Color Homography: theory and applications

Graham D. Finlayson[1,*], Han Gong[1,*], and Robert B. Fisher[2]

[1]School of Computing Sciences, University of East Anglia
[2]School of Informatics, University of Edinburgh
[*]denotes co-first authors.

September 24, 2017

Abstract

Images of co-planar points in 3-dimensional space taken from different camera positions are a homography apart. Homographies are at the heart of geometric methods in computer vision and are used in geometric camera calibration, 3D reconstruction, stereo vision and image mosaicking among other tasks. In this paper we show the surprising result that homographies are the apposite tool for relating image colors of the same scene when the capture conditions – illumination color, shading and device – change.

Three applications of color homographies are investigated. First, we show that color calibration is correctly formulated as a homography problem. Second, we compare the chromaticity distributions of an image of colorful objects to a database of object chromaticity distributions using homography matching. In the color transfer problem, the colors in one image are mapped so that the resulting image color style matches that of a target image. We show that natural image color transfer can be re-interpreted as a color homography mapping.

Experiments demonstrate that solving the color homography problem leads to more accurate calibration, improved color-based object recognition, and we present a new direction for developing natural color transfer algorithms.

1 Introduction

In image formation there are two important parts, the geometry of how points in space map to image locations and the photometry of how illumination, surface reflectances and camera sensors combine to form the colors in an image. Broadly, the mathematical tools underlying our understanding of image geometry are non-linear reflecting the non-linear perspective nature of image formation. Important non-linear concepts include "solving for the homography" (*e.g.* relating subsequent frames in panorama stitching [1]) and epipolar geometry in stereo vision [2]). In contrast, the majority of methods in color/photometric computer vision are linear which, at least for simplified scenes such as the eponymous Mondrian world [3, 4] (the world consists of a patchwork of flat co-planar reflectances), reflects the physics of how images are formed. Linear color problems include, color correction [5–7] (*e.g.* mapping RAW colors from camera to display RGB) and modeling illuminant color change [8] *e.g.* for color object recognition [9].

In Figure 1a, we illustrate a homography as the term applies in geometric computer vision. Here π_1 might denote the image – a perspective projection – of a plane (in 3-dimensions) and π_2 denotes the same plane viewed in a second image. The homography H relates the two planes.

In Figure 1b, Ball_1 is the image of the side-view of a 4-color ball where the ball is lit from behind the camera with a white light. The same ball is lit from above with a bluish light, image Ball_2. The images are carefully registered so they are in pixel-wise correspondence. In an attempt to color correct panel Ball_2 to match Ball_1, we, naively, carry out a linear regression – for locations where both images have non-zero response – and find the best 3×3 matrix M mapping the corresponding pixels. The result of the regression results in image Ball_3. Notice that while the color cast due to the bluish light appears lessened, viewed closely, the colors are incor-
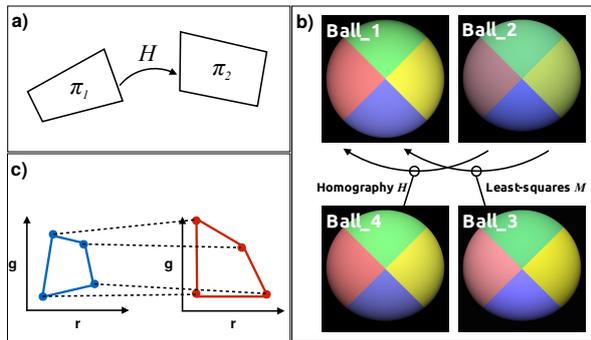
Figure 1: Top left, panel (a), images of two planes are related by a homography. Right, panel (b), 4 images of a colored ball are shown. Ball_1 is the reference image where the illumination color is white and placed behind the camera. Ball_2 is the object illuminated with a blue light from above. Respectively, Ball_3 and Ball_4 are the least-squares mapping and the homography match (in both cases the aim is to *correctly* undo the illumination color) from Ball_2 to Ball_1, Bottom right, panel (c), the chromaticities from Ball_2 matched to corresponding chromaticities in Ball_1.

rectly mapped. In particular, notice that the red color segment looks wrong. Now, we now transform Ball_2 image to the image Ball_4 using the *correct* linear transform H.

In this paper, we propose that to map one *photometric* view to another we must map the colors correctly independent of shading. Since shading only affects the brightness, or magnitude, of the RGB vectors we wish, in effect, to find the 3×3 map which maps the color *rays* (the RGBs with arbitrary scalings) in one photometric view to corresponding rays in another. We note that this "ray matching" is precisely the circumstance in geometric computer vision when co-planar points in an image are mapped, via a homography, to corresponding points in a second image [2]. In analogy to the geometric case, at least 4 non-coplanar rays are required to solve for a color homography.

An RGB measurement without shading can be encoded as the (r, g) chromaticity coordinate: $r = R/(R+G+B)$ and $g = G/(R+G+B)$ (since the vector $[R \ G \ B]^{\mathsf{T}}$ has the same orientation as $[r \ g \ 1-r-g]^{\mathsf{T}}$. In Figure 1c, the 4 reflectances from the ball correspond to 4 points in an rg-chromaticity diagram and these define the quadrilaterals shown in the left and right of the panel (for respectively for the images Ball_2

and Ball_1). Assuming the illumination color change is linear, the mapping between the two chromaticity diagrams is precisely a homography (a fact we formally prove later).

