It is still an unresolved question how the visual system perceives surface lightness given the ambiguity of the sensory input signal. We studied lightness perception using two-dimensional images of variegated checkerboards shown as perspective projections of three-dimensional objects. We manipulated the contrast of a target check relative to its surround either by rendering the image under different viewing conditions or by introducing noncoincidental changes of the reflectance of the surfaces adjacent to the target. We examined the predictive power of the normalized contrast model (Zeiner & Maertens, 2014) for the different viewing conditions (plain view vs. dark and light transparency) as well as for the noncoincidental surround changes (only high or only low reflectances in the surround). The model accounted for lightness matches across different viewing conditions but not for the surround changes. The observed simultaneous contrast effects were smaller than what would be predicted by the model. We evaluated two model extensions that—both relying on contrast—predicted the observed data well. Both model extensions point to the importance of contrast statistics across space and/or time for the computation of lightness, but it awaits future testing to evaluate whether and how the visual system could represent such statistics.

Introduction

The present article addresses how the visual system extracts information about stable properties of objects despite fluctuations in the signal that serves as the sensory input. In particular, we want to understand how, in the achromatic domain, the apparent lightness of a surface is determined from the luminance signal in the retina. A complete model of lightness perception would allow the prediction of surface lightness from the pattern of retinal stimulation. At the moment there is no such model of lightness perception (e.g., Brainard & Radonjić, 2014; Kingdom, 2011, for review).

One can broadly distinguish between two approaches that have been proposed to account for the perception of surface lightness from image luminance. In the inverse optics approach (e.g., Barrow & Tenenbaum, 1978; D’Zmura & Iverson, 1993) it is assumed that the visual system undoes, or inverts, the physical generative process of image formation. In the domain of lightness perception, the generative process is captured by the following equation: \[ L = I \times R + T, \] whereby \( L \) is retinal luminance, \( I \) is illumination, \( R \) is surface reflectance, and \( T \) is a potentially intervening transparent medium such as fog (Adelson, 2000). The task of the visual system is to (explicitly) estimate the external sources (i.e., a light source, the surface reflectances, and a transparent medium) from the retinal luminance signal. This idea is expressed in concepts such as the atmospheric transfer function, which captures the process of image formation, and its inverse, the lightness transfer function, which captures the processes involved in perceiving lightness from luminance (Adelson, 2000). The idea is also evident in equivalent lighting models (Allred & Brainard, 2013; Brainard & Radonjić, 2014; Kingdom, 2011, for review).
Boyaci, Maloney, & Hersh, 2003; Doerschner, Boyaci, & Maloney, 2004, 2007; Murray, 2013).

In the other approach it is assumed that surface lightness is computed from image luminance without an explicit estimation of the illuminant or of transparent layers. Instead, the computations that determine surface lightness from luminance simply get rid of unwanted variability that is introduced by illumination changes and the like. We call this the heuristic or cue-based approach of lightness perception (e.g., Thompson, Fleming, Creem-Regehr, & Stefanucci, 2011). The simplest of such models is the luminance ratio model (Wallach, 1948), according to which lightness matches are accomplished by matching luminance ratios, or contrast. Other representatives of that account are anchoring theory (Gilchrist et al., 1999; Radonjic & Gilchrist, 2010), spatial filtering models (e.g., Blakeslee & McCourt, 2004), edge-integration theory (Rudd, 2013), and other models that are based on contrast normalization (Singh, 2004; Singh & Anderson, 2002; Zeiner & Maertens, 2014).

**Contrast-based models**

The computation of contrast from the retinal luminance signal was suggested as an easy and physiological plausible way to accomplish lightness constancy (Shapley & Enroth-Cugell, 1984), because the contrast signal was said to be invariant with changes in the level of illumination. For example, the Michelson contrast in general is defined as

\[
\frac{(L_{\text{figure}} - L_{\text{background}})}{(L_{\text{figure}} + L_{\text{background}})}
\]

whereby the respective luminances, L, are proportional to \( I \times R \) where I is the (common) illumination and R is the reflectance of either figure or background. Dividing numerator and denominator by I yields

\[
\frac{(R_{\text{figure}} - R_{\text{background}})}{(R_{\text{figure}} + R_{\text{background}})}
\]

the contrast of object reflectances, which is independent of the level of illumination (see Shapley & Enroth-Cugell, 1984, p. 268). When a transparent medium is introduced into a scene instead of a pure illumination change, the contrast signal of a surface is no longer invariant. However, there seems to be still a systematic relationship between contrast and surface reflectance.

Zeiner and Maertens (2014) measured the perceived lightness of target checks of different reflectances that were embedded in variegated checkerboard patterns. The checkerboard was presented in plain view or with some region of the checkerboard being obscured by a shadow or a transparent medium (Figure 1). To account for the observed lightness matches, they proposed the normalized contrast model that comprises three steps: First, the Michelson contrast is computed between a target and its eight surround checks (\( x \), see Figure 1A). That is, the target intensity \( X \) is normalized relative to its average local surround intensities \( S_1, \ldots, S_8 \):

\[
x = \frac{X - \text{mean}(S_1, \ldots, S_8)}{X + \text{mean}(S_1, \ldots, S_8)}
\]

Such contrasts are computed for all checks in the checkerboard. Second, it is assumed that the visual system is sensitive to differences in contrast range and can use them to detect regions seen in plain view, because they have the highest contrast range (Anderson, 1999; Singh, 2004; Singh & Anderson, 2002). Third, if the target is located in a region of reduced contrast, the target contrast is normalized relative to the contrast range in the region of transparency (\( t_{\text{max}} - t_{\text{min}} \)), and this range is subsequently mapped to the contrast range in plain view (\( p_{\text{max}} - p_{\text{min}} \)). The so normalized target contrast depicts the expected lightness match in contrast units. Here, we call it the normalized Michelson contrast (\( NMC \)). It relates to the target contrast \( x \) according to the following equation:

\[
NMC = \frac{x - t_{\text{min}}}{t_{\text{max}} - t_{\text{min}}} \times (p_{\text{max}} - p_{\text{min}}) + p_{\text{min}}.
\]
2002) and uses them to distinguish image regions in plain view from regions that are seen through a transparency. For the range normalization it is additionally assumed that the visual system has some representation of the respective ranges and can map them onto one another. The normalized contrast model successfully explained the data and outperformed a number of a standard lightness models.

