

Perceptually Consistent Color-to-Gray Image Conversion

Shaodi You, Nick Barnes, Janine Walker

National ICT Australia (NICTA), Australian National University

Abstract. In this paper, we propose a color to grayscale image conversion algorithm (C2G) that aims to preserve the perceptual properties of the color image as much as possible. To this end, we propose measures for two perceptual properties based on contemporary research in vision science: brightness and multi-scale contrast. The brightness measurement is based on the idea that the brightness of a grayscale image will affect the perception of the probability of color information. The color contrast measurement is based on the idea that the contrast of a given pixel to its surroundings can be measured as a linear combination of color contrast at different scales. Based on these measures we propose a graph based optimization framework to balance the brightness and contrast measurements. To solve the optimization, an ℓ_1 -norm based method is provided which converts color discontinuities to brightness discontinuities. To validate our methods, we evaluate against the existing $\hat{\text{C}}\text{adik}$ and Color250 datasets, and against NeoColor, a new dataset that improves over existing C2G datasets. NeoColor contains around 300 images from typical C2G scenarios, including: commercial photograph, printing, books, magazines, masterpiece artworks and computer designed graphics. We show improvements in metrics of performance, and further through a user study, we validate the performance of both the algorithm and the metric.

1 Introduction

Color images are frequently presented in grayscale. This is common with digital ink displays and grayscale-only printing, but synthesizing a consistent grayscale image is also important for color blindness, and for prosthetic vision devices such as retinal implants[1] and sensory substitution devices [2] which mostly convey brightness only. One may also wish to present color information in grayscale as part of a virtual or augmented reality application. Such conversion is sometimes considered in image processing tasks such as tone-mapping and compression. For these reasons converting color images to grayscale images (C2G) has attracted a stream of research attention (*e.g.*, [3–5]).

Despite its common usage, C2G is non-trivial. Normal human vision perceives color with three peak spectral sensitivities [6]. Hence, perceivable color-space is homogeneous to \mathbb{R}^3 , whereas gray-scale space is homogeneous to \mathbb{R} . Standard displays present an image in a smaller discrete range though dynamic range and

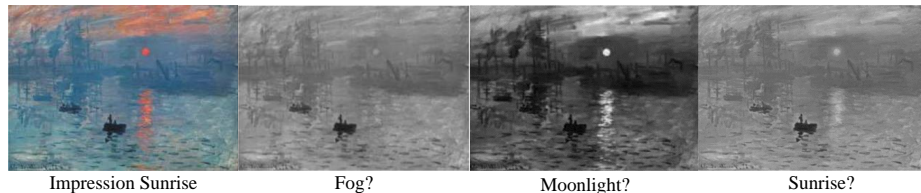


Fig. 1. Converting Claude Monet’s masterpiece ”Impression Sunrise” to grayscale. A conversion without considering the color contrast will change the feeling to haze or fog while a conversion without considering the brightness will change the feeling to moonlight.

quantization restrictions. Perception is restricted further by human observers’ ability to perceive brightness and color difference [7]. However, as perception of the presented grayscale image is also restricted, it is inevitable that information loss will occur in C2G.

More importantly, grayscale space has strong perceptual meaning. Simply quantizing the 3D signal and presenting it as a 1D signal often will not lead to a grayscale image that is perceptually consistent with the original color image . An example is shown in Fig. 1, Claude Monet’s masterpiece ”Impression Sunrise”, which is included in $\hat{\text{C}}\hat{\text{a}}\hat{\text{d}}\hat{\text{i}}\hat{\text{k}}$ dataset [8] and has been actively used for validating C2G algorithms. As can be seen, while the original color image gives the feeling of sunrise, that feeling can easily be lost in conversion to grayscale. A direct grayscale conversion without considering color contrast can change the feeling of the image to fog or haze. On the other hand, strongly enhancing contrast while allowing the brightness to change aggressively can change the feeling of the image to moonlight. Further, some representations of color can lead to an impression of the scene where the appearance is not natural. For example, consider the first row of Fig. 8, where as a result of methods that emphasize contrast over perceptual consistency, orange juice may appear black.

Another important consideration is that ”contrast” is a relative concept that is highly reliant on context. A famous example is the shadow illusion by Edward Adelson *et al.*[9]. While earlier research in C2G focused more on contrast in color space, current research considers spatial context. However, perceiving spatial context is rather complex. Thus, a perceptual consistent contrast model is desired.

In this paper, we propose a C2G algorithm which aims to preserve perceptual consistency as much as possible. To this end, we propose two perceptual consistency measurements based on contemporary research in vision science: brightness and multi-scale contrast measurement. The brightness measurement is based on the idea that brightness in a grayscale image does affect the perception of the probability of color information. The color contrast measurement is based on the finding that the contrast of a given pixel to its surroundings can be measured as a linear combination of color contrast in different scales. Based on these measures, we propose graphical model based optimization framework

to balance the brightness and contrast measurement. To solve the optimization, an ℓ_1 -norm based method is provided.

To validate our methods, we also propose a new dataset, called NeoColor, that improves on existing datasets. The Čadík dataset [8] has only 24 images. Further, the Color 250 dataset [10] is a subset of Berkeley segmentation dataset which was not designed to evaluate C2G algorithms. For NeoColor, we collect about 300 images from typical C2G scenarios including: natural scenes from commercial printing, books, magazines, masterpiece artworks, and computer designed graphics.

The contribution of our work is three-fold. Firstly, we propose a rationale for perceptual consistency of C2G algorithms. Based on perceptual consistency, we propose brightness and multi-scale color contrast measurements. Secondly, we propose an efficient algorithm for C2G which best preserves the perceptual consistency measurement. Thirdly, a new dataset which is specially designed for C2G evaluation is provided to address some shortcomings of the Čadík and Color250 datasets.

The rest of the paper is organized as follows: Section 2 summarizes related work on C2G. In Section 3, we discuss key aspects of human perception. Based on this, in Section 4, the perceptual consistency C2G algorithm is described. Section 5 demonstrates the effectiveness of the proposed method using both automatic evaluation and a user study. Section 6 concludes the paper.

2 Related Work

Existing C2G research can be categorized as global and local methods. Generally, global methods seek a global mapping for each color value to an intensity value, aiming to maximize the overall contrast, whereas local methods enhance spatially local contrast.