We apply color homographies to help solve problems in three applications. In Figure 2a, we show the picture of an image in the RAW RGB space of a camera and the corresponding reproduction when the colors are corrected for display (where both images are also tone mapped for printing in this paper). In computer vision, the idea that pervades color correction is that all we need to do is find the best least-squares transform mapping the color checker shown in the RAW image to pre-measured correct display RGBs with some of the problem focus directed towards automatically finding the checker in the image [11, 12]. However, consistent with other recent work [13, 14], we found that the illumination intensity could vary significantly over the image of a checker. When shading varies, the color correction transform is found by solving for the homography relating the colors from the RAW to reference display RGBs.

Consider that we have a database of colorful objects where the color content of each image is represented by its chromaticity distribution. By matching color distributions we can obtain surprisingly accurate color object recognition [15, 16]. However, object recognition performance degrades when the light color changes [17]. The role of homographies in color object recognition is summarized in Figure 2b. Clearly, image I_1 matches image I_2 (it is of the same object) but image I_3 is an image of a different object. The distribution of rg chromaticities for image I_1 is shown in purple in both chromaticity diagrams in the middle of Figure 2b. The rg distribution of images I_2 and I_3 are overlaid in green. Notice, however, that the distributions do not match in either case. Indeed, even though images I_1 and I_2 are of the same object their chromaticity distributions don't match because images I_2 and I_3 are taken with respect to a warmer illumination color.

Our hypothesis is that the chromaticity distributions for the same scene lit by two different lights but where the image shading might change will be related by a homography. In the last row of Figure 2b, we show the output of homography matching for chromaticity distributions for I_1 matched to I_2 and I_1 matched to I_3. We see the chromaticity distribution of H(I_1, I_2) (the colors in I_1 homographically transformed to match those in I_2) now overlap and we can conclude the two images plausibly have the same color content. In contrast
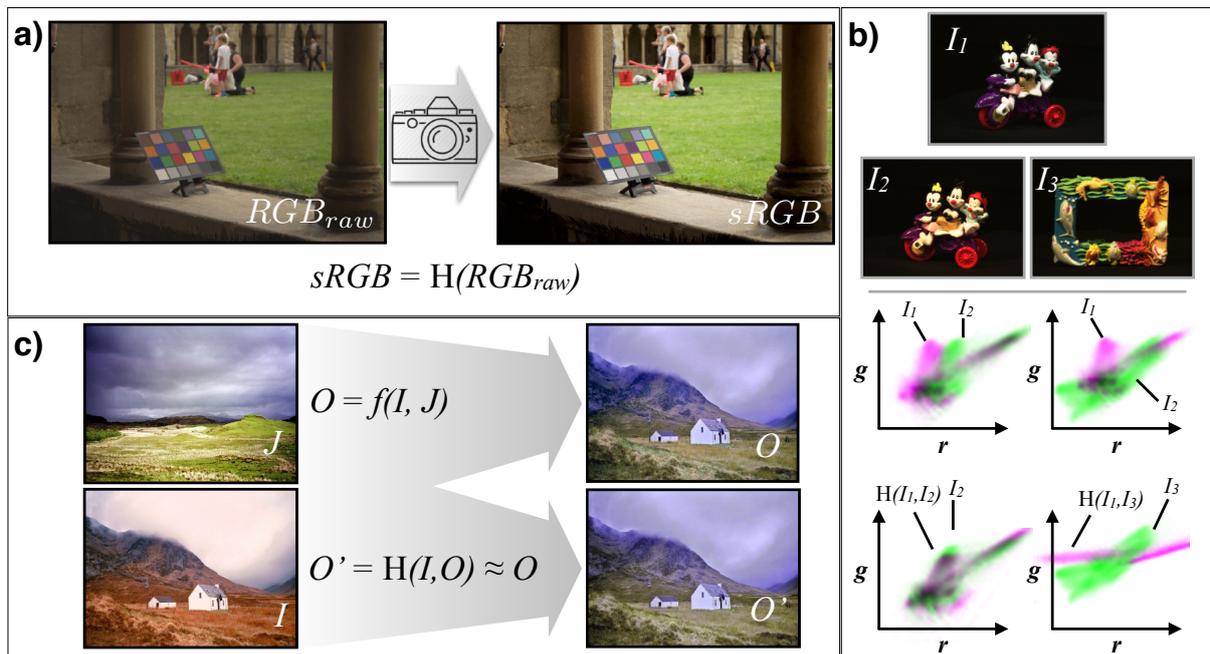
Figure 2: a) Color correction (mapping RAW to display sRGB [10]) is a homography problem. b) The top contains 3 images of colorful objects. Histogram matching of chromaticity distributions of images I_1 with I_2 and I_1 with I_3 are shown in the middle (chromaticity distribution for I_1 is in purple and for I_2 and I_3 is in green). Solving for the color homographies best mapping I_1 to I_2 and I_1 to I_3 (respectively, $H(I_1, I_2)$ and $H(I_1, I_3)$) results in the histograms shown in the bottom row. The distributions for I_1 and I_2 now match and the object is correctly identified. c) Color transfer for matching the colors of an input image I to the colors of a target image J can be reinterpreted as a simple color homography mapping $H(I, O)$.

$H(I_1, I_3)$ fails to match the chromaticity distributions. Note this latter match is poor (even poorer than the input) due to the mechanics of how the match is made which we discuss in Section 3.2. The matching method in this case actually informs us that no good match is possible, but for purposes of illustration, we show the *weak* match that was found.