It is important to note that in both cases an illumination change as well as a change of viewing context, contrast is only a valid cue when (a) all luminances that enter the contrast equation come from surfaces that are seen in the same illumination (viewing context), and (b) when the normalizing surround term includes luminance values from a representative sample of surfaces seen within one illumination (viewing context). A violation of (a) would occur when the luminance of the figure comes from a surface seen in Illumination A, but the luminances of the background come from surfaces that are partially exposed to Illumination A and partially exposed to Illumination B. A violation of (b) would occur, when those luminances that enter the surround term are from surfaces that do not represent the illumination well, because the surfaces might only represent a subset of the representative sample of surfaces (i.e., only the low end; the equivalent is true for a change of viewing context). Traditional stimuli often consist of only two different luminance values (i.e., one for the figure and one for the background). Hence the potential problems have not been relevant. In the case of more natural and articulated stimuli, however, these assumptions and their potential violations should be considered.

Simulations of contrast distributions

To illustrate the systematic relationship between contrast and surface reflectance, we simulated contrast distributions for different arrangements of target reflectances and target surrounds in our checkerboard stimuli. To that end, we randomly assigned reflectance values from our pool of reflectances in the surround of the target (under consideration of the sampling constraints for our stimuli) and computed the resulting target contrasts. This led to a characteristic distribution of contrast values for each target reflectance (see black histograms in Figure 2), and a characteristic shift of the center of mass of each histogram with increasing target reflectance.

This shows that for the above-mentioned criteria (a) an equal illumination of target and surround, and (b) a representative sampling of the surround surfaces, there is a high correlation between contrast and surface reflectance with only minor fluctuations. This is to say that, in a specific context some contrast values are more likely to occur for a particular target reflectance than others, and hence contrast could be an appropriate candidate variable for the computation of a stable percept of surface lightness. On the downside this also means that lightness judgments that are based on contrast alone can go wrong for stimulus configurations that are extremely rare. The dark and light gray arrows in Figure 2 demarcate contrast values that occur less than 5% of the time under random sampling conditions. Such contrast values result from an arrangement of mainly dark or light reflectances in the surround of the target check. For those extreme contrast values, one would expect failures of lightness constancy, because they are more likely to occur for a different target reflectance. For example, the distribution of contrasts for Reflectance 3 is centered at a contrast of $-0.63$, which is indicated by the black arrow below the $x$-axis of the corresponding panel. A contrast value of $-0.07$, indicated by the dark gray arrow, is unlikely to occur for target Reflectance 3 and random sampling of surround reflectances. A contrast value of $-0.07$ is more likely to occur for target Reflectance 7, where the black arrow indicates the center of mass of the corresponding contrast distribution. We would thus expect systematic shifts toward higher or lower perceived target reflectances for...
extremely dark or low surrounds, respectively. This is consistent with a simultaneous contrast effect.

Study outline

Here, we tested the effect of variations in target contrast on perceived lightness. Contrast variations were introduced by the selective sampling of surface reflectances in the surround of a target region. We systematically varied surround reflectances so that they would either all be relatively darker or lighter than the average of the surfaces presented within the rest of the checkerboard. The resulting contrast values for the high and low surround condition are indicated in Figure 2 by the light and dark gray arrows, respectively. We tested the effect of such extreme surround conditions in checkerboards that were seen in plain view and for checkerboards that were partially obscured by a dark or a light transparent medium (see Figure 3).

Methods

Observers

Five naive observers participated in the study; one of them was male. All observers had normal or corrected-to-normal visual ability. Observers’ ages ranged from 24 to 32 years. Observers were reimbursed for their participation. Informed consent was given by all participants prior to the experiment.

Stimuli and apparatus

Stimuli were presented on a linearized 21-in. Siemens SMM2106LS monitor (400 × 300mm, 1024 × 768px, 130 Hz). Presentation was controlled by a DataPixx toolbox (Vpixx Technologies, Inc., Saint-Bruno, QC,
Stimuli were images of customized checkerboards composed of 6 × 6 checks. The images were rendered using Povray (Persistence of Vision Raytracer Pty. Ltd., Williamstown, Victoria, Australia, 2004). The position of the checkerboard, the light source, and the camera were kept constant for each image. Checks were randomly assigned one out of 12 surface reflectance values, except for the target check and those checks directly adjacent to the target, for which the reflectance was specified in the experimental design. The rendered images were converted to a grayscale matrix. The 12 reflectance values were chosen so as to produce perceptually roughly equidistant luminance values on the experimental monitor. This choice resulted in an asymmetric distribution of luminance values around their mean luminance (cd/m²). The side length of each check was 1.1° of visual angle. The background luminance was 141 cd/m².