Global methods: Bala *et al.*[3] start with an image with a limited number of colors, the colors are sorted and brightness values are assigned accordingly. Gooch *et al.*[4] considered brightness variation according to hue, the orientation of the hue-ring was determined by maximizing the global contrast. Kim *et al.*[11] extended Gooch *et al.*'s method: variation from the hue-ring is represented by a Fourier series instead of the original linear mapping. Ancuti *et al.*[12] also extended the method of Gooch *et al.*, so that salient and non-salient areas have different hue-ring variation. A user study was also included to evaluate the effectiveness.

Rosche *et al.*[13,14] sought a continuous global mapping to assign similar colors to similar grayscale values, using pair-wise comparison between sampled colors. Using the same framework, Grundland *et al.*[15] improved the computational efficiency by using PCA to find the 1D component that maximizes contrast.

In more recent work, spatial context has been taken in consideration for global methods. Kuk *et al.*[16] proposed method to find landmark colors though K-means. Adopting similar strategy for spatial context, Lu *et al.*[17, 10] proposed

a bimodal contrast preserving energy function, along with local and non-local constraints. An evaluation metric and dataset were also introduced to automatically evaluate C2G algorithms. Ji *et al.*[18] proposed a hybrid method where color mapping is performance by modulating the hue ring and edges are sharpened by bandpass filtering. Yoo *et al.*[5] considered video C2G instead of single image C2G. They clustered the image sequence in spatio-temporal space to a 1D manifold, and assigned brightness accordingly.

Local methods: Bala *et al.*[19] proposed high pass filtering the color image to strengthen local contrast when converting to grayscale. Subsequently, Smith *et al.*[20] presented a method to pre-assign color variation based on vision science studies, and applied the variation to enhance local contrast. In recent work, Liu *et al.*[21] proposed an enhancement which preserves gradient, while attenuating the halo between inconsistent boundaries and local method of Du *et al.*[22] emphasised color contrast in salient areas.

Existing methods focus on best preserving contrast. However, inadequate focus has been given to ensuring that the image is consistent with human perception.

3 Human Perception of Color and Measurement

Modeling human perception of color in an algorithm is not trivial. It relies not only on physical modeling of object luminance and spectrum, but also on studies of human visual perception. Before we introduce our algorithm, which aims to produce a grayscale image that is consistent with human perception of the original color image, in this section, we discuss the factors that affect human color perception. We first introduce the basic concepts in color space and our notation. In Sec. 3.1 we discuss the perception of color brightness with a probability model, and introduce an energy function to model brightness perception. Then, in Sec. 3.2 we discuss perception of color contrast in complex spatial contexts and introduce an energy function to model complex color contrast perception.

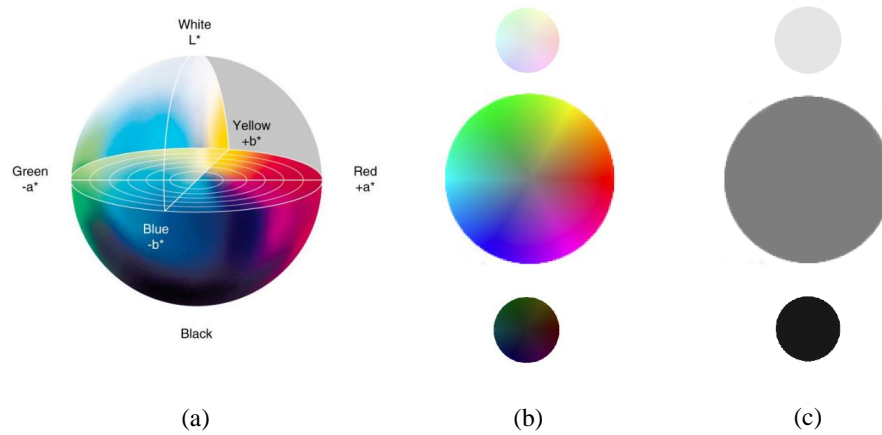
Background and notation

Color space: Since Young-Helmholz theory in the mid-1800s, it has been generally agreed that human perception of color is within a three dimensional space [23–26, 7, 27, 28]. In images, this can be represented in several common color spaces, *e.g.*, RGB, HSL, YUV, CIE $^*a^*b^*$. Because CIE-LAB preserves the perceived difference between colors, in this paper, we will use Euclidean distance in L $^*a^*b^*$ space as the contrast metric, which is common practice for recent C2G research [10, 22, 29, 21]. Figure 2.a is an illustration of the CIE $^*a^*b^*$ color space: a spherical space where lightness l varies from 0 for black to 100 for white. a varies from 100 for red to -100 for green; and b varies from 100 for yellow to -100 for blue.

Notation: Notation is listed in Table 1, for the convenience of the reader. Specifically, we denote the color channels of a pixel as l, a, b respectively and the targeted grayscale value as g . We use lower case for a single pixel, capitalized letter for image matrix and bold lower case for vectors.

Table 1. Notation

Symbols	Single pixel	Image Matrix	Image Vector
CIE L*a*b* Lightness	l	$L = \begin{bmatrix} l_{11} & \dots & l_{1m} \\ \vdots & \ddots & \vdots \\ l_{n1} & \dots & l_{nm} \end{bmatrix}$	$\mathbf{l} = [\dots l_{i+m+j} \dots]^T$
CIE L*a*b* a-channel	a	A	\mathbf{a}
CIE L*a*b* b-channel	b	B	\mathbf{b}
Grayscale image	g	G	\mathbf{g}
Color variance $v = \sqrt{a^2 + b^2}$	v	V	\mathbf{v}


Fig. 2. Color variance as function of brightness.

(a) The CIE L*a*b* color space. (b) A color slice in L*a*b* space with high, middle, low brightness. A high/low brightness color does not give same level of color contrast as mid-brightness color. (c) Desaturated slice of (b), a high/low brightness indicates a small color variation, while middle brightness indicates a large range for color variation.

3.1 Perception of Color Brightness and Perception Measurement

Color as function of brightness: As can be inferred from L*a*b* color space, color components with very low or very high brightness do not provide the same level of color variance as those with mid brightness levels. We show a simple example in Fig. 2.