Finally, we consider color homographies in the context of image color transfer. Automated color transfer is often required in professional photo editing. Artists can choose a desired target picture and manipulate another picture to match its color styles to the target. In Figure 2c, Image I has its colors transferred to map those shown in Image J and the result of the method [18] is denoted $O = f(I, J)$. Using our color homography theorem, we can approximate the color transfer result O by a color homography model which produces a close approximation image $O' = H(I, O)$ (I mapped by a homography to approximate O). By enforcing color space homography the resulting image can be physically interpreted as being an image of the same scene as the input image but under a different illumination. We can re-interpret most color transfer effects using a color homography model. This result is useful because, unlike many of the color transfer algorithms, color homographies can be computed quickly. Indeed we have found that we can run a computationally expensive algorithm on a thumbnail image, compute the homography and then apply the homography to the full-resolution image. Another benefit of our homography-based model is that we often remove the artifacts introduced by some color transfer algorithms. That is, we run a color transfer algorithm (whose output has artifacts) and approximate it as a homography. Because of the simpler transfer that is enforced by the homography many of the spatial artifacts either disappear or are mitigated.

Experiments demonstrate the power of our color space homography idea. Regarding color correction we report a significant improvement in color accuracy compared to the commonly used color correction methods. For object recognition on the Amsterdam Object Image Database [19] (> 100000 images encompassing a range of capture conditions), homographically matching chromaticity distributions supports state-of-the-art color-based object recognition. Finally, we found that existing color transfer algorithms can be re-interpreted as a color homography mapping.

In Section 2, we review the homography problem from geometric computer vision and relate this to linear image formation in color/photometric vision. Section 3 presents the color homography theorem together with a discussion on how to, in practice, solve for a color homography. Experiments in color correction, color object recognition and color transfer are reported in Section 4. We discuss application of homographies to non-RAW images in Section 5. The paper concludes in Section 6.

2   Background

2.1   Geometry

For the geometric planar homography problem, we write:

$$
\begin{bmatrix} \alpha x \\ \alpha y \\ \alpha \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \triangleq \underline{x} = H(\underline{x}')
$$
(1)

In Equation 1, (x, y) and (x', y') denote corresponding image points – the same physical feature – in two images. In homogeneous coordinates the vector $[a\ b\ c]^\intercal$ maps to the coordinates $[a/c\ b/c]^\intercal$ and so, in Equation 1, the scalar α cancels to form the image coordinate (x, y). For all pairs of corresponding points (x, y) and (x', y') that lie on the same plane in 3 dimensional space, Equation 1 exactly characterises the relationship between their images [2]. To solve for a homography (*e.g.* for image mosaicking), we need to find distinctive feature points in pair of images, match them to find candidate corresponding points, then solve for the *best* homography and finally warp the image to bring one image into the coordinate frame of the other [20, 21]. Homographies are at the heart of geometric methods in computer vision and are used in geometric camera calibration [22], 3D reconstruction [23], stereo vision [24] and image mosaicking [25] amongst other tasks.

2.2   Color

A physically accurate model of Lambertian image formation where the illumination impinging on a scene is a single color is written as

$$
\underline{\rho}^x = \alpha^x \int_\omega E(\lambda) S^x(\lambda) \underline{Q}(\lambda) d\lambda
$$
(2)

where respectively $E(\lambda)$, $S^x(\lambda)$ and $\underline{Q}(\lambda)$ denote the spectral power distribution of the light, the spectral reflectance of a surface, and the vector of R-, G- and B-

spectral sensitivities of the camera. And, the integral is taken over the visible spectrum ω. The superscript x denotes spatial dependency and α^x is a scaling factor encoding brightness changes due to the relative interplay between where the lights are positioned and the orientation of the surface in the scene (*e.g.* Lambert's law) and the quantity of light (*e.g.* to model extended light sources). It is well established [26–28] that surface reflectance can often be written to a tolerable approximation as the weighted sum of three basis functions:

$$S(\lambda) \approx \sum_{j=1}^{3} S_j(\lambda)\sigma_j \qquad (3)$$

When $\alpha^x = 1$ (so we can ignore spatial dependency) we have the so-called Mondrian-world [3]. There we have no shading in images and this simplifies consideration of how colors change with illumination and/or imaging device. Under the Mondrian-world assumption, we write:

$$\underline{\rho} = \Lambda^{E(\lambda),\underline{Q}(\lambda)}\underline{\sigma}$$

$$\Lambda_{ij}^{E(\lambda),\underline{Q}(\lambda)} = \int_\omega E(\lambda)S_j(\lambda)Q_i(\lambda)d\lambda \qquad (4)$$

In Equation 4, image formation is a 3×3 linear matrix Λ multiplying a 3×1 weight vector $\underline{\sigma}$. Two important results follow from Equation 4. First, that the 3×3 matrix

$$\Lambda^{E_2(\lambda),\underline{Q}(\lambda)}[\Lambda^{E_1(\lambda),\underline{Q}(\lambda)}]^{-1} \qquad (5)$$

maps colors viewed under illuminant $E_1(\lambda)$ to those recorded under $E_2(\lambda)$ assuming the same camera sensitivities $\underline{Q}(\lambda)$. Second that the 3×3 transform relating colors recorded by cameras with the respective sensitivities $\underline{Q}_1(\lambda)$ and $\underline{Q}_2(\lambda)$ is written as

$$\Lambda^{E(\lambda),\underline{Q}_2(\lambda)}[\Lambda^{E(\lambda),\underline{Q}_1(\lambda)}]^{-1}. \qquad (6)$$

assuming the same illuminant $E(\lambda)$.