In the transparency conditions, a transparent layer was placed between the checkerboard and the camera. The luminance values of the checks that were covered by the transparent medium can be derived according to Metelli's formula: 

\[ LT_t = a \times LP_t + (1 - a) \times LT \]

whereby \( LT_t \) is the luminance of the check seen through the transparent layer, \( LP_t \) is the luminance of the check seen in plain view, \( a \) is the open sector of the rotating episcoverter, and \( LT \) is the luminance of the transparent layer rendered as an opaque surface. The transmittance of the transparent layer was identical in both conditions (\( a = 0.4 \)). In the dark transparency condition the reflectance of the transparent layer was 0.35 povray reflectance (the normalized reflectance value with respect to the highest reflectance value we used in our scenes would be 0.16), which would result in a luminance of 19 cd/m² if the layer would be rendered as an opaque surface. In the light transparency condition the reflectance was 2.0 povray reflectance (normalized reflectance value: 0.9) corresponding to a luminance of 110 cd/m² for an opaque surface.

For each viewing condition we measured the luminance values emitted by the experimental monitor that corresponded to each of the 12 target reflectances (see Table 1). Luminance values were measured using a Konica Minolta LS-100 spot luminance photometer (Osaka, Japan) with a 122 close-up lens on the experimental monitor. Measurements were taken semi-automatically at a distance of approximately 20 cm from the screen. Before starting the measurement, the photometer was manually focused on the target check (E2). Measurements were repeated 10 times for each reflectance and for the three viewing conditions (plain, dark transparency, and light transparency).

To assess observers' lightness matches, an adjustable matching field was presented above the checkerboard. The matching region was embedded in a small frontoparallel checkerboard that was composed of 5 × 5 checks. The luminances of the checks in the matching checkerboard were randomly assigned in each trial and were selected from the same 12 luminances used in the plain condition of the test stimuli. The matching checkerboard subtended 3.0° × 3.0° of visual angle while the comparison field was 1.2° of visual angle wide.

### Design

Perceived lightness as a function of target reflectance was measured in three viewing conditions and for three local surround manipulations (Figure 3). The viewing condition was either plain view or behind a dark or a light transparent medium. The local surround of the target check was defined as the eight checks that were directly adjacent to the target (Figure 1, Panel A). In the full surround condition, the eight surround reflectances were chosen randomly from a subsample of the 12 possible reflectance values (see Table 1) such as that on average the mean surround luminance was close to the mean of all 12 reflectance values used in the scene. Two additional constraints were realized for the sampling of the eight surround luminance values. We made sure that no reflectance value was represented more than twice in the surround and that no two identical reflectances were placed next to each other. In the high surround condition, the reflectances of the surround checks were sampled only from the five highest reflectance values, resulting in a higher than average mean surround luminance and a lower than average contrast. In the low surround condition, the reflectance values of the surround checks

<table>
<thead>
<tr>
<th>Reflectance Value</th>
<th>Plain</th>
<th>Dark</th>
<th>Light</th>
<th>Low</th>
<th>High</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>15</td>
<td>18</td>
<td>69</td>
<td>x</td>
<td>—</td>
<td>x</td>
</tr>
<tr>
<td>0.05</td>
<td>25</td>
<td>22</td>
<td>73</td>
<td>x</td>
<td>—</td>
<td>x</td>
</tr>
<tr>
<td>0.09</td>
<td>40</td>
<td>28</td>
<td>79</td>
<td>x</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>0.14</td>
<td>60</td>
<td>36</td>
<td>87</td>
<td>x</td>
<td>—</td>
<td>x</td>
</tr>
<tr>
<td>0.21</td>
<td>89</td>
<td>48</td>
<td>99</td>
<td>x</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>0.28</td>
<td>120</td>
<td>60</td>
<td>111</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>0.37</td>
<td>155</td>
<td>74</td>
<td>125</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>0.58</td>
<td>242</td>
<td>108</td>
<td>159</td>
<td>—</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>0.68</td>
<td>281</td>
<td>125</td>
<td>176</td>
<td>—</td>
<td>x</td>
<td>—</td>
</tr>
<tr>
<td>0.75</td>
<td>312</td>
<td>137</td>
<td>188</td>
<td>—</td>
<td>x</td>
<td>—</td>
</tr>
<tr>
<td>0.88</td>
<td>365</td>
<td>157</td>
<td>209</td>
<td>—</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>1.00</td>
<td>415</td>
<td>177</td>
<td>229</td>
<td>—</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 1. Target and surround luminances.
were sampled only from the five lowest reflectance values resulting in higher than average contrast values. Figure 2 shows the contrast values in the high, low, and average surround condition for each target reflectance relative to a characteristic distribution of contrasts in the plain view condition.

**Experimental procedure**

In each trial observers matched the perceived lightness of a target check (E2) by adjusting the luminance of an external matching field. To instruct observers, we used a wording common in the lightness literature, where observers are asked to adjust the matching field until it matches the perceived lightness of the target patch, such as if both target and match were cut out from the same piece of paper or made from the same material. The luminance of the matching field was adjusted by pressing one of four buttons on the response box. The upper and the lower buttons allowed coarse adjustments of the luminance by adding or subtracting 10 cd/m² to the current luminance; the right and left buttons were used for fine adjustments in steps of 1 cd/m². The maximum luminance of the monitor was 550 cd/m². Following each button press, the luminance of the matching field was updated to the new value. When the match was satisfactory observers used the middle button of the response box to continue to the next trial. No time limit was imposed on the adjustment procedure.

Each observer completed 540 trials. Each target reflectance was repeated five times in each local surround condition and in each viewing context, respectively. Trials were randomized across surround conditions and across viewing contexts.

**Results**

The goal of this experiment was to test the effect of noncoincidental changes in the distribution of surface reflectances surrounding a target region on the perceived lightness of the target. In particular, we wanted to compare the prediction of the normalized contrast model to observers’ lightness judgments under these extreme surround conditions. Before we address this question we will report that the present results are consistent with those from an earlier study.