Unlike previous C2G, especially global methods, [10, 15], where brightness is considered to be independent of color, we consider brightness as an importance cue for indicating color. As we presented in Fig. 2, although brightness does not allow us to infer the a, b value of a color, it does allow us to infer the possibility of color. This is important for generating a grayscale image that preserves the perceivable feeling of color.

Here we discuss the possibility of inferring color a, b as function of brightness l . Using the metric in CIE L*a*b* space, a color (l, a, b) with brightness $l \in$

$[0, 100]$ has a color variance, denoted as v , limited in the spherical color space:

$$v = \| (a(l), b(l)) \|_2 \leq \sqrt{100^2 - 4 * (l - 50)^2}, \quad (1)$$

Assuming a uniform distribution of color in CIEL*a*b* space, the partial probability of v by l is:¹

$$\begin{aligned} p(v = \tilde{v}|l) &\propto \tilde{v} \sqrt{100^2 - 4 * (l - 50)^2}, \\ 0 \leq \tilde{v} &\leq \sqrt{100^2 - 4 * (l - 50)^2}. \end{aligned} \quad (2)$$

and similarly:

$$\begin{aligned} p(l = \tilde{l}|v) &= \frac{2}{\sqrt{100^2 - v^2}}, \\ 50 - \frac{\sqrt{100^2 - v^2}}{2} &< \tilde{l} < + \frac{\sqrt{100^2 - v^2}}{2}. \end{aligned} \quad (3)$$

As can be seen, a color with a large v value is less likely to have a large variance in l , whereas a color with large l is less likely to have a large variance in v .

Brightness Perceptual Energy Considering Eq. (2) and Eq. (3), we define the brightness perceptual energy for inconsistent brightness between the color and grayscale image as:

$$\begin{aligned} E_B(g) &= w(l, a, b) \| l - g \|_\ell \\ &= \frac{1}{\sqrt{100^2 - a^2 - b^2} + \epsilon} \frac{1}{\sqrt{100^2 - (2l - 100)^2} + \epsilon} \| l - g \|_\ell, \end{aligned} \quad (4)$$

where ϵ is a small constant regulator. Generally, when the brightness in a color image l is not mapped to same grayscale value g , error energy increases. The error energy is reweighted by w which penalizes color with large variance v (because high probability color is not sensed); and color with high or low brightness l (because low probability of color is not sensed).

3.2 Perception of Multi-Scale Spatial Contrast and Measurement

In the previous subsection, we assume that the two areas of color appear close together spatially, and color contrast is only evaluated between the two. Much of the existing C2G research is based on such an assumption. However, color contrast is not absolute, but depends on spatial context. In this section, we start by considering pairwise color contrast and simple spatial context. Specifically, we introduce the Campbell-Robson curve, [23, 30]. Based on this, we introduce perception of complex spatial contrast, specifically considering recent research from vision science [25, 26]. Finally, we study the perception complex spatial color contrast as it appears in many real images.

¹ We use proportion instead of equal mark and neglect a constant for normalization.

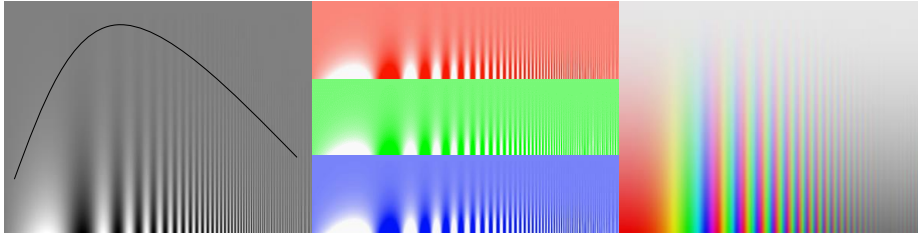


Fig. 3. Campbell-Robson curve.

Left, the original Campbell-Robson curve, brightness contrast sensitivity is a function of spatial frequency. Middle, Campbell-Robson curve on single color. We assume color brightness contrast follows the same curve. Right, curve on hue ring, the color hue contrast follows the same curve.

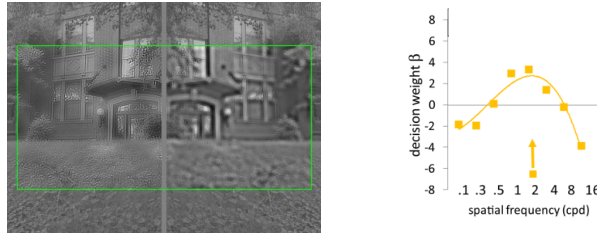


Fig. 4. The Haun-Peli curve for perceiving complex contrast [26]. Left, two gray scale images with difference combinations of intensity of spatial frequency component. Users are asked to choose the one with higher contrast. Right, statistics of importance of different frequency components for the overall perceived contrast.

Perception of simple spatial contrast: As illustrated in Fig. 3.a, the Campbell-Robson curve demonstrates that the perception of contrast is a function of spatial frequency. Specifically, human perception of contrast is most sensitive at 1-2 cycles per degree (cpd). The sensitivity drops for spatial frequency above and below the peak.

The Campbell-Robson study is only of luminance contrast. No equivalent study is available for the perception of color contrast. Although in practice, colour contrast is likely to exhibit spatial sensitivities that differ from luminance contrast, we take the simplest assumption that the perception of color contrast follows the same curve. This is shown in Fig. 3.b for the perceptual sensitivity of color-brightness contrast with respect to spatial frequency, and in Fig. 3.c, for the perception of hue contrast. Our subsequent algorithm could be adapted if curves like that Campbell-Robson became available for color contrast. Thus, **Assumption 1** the perception of color contrast (Brightness, Hue, Saturation) with respect to spatial frequency follows the same Campbell-Robson curve.

Perception of complex spatial contrast: However, the spatial contrast of a natural image is often far more complex than the Campbell-Robson pattern. According to Haun *et al.*[26], human perception of complex contrast follows a linear combination of different spatial frequencies. In Fig. 4, we include a typical result

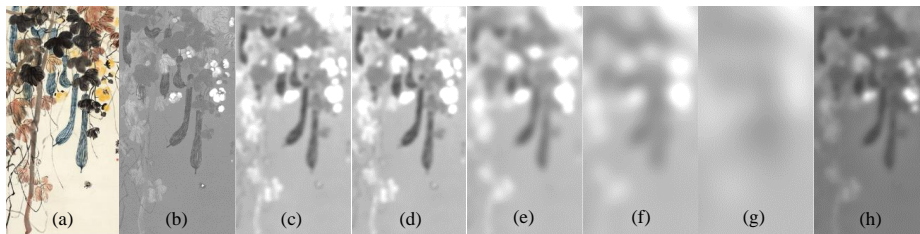


Fig. 5. Complex color contrast. (a) Color image, (b) b channel of L*a*b* spece. (c - g) contrast in different scale, using DoG filter. (h) Combined complex contrast according to Haun-Peli curve.

of such a combination. Note that the coefficients for such a combination need not all be positive.

