010). Our study shows most resemblance to the studies that use the notion of the high-level image interpretation as regions of high interest.

Shorter RTs for High Contrast Energy Change

We also found a significant effect for the amount of contrast energy change on the RTs. Manipulations with a larger change in contrast energy were detected earlier. Neurophysiological data shows that the early visual cortex responds strongly to changes in contrast energy (De Valois and De Valois, 1988; Boynton et al., 1999; Olman et al., 2004; Mante and Carandini, 2005; Dumoulin et al., 2008). We speculate that this preference of early visual cortex underlies the ability to detect these changes. Alternatively, changes in contrast energy might attract more (exogenous) attention. These two explanations are not mutually exclusive and both or either one may explain why changes in contrast might be detected earlier.

Most change blindness studies do not account for effects of stimulus-driven image statistics. Some studies did, but show deviating results. Verma and McOwan (2010) and Spotorno et al. (2013) found faster detection for changes in a region with high salience whereas, Stirk and Underwood (2007) did not. Contrast energy is not the same as saliency. Depending on the definition, saliency is a complex combination of image features such as variations in color, luminance and orientation (Itti and Koch, 2000; Borji et al., 2013). Thus, depending on the

definition, saliency will involve contrast-energy but also many other stimulus-driven features. We focused on contrast energy as it is well defined and a well-known computation of early visual cortex. We show that contrast energy can lead to differences in the detection of changes. This opens the door to investigate other image properties and computations of the early visual system that may potentially also contribute to change blindness, such as orientation and spatial frequency (Hubel, 1982), spatial coherence (Groen et al., 2012, 2013), and contour integration (Field et al., 1993).

Interaction of Change in Contrast Energy and Subjective Importance

Last, we found a significant interaction effect for the amount of change in subjective importance and contrast energy change. By using the combination of the manually labeled images simultaneously with the inherent contrast energy of the images, we were able to investigate not only the effect of stimulus-driven and subjectively important image properties on the change blindness paradigm. Most importantly, we were able to simultaneously measure the effect of both because the levels of contrast change and change in subjective importance were independent in the images.

The interaction we find cannot be explained by the added benefit of two independent observations (race-model). This suggests that the stimulus-driven and subjectively important image properties are not independently processed in the visual system, but instead they facilitate each other. Perception does not only rely on visual input but also on our knowledge of the world (von Helmholtz, 1867). Contrast-energy is solely determined by the visual input, whereas both visual input and our knowledge of the world drive the subjective importance judgments. Therefore, we speculate that this interaction suggests that, like perception, change blindness also relies on visual input and our knowledge of the world.

This interaction between contrast energy and subjective importance may be implemented in different ways in the visual system. For example, different theories propose that higher-order areas feedback to lower-order areas and modulate their responses according to the prior expectations about the visual world (Mumford, 1992; Koch and Poggio, 1999; Rao and Ballard, 1999; Friston, 2005). We speculate that this neural interaction is the origin of the behavioral interaction effect of the subjectively important image representation and the stimulus-driven image statistics.

REFERENCES

- Bex, P. J., and Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 19, 1096–1106. doi: 10.1364/JOSAA.19.001096
- Borji, A., Sihite, D. N., and Itti, L. (2013). Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. *IEEE Trans. Image Process.* 22, 55–69. doi: 10.1109/TIP.2012.2210727
- Boynton, G. M., Demb, J. B., Glover, G. H., and Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Res.* 39, 257–269. doi: 10.1016/S0042-6989(98)00113-8

CONCLUSION

While earlier studies on change blindness conclude that our ability to detect changes in a visual scene is influenced by the subjective image interpretation of the scene, we find that this ability is also influenced by the stimulus-driven image property contrast energy. Specifically, large changes in subjective importance and contrast energy are both detected faster than small changes in these properties. Moreover, these two properties interact; image manipulations that strongly affect both subjective importance and contrast energy are detected fastest, and importantly, faster than expected based on the amount of change of either property alone. We also found an interaction effect; manipulations that were both high in their change of subjective importance and contrast energy were detected earliest. Presumably, these parts of an image are processed in greater detail. The change blindness paradigm can be used to give more insight in how stimulus-driven image statistics and subjective image representations are processed in the visual system. Our results suggest that they are not processed independently, but instead interact in how they are represented in the visual system.

AUTHOR CONTRIBUTIONS

WZ and SD designed the study. All authors contributed to the study design. Stimulus creation and data collection was performed by WZ and JvL. WZ performed the data analysis and interpretation under the supervision of SD. WZ drafted the manuscript, and JvL and SD provided critical revisions. All authors approved the final version of the manuscript for submission.

FUNDING

This work was supported by a Netherlands Organization for Scientific Research (NWO) Vidi grant 452-08-008 to SD.

ACKNOWLEDGMENTS

The Spinoza Centre is a joint initiative of the University of Amsterdam, Academic Medical Center, VU University, VU medical center, Netherlands Institute for Neuroscience and Royal Academy of Arts and Sciences.

- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- De Valois, R. L., and De Valois, K. K. (1988). *Spatial Vision*. Oxford: Oxford University Press.
- Dumoulin, S. O., Dakin, S. C., and Hess, R. F. (2008). Sparsely distributed contours dominate extra-striate responses to complex scenes. *Neuroimage* 42, 890–901. doi: 10.1016/j.neuroimage.2008.04.266
- Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local “association field”. *Vision Res.* 33, 173–193. doi: 10.1016/0042-6989(93)90156-Q

- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Groen, I. I., Ghebreab, S., Lamme, V. A., and Scholte, H. S. (2012). Low-level contrast statistics are diagnostic of invariance of natural textures. *Front. Comput. Neurosci.* 6:34. doi: 10.3389/fncom.2012.00034
- Groen, I. I., Ghebreab, S., Prins, H., Lamme, V. A., and Scholte, H. S. (2013). From image statistics to scene gist: evoked neural activity reveals transition from low-level natural image structure to scene category. *J. Neurosci.* 33, 18814–18824. doi: 10.1523/JNEUROSCI.3128-13.2013
- Hubel, D. H. (1982). Exploration of the primary visual cortex, 1955–78. *Nature* 299, 515–524. doi: 10.1038/299515a0
- Itti, L., and Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.* 40, 1489–1506. doi: 10.1016/S0042-6989(99)00163-7
- Kelley, T. A., Chun, M. M., and Chua, K. P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *J. Vis.* 3, 1–5. doi: 10.1167/3.1.1
- Koch, C., and Poggio, T. (1999). Predicting the visual world: silence is golden. *Nat. Neurosci.* 2, 9–10. doi: 10.1038/4511
- Landman, R., Spekreijse, H., and Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Res.* 43, 149–164. doi: 10.1016/S0042-6989(02)00402-9
- Mante, V., and Carandini, M. (2005). Mapping of stimulus energy in primary visual cortex. *J. Neurophysiol.* 94, 788–798. doi: 10.1152/jn.01094.2004
- Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proceedings of the Eighth International Conference on Computer Vision*, Vol. 2, Vancouver, BC, 416–423. doi: 10.1109/ICCV.2001.937655
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)90010-X
- Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Percept. Psychophys.* 40, 331–343. doi: 10.3758/BF03203025
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241–251. doi: 10.1007/BF00198477
- Olman, C. A., Ugurbil, K., Schrater, P., and Kersten, D. (2004). BOLD fMRI and psychophysical measurements of contrast response to broadband images. *Vision Res.* 44, 669–683. doi: 10.1016/j.visres.2003.10.022
- O’Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of ‘mudsplashes’. *Nature* 398, 34. doi: 10.1038/17953
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Trans. N. Y. Acad. Sci.* 24, 574–590. doi: 10.1111/j.2164-0947.1962.tb01433.x
- Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Res.* 40, 1469–1487. doi: 10.1016/S0042-6989(00)00003-1
- Rensink, R. A. (2002). Change detection. *Annu. Rev. Psychol.* 53, 245–277. doi: 10.1146/annurev.psych.53.100901.135125
- Rensink, R. A., O’Regan, J. K., and Clark, J. J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychol. Sci.* 8, 368–373. doi: 10.1111/J.1467-9280.1997.Tb00427.X
- Sampanes, A. C., Tseng, P., and Bridgeman, B. (2008). The role of gist in scene recognition. *Vis. Res.* 48, 2275–2283. doi: 10.1016/j.visres.2008.07.01
- Shore, D. I., and Klein, R. M. (2000). The effects of scene inversion on change blindness. *J. Gen. Psychol.* 127, 27–43. doi: 10.1080/00221300009598569
- Simons, D. J., and Ambinder, M. S. (2005). Change blindness - theory and consequences. *Curr. Dir. Psychol. Sci.* 14, 44–48. doi: 10.1111/J.0963-7214.2005.00332.X
- Simons, D. J., and Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends Cogn. Sci.* 9, 16–20. doi: 10.1016/j.tics.2004.11.006
- Spotorno, S., Tatler, B. W., and Faure, S. (2013). Semantic consistency versus perceptual salience in visual scenes: findings from change detection. *Acta Psychol.* 142, 168–176. doi: 10.1016/j.actpsy.2012.12.009
- Stirk, J. A., and Underwood, G. (2007). Low-level visual saliency does not predict change detection in natural scenes. *J. Vis.* 7, 3.1–3.10. doi: 10.1167/7.10.3
- Ulrich, R., Miller, J., and Schroter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39, 291–302. doi: 10.3758/BF03193160
- Verma, M., and McOwan, P. W. (2010). A semi-automated approach to balancing of bottom-up salience for predicting change detection performance. *J. Vis.* 10:3. doi: 10.1167/10.6.3
- von Helmholtz, H. (1867). “Treatise on Physiological Optics,” in *Handbuch der Physiologischen Optik*, ed. J. P. C. Southall (New York, NY: Dover).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Zuiderbaan, van Leeuwen and Dumoulin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.