**Replication of Zeiner and Maertens (2014)**

Figure 4 shows the matches in the full surround condition as a function of target luminance for each observer. Matches in the plain view condition fell roughly on the identity line, which means that they can serve as a reference against which the other conditions can be compared. The luminance range of targets seen through the transparent media was reduced as indicated by a smaller spread on the x-axis, but the matches of Observers 1–4 spanned almost the full range of match luminances, which means that the matches were lightness constant. The luminance matches of Observer 5 were more scattered around the identity line. It seems as if this observer was matching the proximal luminance values of the target, or, in other words, was performing brightness instead of lightness matching (Arend & Goldstein, 1987). For this reason, the observer was excluded from all further analyses.

To better be able to judge the degree of lightness constancy, Figure 5A shows the average luminance matches of the remaining four observers as a function of target reflectance. Equal surface reflectances were matched with similar luminances for most of the tested target reflectances. Deviations from lightness constancy were evident only for the two highest reflectances, where the degree of overlap was lower and the variability between observers was higher. Figure 5B shows the lightness matches as a function of the normalized contrast prediction. Here the matches are also expressed in units of Michelson contrast (matches were not normalized in range, because target and match luminance followed a one-to-one mapping in the plain view condition Figure 4). The normalized contrast measure is a good predictor of perceived
surface lightness across viewing conditions because the data points from all three conditions overlap. Zeiner and Maertens (2014) introduced a global $R^2$ measure to quantify the extent to which the data from different viewing conditions are accounted for by one common linear function. The average global $R^2$ value for the present data is 0.89, range = 0.85–0.95. This is very close to the average global $R^2$ parameter of 0.88 reported by (Zeiner & Maertens, 2014).

**Extreme surround conditions**

To examine the effect of the surround manipulation, matches were depicted in units of luminance in Figure 6A. A low mean surround luminance, and hence a higher target contrast, should cause the target to appear lighter and match luminances should also be higher. A high mean surround luminance, and hence a lower target contrast, should cause the target to appear darker and match luminances should be lower. In other words, the matches should be subject to a simultaneous contrast effect. Figure 6A shows the data for one typical observer. Data for all observers can be found in Figure 13 in the Appendix.

Match luminances were indeed higher in the low than in the high surround condition. The difference increased with increasing reflectance, but the effect was of moderate magnitude. This is in line with results reported by Allred, Radonjic, Gilchrist, and Brainard (2012). Our stimuli—like theirs—are articulated in the immediate surround of the target but also in the more remote regions of the checkerboard (comparable to the lowSt and highSt condition in Allred et al., 2012). The authors observed a reduced simultaneous contrast effect in these conditions compared to the low–low or high–high conditions, which were more similar to the traditional simultaneous contrast displays.

In Figure 6B the same data are plotted in units of contrast at both axes. According to the normalized contrast model, all matches should lie on the identity line, regardless of surround type. This is because, in plain view, no additional normalization step is required and thus match and target contrasts should be identical. This was not observed in the data. Instead, the match contrasts deviated substantially from the target contrasts. The nonlinearity for matches in the low surround condition is a consequence of our stimuli.

Given the moderate simultaneous contrast effects in plain view, one might suspect that observers were
doing luminance matching in the extreme surround conditions. However, looking at the data in the transparency conditions that alternative seems unlikely. Figure 7 shows the matches in the extreme surround conditions as a function of target luminance in the transparency conditions. Luminance matches were higher for the low than for the high surround condition in both contexts. So again, there is a moderate simultaneous contrast effect. But critically the matches do not indicate a luminance matching. The gray line indicates the pattern of results if observers were doing luminance matching, whereas the black line indicates reflectance-based matching. The observed data are more consistent with the reflectance-based matching and thus indicative of lightness constancy. The extreme surround conditions and the viewing conditions were factorially combined and presented in a random sequence across trials. It seems unlikely that the visual system did respond to context changes that were induced by the viewing conditions but not to changes induced by the local surround. We conclude that the moderate simultaneous contrast effects were not due to a luminance-based matching strategy.

To account for the observed data we consider two types of modifications of the normalized contrast models. In the following we explain the reasoning for each of them.

Recursive contrast normalization

Figure 7. Luminance matches as a function of target luminance for one observer. Panel A: Dark transparency- matches in the three surround conditions (color coded). The black and the gray line indicate matches that would be purely based on luminance or that would reflect perfect lightness constancy. Panel B: shows the same information for the light transparency condition. Error bars indicate ± one standard deviation.

Figure 8. Target contrast ranges across all reflectances and all images used in the experiment. The left group shows the contrast ranges in plain view for the three surround conditions (full, low, high) and the right group shows the three different viewing conditions for the full surround condition: plain view (p), dark transparency (dt), light transparency (lt).

The normalized contrast model (Zeiner & Maertens, 2014) is composed of three steps. The visual system computes Michelson contrasts across the image. It detects regions of different contrast ranges and identifies the region with the highest range as plain view. If the target contrast is located in a region of reduced contrast range, it is normalized relative to the full range. The normalization works to compute surface...
lightness, because the presence of a transparent medium leads to characteristic changes in the contrast range of the corresponding image regions, which are then undone by the contrast normalization.

The extreme surround conditions did not lead to a reduction in contrast range but to a substantial shift in contrast range relative to the full surround condition (Figure 8). The low and the high surround conditions are accompanied by respective shifts in contrast range relative to the full surround condition. These shifts are of considerable magnitude just as the range reductions that are introduced by the transparent media. We therefore consider the possibility that the shift in contrast range that was introduced by the extreme surround conditions might have signaled the presence of a different viewing condition even in the absence of concomitant geometrical cues. If the extreme surround conditions would be treated as a different viewing condition then they would equally trigger the contrast normalization.