Perception of complex spatial color contrast: Unfortunately, no systematic vision science model is available for complex color contrast. Therefore, here again, we make an assumption that the perception of complex brightness contrast can be extended to complex color contrast:

Assumption 2 the perception of complex color contrast with respect to spatial frequency follows the same Haun-Peli curve.

Complex Color Contrast Consistency Measurement: According to the above assumptions, we introduce the Brightness Contrast Measurement. For a single channel (grayscale) image the complex contrast is defined as:

$$c(G) = \sum \beta_i k_i * G, \quad (5)$$

where k_i is the Difference of Gaussian (DoG) Kernel at scale 2^i :

$$k_i(j, k) = \frac{1}{2\pi 2^{2i}} e^{-\frac{j^2+k^2}{2^{2i}}} - \frac{1}{2\pi 2^{2(i+1)}} e^{-\frac{1}{2} \frac{k^2+k^2}{2^{2(i+1)}}}. \quad (6)$$

β_i are the coefficients derived from the Haun-Peli curve. For example, a screen with 72 dots per inch (dpi), and viewed from 60cm away.

$$\beta_1 = -4, \beta_2 = 1, \beta_3 = 4, \beta_4 = 4, \beta_5 = 1, \beta_6 = -2. \quad (7)$$

When varying the dpi or the viewing distance, the scale by pixels shifts proportionally. Figure 5 shows an example of the contrast at different scales and the combined complex contrast.

We define the complex contrast for the color channels similarly and denote them as $c(L), c(A), c(B)$. The contrast measure is defined as:

$$E_c(G) = \alpha_L \|c(L) - c(G)\|_\ell + \alpha_{AB} \|c(A) - c(G)\|_\ell + \alpha_{AB} \|c(B) - c(G)\|_\ell, \quad (8)$$

where α_L, α_{AB} balances the importance between different channels. As can be seen, the contrast energy functions has three terms that represent the contrast lost for the l, a, b channels respectively. Note that the contrast loss for the a, b channels are directional, this is because we incorporate the understanding of human vision that the brightness of same color is not always perceived in the same way [31].

4 Methodology

In this section, we introduce a graphical model based method which aims to find a gray-scale image which best fits the brightness perceptual consistency as well as the contrast consistency.

Objective Function Combining Eq. (4) and Eq. (8), the proposed objective function seeks the gray-scale image G that optimizes the trade-off between brightness consistency and contrast:

$$E(G) = E_B(G) + E_C(G) \quad (9)$$

Similarly, Eq. (9) can be represented as linear system:

$$\begin{aligned} E(\mathbf{g}) &= E_B(\mathbf{d}) + E_C(\mathbf{g}) \\ &= \| \Lambda(\mathbf{w})(\mathbf{l} - \mathbf{g}) \|_\ell \\ &\quad + \alpha_L \| C\mathbf{l} - C\mathbf{g} \|_\ell + \alpha_{AB} \| C\mathbf{a} - C\mathbf{g} \|_\ell + \alpha_{AB} \| C\mathbf{b} - C\mathbf{g} \|_\ell, \end{aligned} \quad (10)$$

where the bold letters are the vectorized representation of the image matrix. C is the complex contrast operator in matrix form. Λ converts the coefficients array \mathbf{w} to a diagonal matrix. A typical choice of channel importance is $\alpha_L = 0.5, \alpha_{AB} = 1.5$. We will fix such setting for all the experiments.

ℓ_2 -norm Solution: When $\ell = 2$, Eq. (10) takes a quadratic form, which is minimized when:

$$\begin{aligned} \frac{\partial E(\mathbf{g})}{\partial \mathbf{g}} &= 0 \\ &= -\Lambda(\mathbf{w})(\mathbf{l} - \mathbf{g}) \\ &\quad - \alpha_L C^\top (C\mathbf{l} - C\mathbf{g}) - \alpha_{AB} C^\top (C\mathbf{a} - C\mathbf{g}) - \alpha_{AB} C^\top (C\mathbf{b} - C\mathbf{g}), \end{aligned} \quad (11)$$

This can be written as:

$$M\mathbf{g} = \mathbf{u}, \quad (12)$$

where:

$$M = \Lambda(\mathbf{w}) + (\alpha_L + 2\alpha_{AB})C^\top C, \quad (13)$$

and

$$\mathbf{u} = \Lambda(\mathbf{w})\mathbf{l} + \alpha_L C^\top C\mathbf{l} + \alpha_{AB} C^\top C\mathbf{a} + \alpha_{AB} C^\top C\mathbf{b}. \quad (14)$$

Notice that M is invertable because:

1. $\Lambda(\mathbf{w})$ is a positive definite diagonal matrix.
2. $C^\top C$ is semi-positive definite.
3. α_L, α_{AB} are positive.

1, 2, 3 together infer that M is a positive-definite symmetric matrix and is thus invertable [32]. Therefore, the grayscale image is found by:

$$\mathbf{g} = M^{-1}\mathbf{u}. \quad (15)$$

ℓ_1 -norm Solution: It is well known that ℓ_2 norm based objective functions tend to over-smooth boundaries, this problem is significant for color-to-gray conversion. This is because a grayscale image is intended to preserve opposing terms: brightness and multi-channel color contrast. Unfortunately, it is unlikely to find a solution to suit the contrast for all three channels simultaneously. The ℓ_1 solution allows us to neglect a relatively insignificant channel and strongly emphasize