Of course it is well known that Equation 3 is only approximate. Indeed, illuminant metamerism [29] – the phenomenon that two surfaces look the same under one light can look different under another – cannot be described by a 3×3 matrix of illuminant change. Yet, metamerism (illuminant or sensor) is rare. Moreover Marimont and Wandell [30] extended the linear model formalism to incorporate image formation into the derivation of the optimal linear basis and found that 3×3 matrices could well account for illuminant and sensor change. A similar result was reported by Funt

et al. [31, 32]. While most of the literature assumes a single global illuminant, other recent work are also proposed to solve multi-illuminant color constancy [33–35], which is an emerging and challenging area of research.

That illumination (or device) colors map across images using a linear transform is a common assumption and is widely reported in the literature including, in color correction [5–7] color object recognition [9, 36] and illuminant estimation [4]. However, the Mondrian-world does not generally hold. Indeed, color intensity can and does vary on a per-pixel basis due to the relative position of light and surface and also due to the quantity of light varying across a scene. When illumination color is held fixed then a chromaticity representation of color can be used to factor out per pixel shading variation *e.g.* [37].

To deal with a changing light color, the so-called "diagonal" model of image formation is often employed with respect to which color change across images is modeled as a diagonal matrix [38, 39] multiplying the image colors. The diagonal model is at the heart of the "comprehensive color image normalization" [40] and the $m1m2m3$ coordinate system [41] both of which are light color plus shading invariant image features. Further, specially chosen spectral band ratios (for example, R/G and B/G) have an analogous diagonal property (the 2-d band ratios are mapped across illumination by a 2×2 diagonal matrix). This property is exploited in illuminant color and shading-independent histogram matching for object recognition [42] and illuminant estimation [43, 44].

Of course moving to a chromaticity representation means one of the 3 degrees of freedom measured in an RGB image has been lost. In calibrated color correction – mapping the RAW RGBs recorded for a known color chart to a standard color space – it is possible to find the full 3×3 matrix mapping the orientation of input color vectors to the orientation of output colors by a searching strategy [13]. Alternately, in [14], solving for the best shading while simultaneously finding the color correction matrix was formulated as an Alternating Least-squares approach. Both these methods deliver significantly lower correction error compared with a shading blind linear least-squares optimization. In display calibration [45] it was shown that 4 chromaticities sufficed for color calibration.

5

3 Color Homography

Let us map an RGB $\underline{\rho}^{\mathsf{T}}$ to a corresponding RGI (red-green-intensity) $\underline{c}^{\mathsf{T}}$ using a 3×3 matrix C:

$$\underline{\rho}^{\mathsf{T}} C = \underline{c}^{\mathsf{T}}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R \\ G \\ R+G+B \end{bmatrix}^{\mathsf{T}} \quad (7)$$

interpreting the right-hand-side of Equation 7 as a homogeneous coordinate (e.g here, as chromaticities) we see that

$$\underline{c} \propto \begin{bmatrix} r & g & 1 \end{bmatrix}^{\mathsf{T}}, \ r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B} \quad (8)$$

In the following proof it is useful to represent 2-d chromaticities by their corresponding 3-d homogeneous coordinates.

Theorem 1 (Color Homography). *Chromaticities across a linear change in capture condition (light color, shading and imaging device) are a homography apart.*

Proof. First we assume that across a change in illumination or a change in device where the shading is the same (e.g. for the Mondrian-world) the corresponding RGBs are related by a linear transform M (*i.e.* Equations 5 and 6 hold). Clearly, $H = C^{-1}MC$ maps colors in RGI form between illuminants. Due to different shading, the RGI triple under a second light is represented as $\underline{c}'^{\mathsf{T}} = \alpha \underline{c}^{\mathsf{T}} H$, where α denotes the unknown scaling and T denotes transpose. Without loss of generality let us interpret $\underline{c}$ as a homogeneous coordinate *i.e.* assume its third component is 1. Then, chromaticity coordinates $\underline{c}^{\mathsf{T}}$ and $\underline{c}'^{\mathsf{T}}$ are a homography apart. $\square$

In geometry, homographies are applied for mapping spatial coordinates in one image to correspondences in another. In color homography we are also interested in this 2D-2D match problem (chromaticity to chromaticity mapping). However, for some applications – *e.g.* color transfer – the homography that maps 3D colors to 3D color matches (in a shading independent way) is the apposite tool.