So we applied the contrast normalization to the new contrast range that was introduced by the local surround. We called that the recursive contrast normalization because the local surround was nested in the viewing condition. Figure 9 shows the mean contrast matches of one observer as a function of target contrast together with the predictions from the recursive contrast computation. The model predictions mostly capture the magnitude of the differences between different surrounds in all viewing conditions. Data for individual observers are shown in Figure 14 of the Appendix. We quantified the predictive power of the recursive contrast computation in terms of the residual sum between model predictions and data. For plain view the average sum of residuals across all four observers is 0.16 (full), 0.26 (high), and 0.97 (low). For the dark transparency, it is 0.91 (full), 0.37 (high), and 1.07 (low). In the light transparency, we found the average

Figure 9. Mean matched contrasts and ± 1 SD as a function of target contrast. The recursive normalized contrast predictions for all three target surround conditions are plotted as open circles. Panel A shows the data for plain view. Panel B shows the data for the dark transparent context. Panel C shows the data for the light transparent context.

Figure 10. Mean matched contrasts and ± 1 SD as a function of target contrast. The global normalized contrast predictions for all three target surround conditions are plotted as open circles. Panel A shows the data for plain view. Panel B shows the data for the dark transparent context. Panel C shows the data for the light transparent context.
sum of residuals to be 1.14 (full), 0.48 (high), and 1.89 (low).

Global contrast normalization

The second model modification was motivated by the observation that lightness matches were predicted better when the surround term of the contrast equation included eight checks than when it included only the four checks that shared an edge with the target (Maertens & Shapley, 2013). We generalized this idea and computed a contrast measure in which the surround term included all image regions that belonged to the same viewing condition (i.e., the entire checkerboard in plain view and all checks behind the transparent medium in the transparency conditions). As a consequence, the extreme surround reflectances had a smaller effect on the contrast. So instead of computing the local target contrast we computed the global target contrast. The global target contrast was then normalized as before to account for range differences between the transparent media and

Figure 11. Contrast information present in each image presented to one observer in plain view for the high surround condition (Panel A) and the low surround condition (Panel B). Images (trials) were sorted according to target–target contrasts. The black lines indicate the overall contrast range within each image (trial). The filled black circle indicates the median of this distribution. Blue dots show the contrast values for the eight surround checks in each trial. The red dot indicates the target contrast.

Figure 12. Panel A shows the average data for each of our four observers from the dark transparency condition. Model predictions are plotted as a function of target contrast on top of the data. In Panel B, data from the light transparency condition are shown, in the same way as in Panel A.
plain view. Figure 10 plots matched contrasts as a function of target contrasts for the same observer as in Figure 9, together with the predictions from the global contrast normalization. The model captured the data well. Data for individual observers are shown in Figure 15 of the Appendix. We quantified the predictive power of the global contrast normalization model in terms of the residual sum between model predictions and data. For plain view the average sum of residuals across all four observers is 0.16 (full), 0.27 (high), and 0.23 (low). For the dark transparency, we found the average sum of residuals to be 1.07 (full), 0.71 (high), and 0.94 (low). In the light transparency, it is 1.16 (full), 0.77 (high), and 0.87 (low).

**Discussion**

In the present work we examined the role of Michelson contrast for perceived surface lightness in variegated checkerboard stimuli. In particular, we tested the effects of two context manipulations, both of which involved systematic changes in target contrast, on the perceived lightness of a target. First, we varied the viewing conditions under which a target surface was seen. The target was either presented in plain view, or behind a dark or a light transparent medium. The insertion of a transparent medium between target and observer resulted in a reduced and shifted contrast range in the projected image (Figures 2 and 8). Second, we manipulated the reflectances of the surfaces that surrounded the target. They were either randomly sampled from all possible reflectances, or they were selectively drawn from the five lowest or highest of all 12 possible reflectance values. The manipulation of surround reflectances also resulted in a shift in contrast ranges (Figure 8) but in the absence of geometric cues to occlusion.

Observers were lightness constant across the different viewing conditions. This was in line with previous results (Zeiner & Maertens, 2014) and predicted by the normalized contrast model. The extreme surround conditions had only a moderate effect on observers’ lightness judgments. The observed simultaneous contrast effects, or deviations from lightness constancy, were smaller than expected according to the normalized contrast model. Simultaneous contrast effects have been shown to vary in magnitude depending on the experimental conditions. Schmid and Anderson (2014) reported that simultaneous contrast effects were smaller and thus lightness constancy better for targets that were shown on realistic surfaces that had depth structure and were made of shiny material compared to targets shown on flat and matte surfaces. This is consistent with our findings and could imply that under more naturalistic viewing conditions, simultaneous contrast effects might be less prominent than previously thought.

**Model modifications**

To account for the observed effects of context, we suggested two model modifications. For the recursive contrast normalization we assumed that the local darkening or lightening of the check’s surround was interpreted by the visual system as a change in the viewing situation, such as an illumination change or the presence of a transparent medium. Accordingly we assumed that the concomitant change in contrast range would be undone by applying the contrast normalization recursively to the local surround. The corresponding model prediction resulted in good quantitative agreement with the data across the different surround conditions (Figure 9).

The idea that the visual system might treat photometric variations as spatial variations in the illumination has been suggested before (Allred et al., 2012), and Ekroll and Faul (2013) have demonstrated empirically that certain combinations of luminances at low contrast evoke perceptions of transparency that were not intended in the stimulus design. So it is possible that, unbeknownst to the experimenter, in the present stimulus the respective local darkening or lightening of the target surround has acted as a cue to image segmentation. However, this raises the question of how exactly the visual system performs the normalization of contrast ranges.