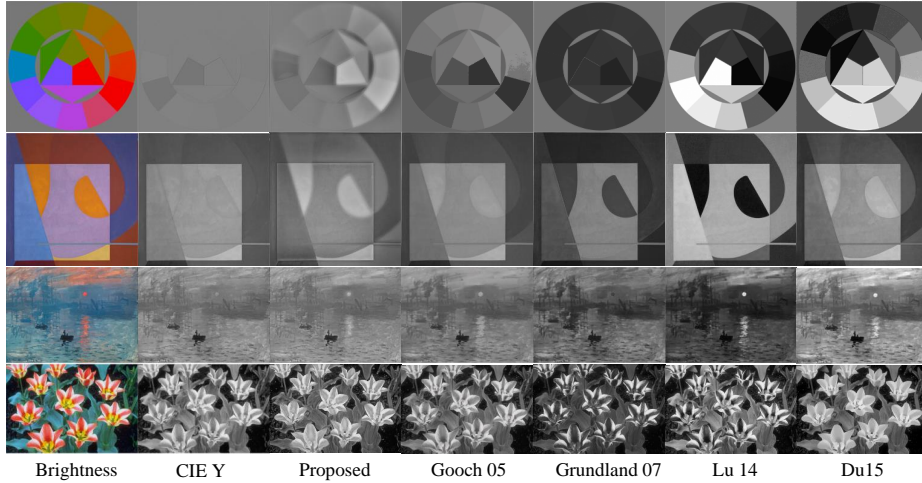


Fig. 6. Visual results on $\hat{\text{Cadfk}}$ dataset. (Optimized for A4 paper viewed from 40cm away).

the contrast of a more significant channel. In addition we take advantage of using ℓ_1 where discontinuity at boundaries is implicitly handled in the graphical model.

To solve the ℓ_1 objective function, we use the Iteratively Re-weighted Least Square method [33]. For i -th iteration, we solve a reweighted least square problem of Eq. (10):

$$\min E(\mathbf{g}_i) = \min \Lambda(\mathbf{q}_i)[E_B(\mathbf{d}) + E_C(\mathbf{g}_i)], \quad (16)$$

where the weight is defined as:

$$\mathbf{q}_{i+1} = \frac{1}{|\mathbf{r}_i| + 1}, \quad (17)$$

w_0 is set to 1 for all the pixels, and \mathbf{r} is the residual of each iteration:

$$\mathbf{r}_i = \Lambda(\mathbf{q}_i)[|\Lambda(\mathbf{w})(\mathbf{l} - \mathbf{g}_i)| + \alpha_L |\mathbf{C}\mathbf{l} - \mathbf{C}\mathbf{g}_i| + \alpha_{AB} |\mathbf{C}\mathbf{a} - \mathbf{C}\mathbf{g}_i| + \alpha_{AB} |\mathbf{C}\mathbf{b} - \mathbf{C}\mathbf{g}_i|]. \quad (18)$$

Speedup: Constructing the complex contrast matrix C will yield a large dense matrix with N^2 elements where N is the number of pixels, as a large scale DoG compares the difference over a large range. Fortunately, we find that the blurring matrix C can be pre-computed using decomposition of the Gaussian kernel and the construction of an image pyramid. Details of the speedup can be found in the supplementary material.

5 Experiments

We validate the proposed method systematically. We compare it with methods for which code are available, on existing datasets as well as the new proposed dataset. We evaluate qualitative performance, as well as performance with respect to metrics. Further, we include a user study to validate the method and the proposed metric.

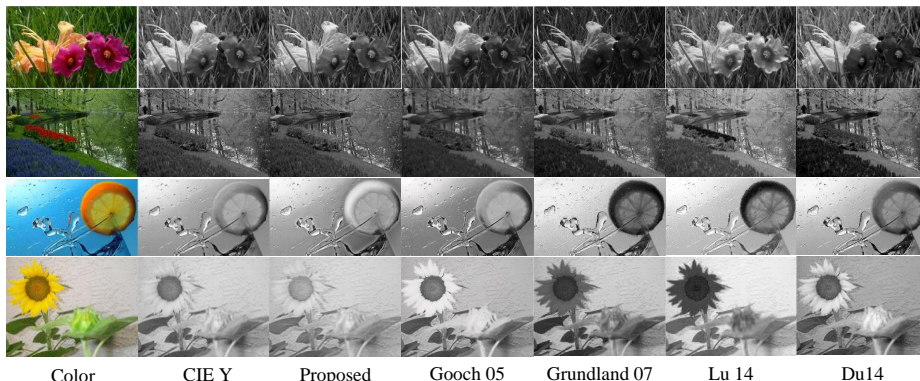


Fig. 7. Visual results on Color250 dataset. (Optimized for A4 paper viewed from 40cm away).

5.1 Datasets

Čadík [8] is the earliest dataset for C2G evaluation, which contains 25 highly saturated images. The contents varies from geometric pattern to real scenes. Because the Čadík dataset has only 25 images, as is common practice, we evaluate the performance on the *Color250* dataset[10]. This data set is a subset of the 2001 Berkeley Segmentation Dataset [34]. *CSDD* [22] contains 22 highly saturated and fully chromatic geometric patterns. Because the number of images is limited and no natural image are included, we do not use it in our evaluation. However, we include the results on the CSDD dataset in the supplementary material.

NeoColor The Berkeley Segmentation Dataset was designed for evaluating segmentation performance rather than evaluating C2G algorithms. Images in the dataset tend to have strong contrast on object boundaries, only contain a limited number of segments with different colors, and the contents are relatively simple. On the other hand, C2G algorithms must be able to deal with more complex images that frequently arise.

To fully evaluate the performance of C2G algorithms, we have collected high quality digital images of natural scenes, advertisements, computer graphic designs as well as fine art. We consider that this new dataset, NeoColor, extends upon the existing datasets by including images of greater color complexity. A full collection of the dataset appears in the supplementary material. Some examples are shown in Fig. 8.

5.2 Evaluation metric

We evaluate our methods using *EScore* proposed by Lu *et al.*[10] which has a joint evaluation between color contrast preservation and color fidelity. However, the metric emphasizes preservation of high contrast, and measurement of spatial contrast preservation is based on random sampling which does not fully reflect the concept on visual perception preservation. The *QScore* Ma *et al.*[29], has better quantization for color contrast measurement, however, scale dependent contrast is not emphasized.