3.1 Color Homography Estimation by Alternating Least Squares

Suppose A and B denote respectively $n \times 3$ matrices of n corresponding pixel RGBs with respect to two images of the same scene where the illumination changes

(and, also possibly the camera properties). The color change is modeled as a linear transform (Equations 5 and 6) but because of the relative positions of light and surfaces there might also be per-pixel shading perturbations. Assuming that the Lambertian image formation is an accurate physical model, the relationship between A and B is written as

$$DAH \approx B \quad (9)$$

where D is an $n \times n$ diagonal matrix of scalar shading factors and H is a 3×3 homography color correction matrix. If both D and H are applied to A, then we call this a *shading homography*. We can solve Equation 9 by using Alternating Least-Squares (ALS) [14, 46, 47] described in Algorithm 1. There $\|.\|_F$ denotes the Frobenius norm. And, at iteration i, H^i and D^i are found using the closed form Moore-Penrose inverse. The least squares fit of A to B is equal to $(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}B$. Each scalar component at the jth diagonal element of D^i is the result of a least squares fit of the jth row of A^{i-1} to the jth row of B (where again we use Moore-Penrose inverse). The effect of the individual H^i and

---

Algorithm 1: Homography from alternating least-squares

---
1 $i = 0, \min_{D^0} \left\| D^0 A - B \right\|_F, A^0 = D^0 A$;

2 **repeat**

3     $i = i + 1$;

4     $\min_{H^i} \left\| A^{i-1} H^i - B \right\|_F$;

5     $\min_{D^i} \left\| D^i A^{i-1} H^i - B \right\|_F$;

6     $A^i = D^i A^{i-1} H^i$;

7 **until** $\left\| A^i - A^{i-1} \right\|_F < \epsilon$;

D^i can be combined into a single matrix $D = \prod_i D^i$ and $H = \prod_i H^i$ (where the product is taken by post-multiplying matrices). That is, $DAH \approx B$. ALS converges to a local minimum [48].

Finally, note how we initialize the matrix A^0 to be the closest least-squares fit of the rows of A to the rows of B. This initialization is performed for two reasons. First, we find slightly better visual results if we applied D before H. Secondly, it simplifies the proof of Theorem 2.

To motivate the theorem below we know from geometric computer vision that given the images of the two planar regions each enclosed by 4 corner points in two views (assuming certain assumptions) there is a unique homography and that this can be found directly [2]. In

6

Theorem [2], we capture the conditions where the ALS method can find the unique homography given 4 pairs of colors captured across viewing conditions.

Theorem 2 (ALS Uniqueness). *The ALS method finds the homography – unique, up to a scalar multiplier – when there are 4 pairs of corresponding non-collinear colors A and B each of which is a 4×3 full-rank matrix (i.e. $rank(A) = rank(B) = 3$) and the null space – which is represented by the 4-vector $\underline{v}^i$ – of $(A^i - B)$ at convergence is not sparse. A vector $\underline{v}^i$ is sparse if 1, 2 or 3 (but not all 4) of its entries are 0.*

Proof. In this case, A and B are 4×3 matrices. Because the elements of A and B are strictly positive and full rank, the 3×3 cross product matrices $A^\mathsf{T}A$ and $A^\mathsf{T}B$ are also full rank. We point this out because this in turn implies that at every time we calculate D^i and H^i in the ALS method that these matrices themselves are full rank. That is, so long as A and B begin as full rank matrices the matrices A^i, B, D^i and H^i are all themselves full rank *i.e.* we never encounter the circumstance where an individual least-squares fit introduces rank deficiency.

Now let us suppose that, the stopping condition of the ALS procedure is met in the ith iteration. To prove Theorem [2], we need to first prove the following lemma.

Lemma 1. *When the ALS algorithm converges at step i then $D^i = \mathcal{I}_{4\times4}$ and $H^i = \mathcal{I}_{3\times3}$ where $\mathcal{I}$ denotes the identity matrix.*

Proof. Let us assume Lemma [1] is false then when the algorithm converges it must follow we can write $DA^{i-1}H = A^i$ (where here D and H are respectively a 4×4 diagonal matrix and an arbitrary 3×3 linear transform). This can only be true if the rows of A^{i-1} are row Eigenvectors of H. This cannot be the case because then two of the rows of A^{i-1} would be the same (up to a scalar) *i.e.* collinear and in the statement of the theorem this was assumed not to be the case (*i.e.* to get to A^i involves multiplying A by a series of full rank matrices. If the rows of A are not collinear then neither are the rows of A^i). □

Let us now prove Theorem [2] by contradiction. We will show that if the alternating least-squares procedure has converged and the unique homography has not been found then the assumption of convergence cannot hold (and, yet, we know the algorithm converges [48]).

Let us assume that $A^i \neq B$. It follows that

$$B = A^i + \underline{v}\,\underline{w}^\mathsf{T} \quad \text{s.t.} \quad \underline{v}^\mathsf{T}A^i = \mathbf{0}$$

where respectively $\underline{v}$ and $\underline{w}$ are 4×1 and 3×1 vectors, and $\underline{v}$ spans the column null-space of A^i. From Lemma [1], on convergence, both D^i and H^i are identity matrices. Let us consider the final step in the algorithm (step 5 just before the algorithm stops). By Assumption the algorithm converges and $A^i \neq B$ and $H^i = \mathcal{I}_{3\times3}$ (from Lemma [1]):

$$A^i + \underline{v}\,\underline{w}^\mathsf{T} = B = D^i A^i H^i = D^i A^i \mathcal{I}.$$