In order to compute the model predictions we have used contrast ranges that were derived from the distribution of target contrasts across trials. The resulting ranges for contrasts in the different surround and viewing conditions are depicted in Figure 8. However, in any single trial, the visual system sees only the contrasts that are present in that particular realization of a checkerboard. This comprises a limited sample of contrasts in plain view, a limited sample of contrasts in the region of transparency, and eight contrast values for the extreme surround checks. Figure 11 depicts these quantities for each individual trial in the plain view condition for the high (A) and the low surround condition (B). The range of contrasts of all checks in the checkerboard is shown as black lines, the median of that range as black dot, the contrasts of the eight surround checks in blue, and the contrast of the target in red.

As described above we used the minimum and maximum value of the red distribution, the target contrasts across trials, as a proxy for the range of contrasts in the extreme surround condition. However, the visual system is confronted only with a single image per trial, which is one individual line in the plot. In many trials the minimum and the maximum of the surround contrasts (blue dots) do not coincide with the minimum and maximum of the target contrasts (red dots). Thus, the range normalization, as we implemented it here, would fail if it were done based on the range of surround contrasts in a single trial. So, while the figure suggests that it could be possible to segregate the distribution of surround contrasts from the distribution of contrasts in...
the rest of the checkerboard, it is also evident that a normalization based on the contrast range within a single trial would vary from trial to trial depending on the statistics of that particular image. The present analysis points to the importance of contrast statistics for the process of segmentation. Using more naturalistic stimuli, it needs to be tested systematically how the spatial variation of contrast within an image affects perceived lightness in individual trials.

The second model modification was to include a larger area in the surround term of the contrast computation. As a consequence, the local surround manipulation had a smaller influence on target contrast (Figure 10). We computed what we called a global contrast normalization and it accounted well for the data. We have pointed out in the Introduction that a contrast measure can only be useful as a cue to surface reflectance if (a) all luminances that enter the contrast equation come from surfaces that are in the same illumination, and (b) when the normalizing surround term includes luminance values from a representative sample of surfaces seen within one illumination. Our extreme surround manipulations are examples for a violation of the second condition. By integrating across a larger area of space for the target surround, the visual system could partially compensate for this violation. This idea is conceptually similar to Adelson’s adaptive windows (Adelson, 2000). However, Adelson used the idea to explain why in his demonstration the simultaneous contrast effect was stronger for an articulated surround than for a homogeneous surround. He argued that the visual system requires a representative sample of luminance values to stabilize perception against coincidental luminance variations. When the surround is homogeneous, the number of available luminance values is too limited and therefore the visual system increases the size of the adaptive window. The adaptive window will then include a sufficient number of luminance values to reduce or even avoid a simultaneous contrast effect. In case of our global contrast normalization we also presume that the visual system enlarges the surround in the contrast computation to counteract the effect of the extreme surround conditions.

It is difficult to imagine a mechanism that performs a contrast computation across such a wide surround area, but the present data as well as the original conception of the usefulness of the contrast computation seem to imply this. Further study is required to test the effect of local versus global surround manipulation on perceived surface lightness in realistic stimuli.

Lightness perception through transparent media

So far we have discussed potential mechanisms of the visual system to deal with input variations that might have resulted from variation in the level of illumination. However, here we tested both local context changes and changes in the viewing conditions that were unambiguously due to the insertion of a transparent medium.

Contrast has been suggested not only as a cue to segment scenes into different contextual frameworks (Anderson, 1999; Singh, 2004; Singh & Anderson, 2002), but also to describe the perceived transmittance and the perceived reflectance of surfaces seen through the transparent medium (Robilotto, Khang, & Zaidi, 2002; Robilotto & Zaidi, 2004; Singh, 2004; Singh & Anderson, 2002). Singh (2004) and Singh and Anderson (2002) showed that when the episcotister model of transparency (Metelli, 1970) is reformulated in terms of Michelson contrast, it provides a more accurate account of perceived lightness through transparency and of perceived transmittance as well. The mapping between luminances seen in plain view and luminances seen behind a transparency was captured by the contrast ratio model (Singh, 2004). The model compensates for the contrast reduction in the region of transparency by means of a scaling between the contrast range in the region of transparency and the contrast range in plain view. Lightness matches according to the model are described as follows:

\[ y = \alpha_c \times x \]

whereby \( y \) is the match and \( x \) the target contrast. \( \alpha_c \) captures the magnitude of the contrast reduction in the region of transparency relative to plain view and it scales the contrasts accordingly. It is defined as follows:

\[ \alpha_c = \frac{\ln_{\text{max}}}{P_{\text{max}}} = \frac{T_{\text{max}} - T_{\text{min}}}{T_{\text{max}} + T_{\text{min}}} \]

whereby \( \ln_{\text{max}} \) is the maximum contrast in the region of transparency and \( P_{\text{max}} \) the maximum contrast in plain view. \( \ln_{\text{max}} \) is derived from the minimum (\( T_{\text{min}} \)) and maximum luminance (\( T_{\text{max}} \)) in the region of transparency, while \( P_{\text{max}} \) is derived from the minimum (\( P_{\text{min}} \)) and maximum luminance (\( P_{\text{max}} \)) in plain view.

Evidently, the contrast ratio model and our normalized contrast model incorporate the same conceptual mechanism. Given the high similarity between the two models, we compared the two models mathematically and with respect to their predictive power for our data. Singh (2004) developed the model for stimuli that were composed of only two luminance values in the surround. To create an equivalent situation for the stimuli used here, the target check would have to have the highest luminance and the eight surround checks would have to all have the same lowest luminance (and vice versa for the minimum contrast). We showed that for this special case the normalized contrast model and the contrast ratio model are mathematically equivalent (see Appendix).