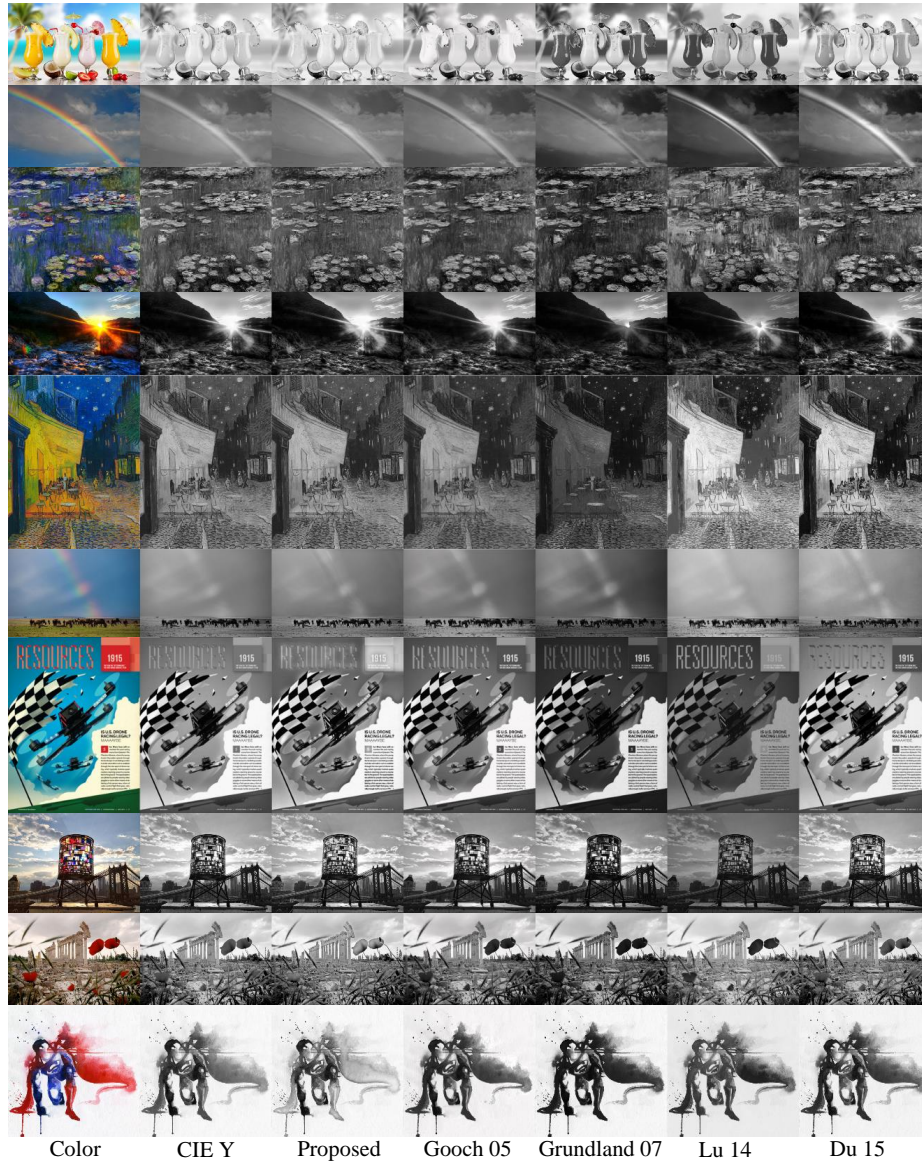


Fig. 8. Visual results on NeoColor dataset. (Optimized for A4 paper viewed from 40cm away).

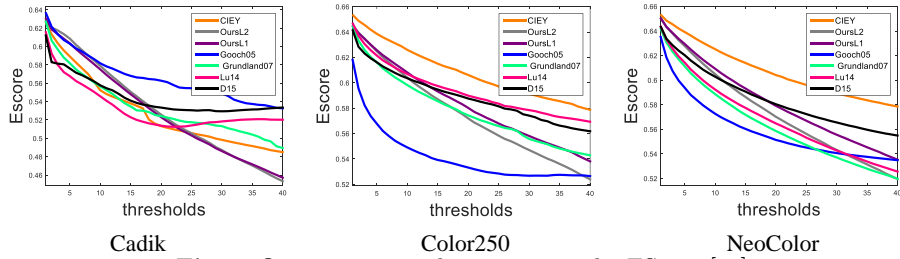


Fig. 9. Quantitative evaluation using the EScore [10].

The y-axis the EScore, where as the x-axis the contrast threshold level.

Visual Perception Metric (PScore) Because both metrics focus more on contrast preservation rather than the overall perception of color, we provide further evaluation using the proposed perceptually consistent color energy function Eq. (9). For a fair comparison, we weight all scales of contrast preservation separately, we set $\beta_i = 1$ (Eq. (5)) for the evaluated scale and rest as 0.

5.3 Results

We compare with classic methods of Gooch *et al.*[4] and Grundland *et al.*[15] as well as the state-of-the-art methods of Lu *et al.*[10] and Du *et al.*[22]. We use the default setting for all the datasets. For our method, scale is set to optimize contrast for A4 paper viewed from 40cm away, and all parameters are fixed for all the experiments. Code is not available² for the methods of Liu *et al.*[21] and Ji *et al.*[18] and thus they are not evaluated. We report both visual and quantitative results.

Visual results Visual results on the $\hat{\text{C}}\text{adik}$ dataset, the Color250 dataset and NeoColor dataset are shown in Fig. 6, Fig. 7 and Fig. 8 respectively. A full collection of the results can be found in the supplementary material.

As can be seen, methods that do not consider perceptual consistency are likely to produce strong contrast but unnatural images. For example in Fig. 8, black orange juice, a rainbow with black and white stripes, and a black sun in a sunset image, all lose their natural impression. Further, the impressionist painting "Lotus" which now has more emphasis on weeds has lost the original feeling of the image, and the intention of the artist. On the contrary, the proposed methods allows a reasonable sense of color contrast as well as the original feeling of the images.

Quantitative results: Quantitative evaluation using the E-score is shown in Fig. 9, a higher curve indicates better performance. As can be seen, the $\hat{\text{C}}\text{adik}$ dataset is not large enough to produce a smooth curve. On the Color250 dataset, our method performs well on contrast in the range 0-20 (where total black to white was a scale of 100).

The Q-score proposed by Ma *et al.*[29] is shown in Fig. 10. Whereas the evaluation for most C2G algorithms was reasonable, CIE Y which does not consider color contrast is evaluated as the best on all the datasets.