The least-squares per-row scalar returned in Line 5 of Algorithm [1] can be written in closed form (we are simply using the Moore-Penrose inverse)

$$D^i_{jj} = \frac{\underline{a}_j^\mathsf{T}(\underline{a}_j + v_j\underline{w})}{\underline{a}_j^\mathsf{T}\underline{a}_j} = \frac{\left\|\underline{a}_j\right\|^2 + v_j\underline{a}_j^\mathsf{T}\underline{w}}{\left\|\underline{a}_j\right\|^2}$$

where the jth row of A^i, denoted $\underline{a}_j^\mathsf{T}$, has a magnitude $\left\|\underline{a}_j\right\| > 0$, because A^i has full rank and no two points are collinear. It can be shown that the scaling of the jth row of A^i that best matches the jth row of B in a least-squares sense can be written as:

$$D^i_{jj} = 1 + \frac{v_j}{\left\|\underline{a}_j\right\|^2}\underline{a}_j^\mathsf{T}\underline{w}$$

which implies $D^i_{jj} = 1 \iff \underline{a}_j\underline{w} = \mathbf{0}$. Because A^i is assumed to be full rank, it is possible that three of its four rows can satisfy $\underline{a}_j^\mathsf{T}\,\underline{w} = 0$. However, this can not hold for the fourth row otherwise A is not full rank. And by our assumption of *non-sparsity* all elements of $\underline{v}$ are either non-zero or zero. If they are all non-zero *i.e.* the algorithm has not converged (when by assumption it has). We have a contradiction that $D^i \neq \mathcal{I}$. □

Of course the reader will have noticed a sleight of hand on our part. We made a *non-sparsity* assumption on the null-space of $A^i - B$ at convergence. Albeit rarely, we found the alternating least-squares procedure might terminate with a sparse null space vector and in this case with $A^i \neq B$ (the algorithm stops but we do not find the unique homography).

To investigate convergence empirically, we uniformly randomly generated 4×3 matrices A (matrix elements in the interval $[0, 1]$ to 2 decimal places) and then randomly generated matrices D (elements chosen uniformly and randomly from $[0, 1]$) and H (elements

chosen from the standard normal distribution $N(0, 1)$). Then, we compute $B = DAH$. The matrix B created in this way could plausibly be matching colors under different lights or different devices. We then ran the ALS procedure to solve for the homography, to discover D and H.

Over 10,000 runs the ALS procedure converged to the correct answer (the non-sparsity condition was satisfied) over 96.5% of the time. For the remaining 3.5% of the cases that did not converge we mapped A and B to new matrices A' and B' according to:

$$A' = D^A A H^A \ , \ B' = D^B A H^B$$

where D^A and D^B are randomly chosen positive matrices (entries in $[0, 1]$) and H^A and H^B are 3×3 matrices with elements drawn from $N(0, 1)$. We then run our procedure a second time. Assuming convergence to the correct answer *i.e.* we find matrices D' and H' such that $D'A'H' = B'$. It follows that $D = [D^B]^{-1} D' D^A$ and $H = H^A H' [H^B]^{-1}$. In all cases, this kind of random perturbation sufficed to make ALS converge to the correct answer. In all cases (even without the perturbation), the % difference $\|DAH - B\| / \|B\|$ between the actual and fitted homography, between DAH and B was less than 0.5% (and usually orders of magnitude smaller). Indeed, we have not found a compelling visual example where the output of the ALS procedure – when it converges to the wrong answer – appears different from the overall best answer (zero error for the 4 point homography fitting). In Appendix A we present a numerical example where the 4 point ALS minimization fails to solve for the homography.

Finally, in thinking about solving for a homography by mapping the original matrices A and B to counterparts (which, for ALS then converges correctly) we were, as a side-effect, able to write the homography calculation directly as a simple matrix computation (see Appendix B). While wholly equivalent to the "direct method" [2], the form of the equation appears novel.

3.2 Homography-based Chromaticity Matching

In geometric computer vision we find the homography relating two images in three steps. First, we find distinctive features, second we find candidate matching locations by pairing distinctive features which also have similar underlying image structure and third, we find the best homography match accounting for as many of our paired image points as possible. See [20] for a general discussion for finding corresponding feature points between two images. Here, we treat the chromaticity distributions *as images* and seek to find and match these distributions analogously to the geometric case (*i.e.* we find and match interest points found for chromaticity distributions). An interesting technical issue is that we would, ideally, like to be able to find the same feature points robust against the image transformation we are trying to discover. The ASIFT [21] algorithm is a methodology for matching images that is fully invariant to affine image modification. In the top of Figure 3, we show the ASIFT match for the color distributions of images I_1 and I_2 from Figure 2. Visually, we see that structurally similar (corresponding to the shape of the chromaticity distribution) points are found in both images. Pairs of correspondences found by RANSAC [2] matching are also shown (as the lines joining the ASIFT points in the two distributions). Among other popular feature point matching algorithms, such as SIFT [20], MSER [49], SURF [50], and Harris corner feature [51], we found ASIFT delivers the best chromaticity matching result.

The result of solving for the homography relating matched points is shown in the chromaticity diagram shown bottom left of Figure 2b. There the green distribution shows the chromaticities of image I_2 and in purple the homographically transformed chromaticities from image I_1. The distributions, and so the images, match. The result of applying the best homography relating the distributions for images I_1 and I_3 (based on the matched points shown in the bottom of Figure 3), does not match the chromaticity distributions (see bottom right, Figure 2b). We can conclude that the object in image I_1 does not match the object in image I_3.