For the stimuli used in the present experiment, the two models made different quantitative predictions. Figure 12 illustrates each model’s prediction as a function of target contrast for the dark (A) and the light (B) transparency and the full surround condition.
Average contrast matches are plotted together with the model predictions. The largest deviations between the two models were observed in the light transparency condition. Here, model predictions for the lowest target reflectances were almost 0.3 contrast units apart, while for the highest target reflectances, the model predictions were almost completely overlapping. To further evaluate the goodness of the two models in predicting our data we computed paired-sample t tests between each model prediction and the observed data for each reflectance condition. Significant differences between the predictions of the contrast ratio model and the observed data were found for the first four reflectance conditions in the light transparency condition \((p < 0.05)\). In the dark transparency condition only Reflectance 6 revealed a significant difference \((p < 0.05)\). For the normalized contrast model, significant differences between model predictions and observed data were shown for Reflectance 6 \((p < 0.05)\) in the light transparency. For the dark transparency, significant differences were found for Reflectances 3, 5, 6, 8, and 9 \((p < 0.05)\). All relevant test statistics can be found in Table 2 in the Appendix. To summarize, the normalized contrast model was better able to account for matches at low target contrasts. The contrast ratio model provided a better account for matches at intermediate contrasts. These differences in the predictions can be attributed to differences in the anchoring of the contrast ranges in each model. The contrast ratio model anchors the maximum contrast while the normalized contrast model anchors both minimum and maximum. To decide between the two models one needs to design stimuli for which the models would make maximally different quantitative predictions. This was beyond the scope of the present paper and remains to be tested in the future.

**Conclusions**

There have been two predominant theoretical approaches in lightness research: the inverse optics approach and the cue-based approach. The cue-based approach assumes that the visual system exploits cues in the retinal image to construct perceptual variables that systematically co-vary with invariant properties of the external world.

Here we show that a contrast-based model, which is a representative of the cue-based approach, can account for lightness constant matches across different viewing conditions, and for deviations from lightness constancy such as simultaneous contrast effects.

The model that is advocated here has been suggested in different contexts. Singh (2004) and Singh and Anderson (2002) introduced the contrast ratio model to account for perceived transmittance and perceived lightness in simple stimuli. Maertens and Shapley (2013) and Zeiner and Maertens (2014) suggested the normalized contrast model to account for lightness constancy across different viewing conditions in more naturalistic stimuli. We show (a) that the two models are formally equivalent for the kind of simple stimuli originally used by Singh (2004) and Singh and Anderson (2002), and (b) that they make similar quantitative predictions for the more naturalistic stimuli used here. The crucial factor is that the main currency of both models is contrast. The importance of contrast for the perception of lightness has been repeatedly pointed out in the recent past (Ekroll & Faul, 2013; Schmid & Anderson, 2014). However, the present results also raise questions about how the visual system might extract and represent information about contrast statistics that might either have to be pooled over a relatively large area in an image or, if not present in single images, would require some representation of those statistics over time.

**Keywords:** surface lightness, Michelson contrast, transparency, simultaneous contrast effect

**Acknowledgments**

This work was supported by the German Research Foundation (DFG MA5127/1-1 to Marianne Maertens).

Commercial relationships: None.

Corresponding Author: Christiane B. Wiebel.

Email: christiane.wiebel@tuberlin.de.

Address: Modelling of Cognitive Processes Group, Department of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, Berlin, Germany.

**References**


Appendix

In the study by Singh (2004) match and target were reversed, which is why the target contrast is multiplied with 1/z.
\[ y = \frac{x - t_{\text{min}}}{t_{\text{max}} - t_{\text{min}}} \cdot (p_{\text{max}} - p_{\text{min}}) + p_{\text{min}} \]
\[ y = (x - t_{\text{min}}) \cdot \frac{p_{\text{max}} - p_{\text{min}}}{t_{\text{max}} - t_{\text{min}}} + p_{\text{min}} \]
\[ y = (x - t_{\text{min}}) \cdot \frac{p_{\text{max}} - p_{\text{min}}}{t_{\text{max}} - t_{\text{min}}} - \left( \frac{p_{\text{min}} - p_{\text{max}}}{p_{\text{max}} + p_{\text{min}}} \right) \]
\[ y = (x - t_{\text{min}}) \cdot \frac{p_{\text{max}} - p_{\text{min}}}{t_{\text{max}} - t_{\text{min}}} - \left( \frac{p_{\text{min}} - p_{\text{max}}}{p_{\text{max}} + p_{\text{min}}} \right) \frac{T_{\text{max}} - T_{\text{min}}}{T_{\text{max}} + T_{\text{min}}} + p_{\text{min}} \]
\[ y = (x - t_{\text{min}}) \cdot \frac{p_{\text{max}} - p_{\text{min}}}{T_{\text{max}} - T_{\text{min}}} + p_{\text{min}} \]
\[ y = x \cdot \frac{T_{\text{min}} - T_{\text{max}}}{T_{\text{max}} + T_{\text{min}}} \cdot \frac{p_{\text{max}} - p_{\text{min}}}{p_{\text{max}} + p_{\text{min}}} - \frac{T_{\text{min}} - T_{\text{max}}}{T_{\text{max}} + T_{\text{min}}} \cdot \frac{p_{\text{max}} - p_{\text{min}}}{p_{\text{max}} + p_{\text{min}}} + p_{\text{min}} \]
\[ y = x \cdot \frac{T_{\text{min}} - T_{\text{max}}}{T_{\text{max}} + T_{\text{min}}} \cdot \frac{p_{\text{max}} - p_{\text{min}}}{p_{\text{max}} + p_{\text{min}}} - \left( 1 \cdot \frac{p_{\text{max}} - p_{\text{min}}}{p_{\text{max}} + p_{\text{min}}} \right) + p_{\text{min}} \]
\[ y = x \cdot \frac{T_{\text{min}} - T_{\text{max}}}{T_{\text{max}} + T_{\text{min}}} + \frac{p_{\text{max}} - p_{\text{min}}}{p_{\text{max}} + p_{\text{min}}} + \frac{p_{\text{min}} - p_{\text{max}}}{p_{\text{min}} + p_{\text{max}}} \]
\[ y = x \cdot \frac{T_{\text{min}} - T_{\text{max}}}{T_{\text{max}} + T_{\text{min}}} \]
\[ y = x \cdot \frac{p_{\text{max}}}{t_{\text{max}}} \]
\[ y = x \times \frac{1}{x_{c}} \]
Simulation of contrast distributions