² Not available on webpage as well as through email contact

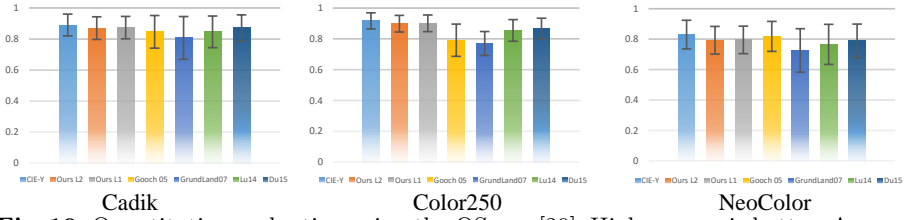


Fig. 10. Quantitative evaluation using the QScore [29]. Higher score is better. Average score and standard variance is shown.

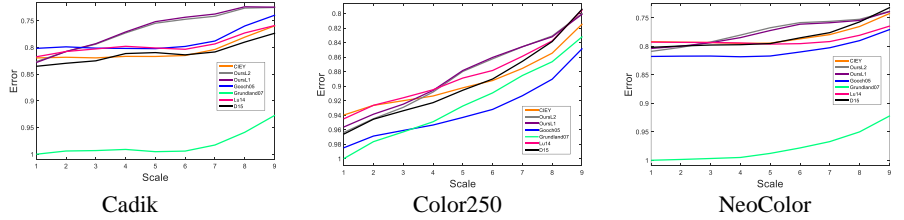


Fig. 11. Quantitative evaluation using proposed PScore. The y-axis is the PScore. The x-axis represents the natural contrast preservation at scale $2^{\frac{i+2}{2}}$ pixels.

Results for consistency (P-Score) are shown in Fig. 11. Again, a higher curve shows better performance. The x-axis represents contrast preservation at scale $2^{\frac{i+2}{2}}$ pixels. We plot the contrast preservation for different scales. As can be seen, the proposed method best preserves the color contrast at the scale which is most sensitive for perception. Whereas for the scales which are less contributive to overall contrast, the proposed method has lower contrast preservation in comparing with other methods. The ordering is accordance with the feeling on visual results.

5.4 User Study

Ten volunteers (aged 22-41) with normal or corrected-to-normal vision took part, with data collection in March 2016. Written consent was obtained from participants before they began the experiment. The research was approved by the *institution removed* Ethics Committee, and adhered to the tenets of the Declaration of Helsinki.

Considering the complexity of the task, four methods were compared: the proposed, CIE Y (as a widely used method), Lu *et al.*(as the state-of-the-art), Grundland *et al.*(as reported relatively worse from the quantitative metric). Participants were masked to the algorithm. The color image was presented in the middle of the screen, with a white background, and the four grayscale images presented at identical distances from the original image. Please refer to supplementary material for screen layout and experimental details. We selected 50 hard images from Color250, and 200 images from NeoColor, excluding more difficult images. Then, 150 images were randomly selected from these subsets with a 30% selection probability from Color250 and a 70% probability from NeoColor. Each user was shown these images in random order, and placement of the grayscale images was randomly assigned for each trial.

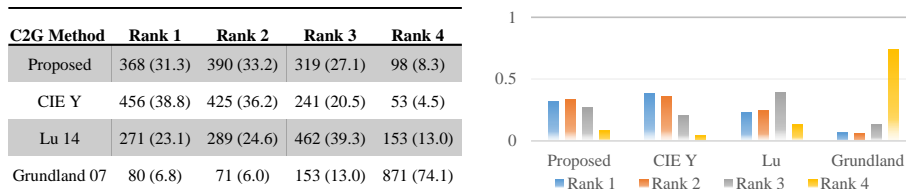


Fig. 12. User preference on C2G algorithms.

Left, statics of users’ ranking by number of choices. Parenthesis list the percentage. Right, ranking frequency distribution of all the methods.

Participants were required to rank the grayscale images from best (1) to worst (4) according to the instruction: “Which black-and-white image best represents the original color image”. When answering, the participants were asked to consider: which grayscale image best retains the contrast between colors of the original color image; and, which grayscale image best retains the overall sense or feeling of the original color image. The chance rate of ranking a image as “best” (1 and 2) representation of the color image was 50%, and 75% was the criterion set as a benchmark as a reliable grayscale representation of the color image (*i.e.*, participants ranking a Color2Gray vision processing method “1” or “2” for 75% of the trials). Descriptive analyses were used to characterize the counts and percentages of responses. Comparisons between participants and color-to-gray methods for the average rankings of the preferred images were calculated using non-parametric statistics (e.g., Kruskal-Wallis H test). Windows SPSS v23 (IBM Corporation, Somers, NY) was used for all statistical analyses.

Results: The data for ten participants was pooled for further analyses as there was no significant difference in overall preferences between participants ($p < 1.00$). Overall preference is shown in Fig. 12. Our proposed C2G method was ranked as a significantly ($p < 0.0001$) better representation of the original color images compared to the Grundland and Lu methods. However, the CIE-Y approach was selected significantly ($p < 0.0001$) more often compared to the proposed method. Our proposed method was ranked as being a better (ranked 1 or 2) representation of the original color image for 64.5% of responses, and CIE-Y achieved 75.0% of responses ranked as 1 or 2. The CIE-Y approach is similar to approaches that are ubiquitous in print media (*e.g.*, newspapers, kindle). Among the specialized C2G algorithms, our proposed method was highest ranked. The Grundland and Lu methods were consistently ranked as the least preferred methods with only 12.8% and 47.7% of responses ranked 1 or 2.

6 Conclusion

In this paper, we proposed a method for converting color images to gray scale while preserving brightness consistency as much as possible. The key ideas of brightness consistency is that the feeling and naturalness of the original color is preserved, and that color contrast is preserved. These key ideas were used to derive quantitative metrics based on recent studies in vision science. An ℓ_1 optimization framework was proposed to find the grayscale image which optimised

the proposed brightness consistency metric. To evaluate the proposed methods, we used both existing datasets and a proposed new dataset. We also validated both the algorithm and the metric with a user study. For future work, we will exploit brightness consistency with more emphasis on visual augmentation.