While there are many ways of matching color distributions we found that a simple structural match score – a measure of how well a homography matches ASIFT points – provided a powerful way to determine whether one chromaticity distribution matched another. Suppose ASIFT returned N and M points of interests from a pair of chromaticity distributions and that (via RANSAC) we found that we could match m pairs of points by finding the best homography. Then, our structural match is defined as:

$$\text{structural\_match} = \frac{m}{\sqrt{MN}} \quad (10)$$

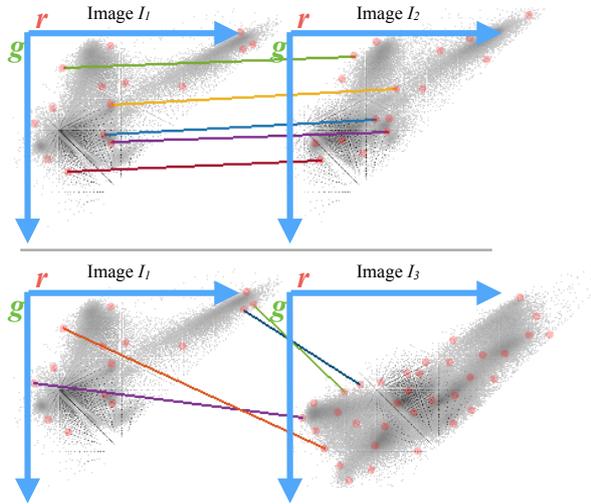A structural match score of 1 means all ASIFT points found in both chromaticity distributions are, placed

Figure 3: Top: ASIFT [21] can match the chromaticity distributions of images I_1 and I_2. Bottom: ASIFT cannot match the chromaticity distributions of image I_1 to image I_3 (images I_1, I_2 and I_3 shown in Figure 2b)

in correspondence, a homography apart. Section 4.2 shows an application of object recognition which adopts this structural match score as a measure of image similarity. Later we will see that the number of ASIFT points extracted from a chromaticity distribution – which we call its *structural complexity* – can be used to predict whether a homography-based distribution match will, in fact, be possible.

It is known that ASIFT is about 13.5 times the cost of SIFT [21]. Importantly, efficient implementations of SIFT exist including in real-time on mobile devices. Also, in general SIFT operates on large (> 1 megapixel) images. In contrast, though we use the more expensive ASIFT to match thumbnail images: our "images" are small: chromaticity distribution that have just 320×320 bins. Thus, despite the $13.5 \times$ cost multiplier compared to SIFT, the cost of ASIFT in matching chromaticity distributions is similar to running SIFT on full size images.

4 Experiments

4.1 Color Correction

4.1.1 Color Correction using a Color Chart

In color correction – mapping RAW RGBs to a display color space – the target RGBs are known to vary in intensity. Indeed, serious photographers will take a picture of a color target (such as the Macbeth color checker [52] shown in Figure 2a) and a second picture of a uniform gray target with same size in the same location. By dividing the RGB image of the checker by the image of the gray-target the shading is removed and then the shading corrected RGBs can be mapped to known reference display color coordinates using simple least-squares. However, in computer vision, and even for the vast majority of photographers, this two-step approach to calibration is rarely taken (*e.g.* see [11, 12]). If nothing else it is invasive and in some cases cannot be done at all (*e.g.* in an on-going surveillance situation). However, the photographic "best practice" allows us to measure useful data to evaluate homography vs. linear least-squares color correction. We can find the best least-squares 3×3 matrix mapping the non-shading corrected checker to a reference target and then apply this matrix to the shading corrected target. Or, we transform the target using a homography. This experimental methodology is described in detail in [13].

Because the display RGBs are measured in coordinates relevant to human vision, *e.g.* sRGB [10], the color error can be converted to the CIE Lab, CIE Luv, RGB error metrics. The CIE formulas [29] return a ΔE error of 1 if two patches are just discernible from each other. The ΔE RGB error between two RGB vectors $\underline{p}$ and $\underline{q}$ is calculated as $||\underline{p} - \underline{q}||$. Similarly, ΔE in Table 1 is also for the Luv and Lab representations. In Figure 2a, we show the RAW and rendered (JPEG) images of one of our images. Two examples containing outdoor and indoor lighting conditions for our color correction evaluation are shown in Figure 4. In total, we captured 13 images at Norwich cathedral and around the University of East Anglia campus. Our image set contains both indoor and outdoor illumination. To calculate the best homography we randomly chose 4 matching colors (according to the known correspondences) and using RANSAC, chose the homography that minimized correction error (ΔE Luv). The mean, median, 95% quantile and max ΔE errors calculated over the 13 images are reported in Table 1 where we compare per-

Figure 4: Two example images with non-uniformly shading used for color correction test.
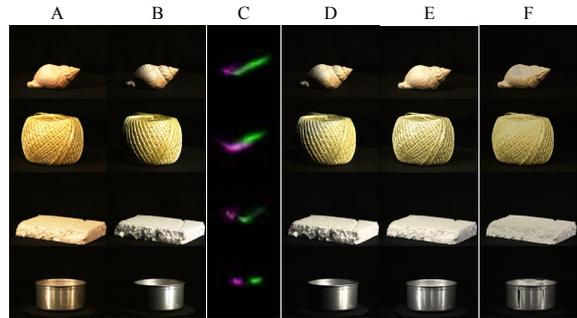

Figure 5: Given pixel-wise correspondence between Images A and B, a color homography is solved for to convert Image A to Images D (with shading correction) and E (without shading). Image F is the worse conversion result when linear least-squares is adopted. Image C shows the rg chromaticity spread (Image A in green, Image B in pink).

formance to simple least-squares, root-polynomial [53], and ALS [14] (which is shown to be improved on [13]). Compared with the ALS color correction, all mean errors are improved by about 9%, and all median errors are improved by about 21%, at the cost of getting a slightly higher 95% quantile error and maximum error (except for RGB metric). Compared to the least squares and root-polynomial methods – which are shading dependent – the advantage of using the homography based formalism is quite significant (*e.g.* the median errors – for all 3 error metrics – are halved).