To get an idea of the possible distribution of contrast values for the reflectance values used in our stimuli we computed the contrasts for \( n = 1,000 \) simulated configurations of target and surround reflectances. The surround reflectances were randomly sampled from all 12 possible reflectance values (see Table 1) considering the constraints that no two neighboring reflectances were identical and that the same reflectance did not occur more than twice (see Figure 16). We then used the luminances of one such sample of eight surround reflectances to calculate the respective Michelson contrast with each of the 12 target luminances. This procedure was repeated 1,000 times. For each target reflectance we plotted a histogram of the respective contrast values. The average contrast values for the two extreme and the full surround condition were computed with the average luminance in each condition in the surround term of the contrast equation.

Figure 13. The left panel shows mean luminance matches as a function of target luminance. The right panel shows mean contrast matches as a function of target contrast. The contrast-based model predictions are indicated by the unity line in the contrast–contrast plot. Data for each observer are illustrated row-wise. The first observer is equivalent to the one illustrated in the main text.
Figure 14. Mean matched contrasts as a function of target contrast. The recursive normalized contrast predictions for all three target surround conditions are plotted as linear fits, respectively. Average model predictions are plotted as open circles on top. Each viewing context is displayed in a separate panel (column-wise). Data for each observer are illustrated row-wise. The first observer is equivalent to the one illustrated in the main text.
Figure 15. Mean matched contrasts as a function of target contrast. The global normalized contrast predictions for all three target surround conditions are plotted as linear fits, respectively. Average model predictions are plotted as open circles on top. Each viewing context is displayed in a separate panel (column-wise). Data for each observer are illustrated row-wise. The first observer is equivalent to the one illustrated in the main text.
Figure 16. Illustration of the constraints for selecting checks surrounding the target. (A) Violation of position constraint: two identical reflectances were not allowed to be placed adjacent to each other as in Positions 1 and 2 in the Figure. (B) Violation of frequency constraint: The same reflectance value did not appear more than twice in the surround of the target. In the Figure the constraint is violated because the same reflectance is shown on three positions (1, 2, and 3).

<table>
<thead>
<tr>
<th>R</th>
<th>CRM t</th>
<th>p</th>
<th>NCM t</th>
<th>p</th>
<th>CRM t</th>
<th>p</th>
<th>NCM t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-3.19</td>
<td>0.05</td>
<td>1.59</td>
<td>0.21</td>
<td>-4.76</td>
<td>0.02</td>
<td>3.28</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>-3.19</td>
<td>0.05</td>
<td>3.27</td>
<td>0.05</td>
<td>-4.66</td>
<td>0.02</td>
<td>1.73</td>
<td>0.18</td>
</tr>
<tr>
<td>3</td>
<td>3.11</td>
<td>0.05</td>
<td>5.78</td>
<td>0.01</td>
<td>-4.94</td>
<td>0.02</td>
<td>0.83</td>
<td>0.47</td>
</tr>
<tr>
<td>4</td>
<td>0.86</td>
<td>0.45</td>
<td>2.89</td>
<td>0.06</td>
<td>-3.34</td>
<td>0.04</td>
<td>2.15</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>2.81</td>
<td>0.07</td>
<td>4.70</td>
<td>0.02</td>
<td>-1.46</td>
<td>0.24</td>
<td>1.19</td>
<td>0.32</td>
</tr>
<tr>
<td>6</td>
<td>9.88</td>
<td>0.00</td>
<td>16.00</td>
<td>0.00</td>
<td>-0.23</td>
<td>0.83</td>
<td>13.97</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.99</td>
<td>0.40</td>
<td>2.01</td>
<td>0.14</td>
<td>2.10</td>
<td>0.13</td>
<td>7.50</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>2.07</td>
<td>0.13</td>
<td>4.78</td>
<td>0.02</td>
<td>0.41</td>
<td>0.71</td>
<td>3.10</td>
<td>0.05</td>
</tr>
<tr>
<td>9</td>
<td>0.98</td>
<td>0.40</td>
<td>5.21</td>
<td>0.01</td>
<td>-0.94</td>
<td>0.42</td>
<td>1.57</td>
<td>0.21</td>
</tr>
<tr>
<td>10</td>
<td>-0.22</td>
<td>0.84</td>
<td>0.70</td>
<td>0.54</td>
<td>-0.55</td>
<td>0.62</td>
<td>2.77</td>
<td>0.07</td>
</tr>
<tr>
<td>11</td>
<td>-0.30</td>
<td>0.78</td>
<td>0.20</td>
<td>0.85</td>
<td>-2.00</td>
<td>0.14</td>
<td>0.94</td>
<td>0.42</td>
</tr>
<tr>
<td>12</td>
<td>-3.23</td>
<td>0.05</td>
<td>-2.30</td>
<td>0.10</td>
<td>-2.71</td>
<td>0.07</td>
<td>0.39</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 2. Paired-sample t test statistics: Model predictions versus observed data for the contrast ratio model (CRM) and the normalized contrast model (NCM) at every reflectance level (R).