References

1. Ayton, L.N., Blamey, P.J., Guymer, R.H., Luu, C.D., Nayagam, D.A.X., Sinclair, N.C., Shivdasini, M.N., Yeoh, J., McCombe, M.F., Briggs, R.J., Opie, N.L., Villabos, J., Dimitrov, P.N., Varsamidis, M., Petoe, M.A., McCarthy, C.D., Walker, J.G., Barnes, N., Burkitt, A.N., Williams, C.E., Shepherd, R.K., Allen, P.J., the Bionic Vision Australia Consortium: First-in-human trial of a novel suprachoroidal retinal prosthesis. *PLOS ONE* **9** (2014)
2. Stronks, H., Parker, D., Walker, J., Lieby, P., Barnes, N.: The feasibility of coin motors for use in a vibrotactile display for the blind. *Artif. Organs* **39** (2015) 480–91
3. Bala, R., Braun, K.M.: Color-to-grayscale conversion to maintain discriminability. In: *Electronic Imaging 2004, International Society for Optics and Photonics* (2003) 196–202
4. Gooch, A.A., Olsen, S.C., Tumblin, J., Gooch, B.: Color2gray: salience-preserving color removal. In: *ACM Transactions on Graphics (TOG)*. Volume 24., ACM (2005) 634–639
5. Yoo, M.J., Lee, I.K., Lee, S.: Color sequence preserving decolorization. In: *Computer Graphics Forum*. Volume 34., Wiley Online Library (2015) 373–383
6. Pokorny, J., Smith, V.C.: Chromatic Discrimination. In: *The Visual Neurosciences: Volume 2*. MIT Press, Cambridge Massachusetts (2004) 908–923
7. Stevens, S.S.: On the psychophysical law. *Psychological review* **64** (1957) 153
8. Čadík, M.: Perceptual evaluation of color-to-grayscale image conversions. In: *Computer Graphics Forum*. Volume 27., Wiley Online Library (2008) 1745–1754
9. Adelson, E.H.: $Z_j \pm$ lightness perception and lightness illusions. *New Cogn. Neurosci* **339** (2000)
10. Lu, C., Xu, L., Jia, J.: Contrast preserving decolorization with perception-based quality metrics. *International journal of computer vision* **110** (2014) 222–239
11. Kim, Y., Jang, C., Demouth, J., Lee, S.: Robust color-to-gray via nonlinear global mapping. In: *ACM Transactions on Graphics (TOG)*. Volume 28., ACM (2009) 161
12. Ancuti, C.O., Ancuti, C., Bekaert, P.: Enhancing by saliency-guided decolorization. In: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE* (2011) 257–264
13. Rasche, K., Geist, R., Westall, J.: Detail preserving reproduction of color images for monochromats and dichromats. *Computer Graphics and Applications, IEEE* **25** (2005) 22–30
14. Rasche, K., Geist, R., Westall, J.: Re-coloring images for gamuts of lower dimension. In: *Computer Graphics Forum*. Volume 24., Wiley Online Library (2005) 423–432
15. Grundland, M., Dodgson, N.A.: Decolorize: Fast, contrast enhancing, color to grayscale conversion. *Pattern Recognition* **40** (2007) 2891–2896
16. Kuk, J.G., Ahn, J.H., Cho, N.I.: A color to grayscale conversion considering local and global contrast. In: *Computer Vision–ACCV 2010*. Springer (2010) 513–524
17. Lu, C., Xu, L., Jia, J.: Real-time contrast preserving decolorization. In: *SIGGRAPH Asia 2012 Technical Briefs, ACM* (2012) 34
18. Ji, Z., Fang, M.e., Wang, Y., Ma, W.: Efficient decolorization preserving dominant distinctions. *The Visual Computer* (2015) 1–11
19. Bala, R., Eschbach, R.: Spatial color-to-grayscale transform preserving chrominance edge information. In: *Color and Imaging Conference*. Volume 2004., Society for Imaging Science and Technology (2004) 82–86

20. Smith, K., Landes, P.E., Thollot, J., Myszkowski, K.: Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. In: Computer Graphics Forum. Volume 27., Wiley Online Library (2008) 193–200
21. Liu, Q., Liu, P.X., Xie, W., Wang, Y., Liang, D.: Gcsdecolor: Gradient correlation similarity for efficient contrast preserving decolorization. Image Processing, IEEE Transactions on **24** (2015) 2889–2904
22. Du, H., He, S., Sheng, B., Ma, L., Lau, R.W.: Saliency-guided color-to-gray conversion using region-based optimization. Image Processing, IEEE Transactions on **24** (2015) 434–443
23. Campbell, F.W., Robson, J.: Application of fourier analysis to the visibility of gratings. The Journal of physiology **197** (1968) 551
24. Lotto, R.B., Purves, D.: A rationale for the structure of color space. Trends in neurosciences **25** (2002) 84–89
25. Peli, E.: Contrast in complex images. JOSA A **7** (1990) 2032–2040
26. Haun, A.M., Peli, E.: Perceived contrast in complex images. Journal of vision **13** (2013) 3–3
27. Majumder, A., Irani, S.: Perception-based contrast enhancement of images. ACM Transactions on Applied Perception (TAP) **4** (2007) 17
28. Jaroensri, R., Paris, S., Hertzmann, A., Bychkovsky, V., Durand, F.: Predicting range of acceptable photographic tonal adjustments. In: Computational Photography (ICCP), 2015 IEEE International Conference on, IEEE (2015) 1–9
29. Ma, K., Zhao, T., Zeng, K., Wang, Z.: Objective quality assessment for color-to-gray image conversion. Image Processing, IEEE Transactions on **24** (2015) 4673–4685
30. Oliva, A., Torralba, A., Schyns, P.G.: Hybrid images. ACM Transactions on Graphics (TOG) **25** (2006) 527–532
31. Corney, D., Haynes, J.D., Rees, G., Lotto, R.B.: The brightness of colour. PloS one **4** (2009) e5091
32. Lang, S.: Algebra, volume 211 of graduate texts in mathematics (2002)
33. Candes, E.J., Wakin, M.B., Boyd, S.P.: Enhancing sparsity by reweighted ℓ_1 minimization. Journal of Fourier analysis and applications **14** (2008) 877–905
34. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proc. 8th Int’l Conf. Computer Vision. Volume 2. (2001) 416–423