4.1.2 Color Correction for Single Color Objects

The solution for a color homography requires at least 4 non-collinear chromaticities. For a perfect convex-shaped monochrome Lambertian surface viewed under a single light source, the resulting chromaticity distribution – in the idealized case – will comprise a single point. However, most objects are not perfectly Lambertian – there is a highlight component – and are not convex and so there could be inter-reflections. Both specular highlights [54, 55] and inter-reflections [56–58] lead to the chromaticity distribution of a single object being more than a single point. Indeed, in the presence of specular highlights and inter-reflections we expect a spread of points in 3D RGB space and a spread (not all collinear) points in chromaticity space.

In columns A and B of Figure 5 we show, respectively, 4 monochrome objects viewed in 2 viewing conditions (different colored light in different positions). In column C we show, in green, the chromaticity distributions of images in column A and in pink the chromaticities of the images in B. Clearly, there is a spread in chromaticities in both viewing conditions. If we are to match chromaticity distributions by solving for a homography then it is a necessary condition that the images (and their chromaticities) in columns A and B are related by a homography. Using the ALS method (the

pixels are in correspondence) we solve for the best 3×3 matrix and shading correction that relates the images in column A to counterparts in column B. Mapping the images in A with the solved shading and color corrections results in the fitted images, shown in column D. Visually, the fit is excellent. In column E we apply just the 3×3 homography matrix (no shading correction) to the inputs from A. Now we have the object mapped to viewing condition B but the original shading preserved. Finally, in column F we find a pixel-wise least-squares fit. Here the wrong 3×3 matrix is found (since it is attempting to best compensate shading and colour correction in a single matrix).

4.2 Color Object Recognition

The Amsterdam Object Image Database [19] is a large 1000 object database widely used to benchmark color-based object recognition. Importantly, each object in the database has its image captured with respect to 72 rotational views (range from $0°$ to $355°$, at $5°$ resolution) and 24 illumination angles, and 12 illuminant color temperatures (2175K to 3075K). In total the ALOI dataset comprises in excess of 100,000 images. We match objects by matching their underlying distributions. Simply, in color-distribution-based object recognition, if two images have similar underlying color distributions then this is taken as evidence of a possible object match.

Table 1: Errors for Color Correction (X-Rite Classic Color Checker)

Method	Mean	Median	95%	Max
ΔE (Lab)				
Least-Squares	6.16	5.67	12.27	13.83
Root-Polynomial	5.67	4.67	14.60	16.97
ALS	3.71	3.27	**8.24**	**9.02**
Homography	**3.40**	**2.59**	9.20	10.28
ΔE (Luv)				
Least-Squares	7.02	6.63	14.23	15.55
Root-Polynomial	6.69	5.46	16.88	19.19
ALS	4.17	3.70	**8.93**	**9.89**
Homography	**3.88**	**2.97**	9.97	10.76
ΔE RGB ($\times 10^{-2}$)				
Least-Squares	7.58	6.34	19.07	30.11
Root-Polynomial	9.07	7.21	24.72	30.60
ALS	4.36	3.16	**14.01**	26.05
Homography	**4.01**	**2.82**	14.62	**24.44**

We wished to evaluate how well homography-based chromaticity matching would support color-based object recognition. In our approach we use the ASIFT procedure to find distinctive points in the chromaticity distributions of all the images in the database. Then, for each testing condition (viewing angle, illumination angle or illumination color), we use a standard reference dataset and use the remaining images as query images. The standard reference capture condition is: frontal view (*i.e.* 0 degree of rotation), frontally lit and the light has a 3075K yellow color.

We compare the homography-based color object recognition (H) with Swain's color indexing algorithm which uses an rg-chromaticity space [15] (rg), comprehensive color image normalization [40] (CN) and Gevers and Smeulders' [41] $m1m2m3$ approach. The Comprehensive Normalization and m1m2m3 techniques are chosen both because they explicitly incorporate shading invariance into their formalisms and also they supersede early work (*e.g.* Comprehensive Normalization delivers better results than aligning chromaticity distributions by matching their means with a diagonal matrix [38, 39]).

The recommended bin sizes for rg and CN are 16×16. The recommended bin size for m1m2m3 is 32×32 (note the distribution of these bins is non-linear with the partition found by a calibration procedure [41]). For our ASIFT-based chromaticity distribution match,

a 320×320 histogram is used. For all methods we assess recognition performance using the simple Average Match Percentile (AMP). An average match percentile of 99% informs us that the correct matching answer is in the top 1% of matches.

We run the recognition experiment given all 1000 objects where, relative to the standard viewing condition, either the viewing angle, illumination angle or illumination color are individually varied. As shown in Table 2, the pure homography-based chromaticity matching supports the second best performance for viewing angle and illumination angle tests. Of course, as alluded to in the color correction section, to match chromaticity histograms using homographies the underlying distributions need to have sufficient 2-dimensional structure. Visually, we found that $\sim 80\%$ objects in the ALOI dataset are comprised of objects with 3 or fewer reflectance colors and we know we need 4 points to solve for a homography. This said, the overall performance of the homography-based method is good *i.e.* it performs well for many objects with fewer than 3 colors present. Potentially, a good match remains possible because the effect of specular highlights and inter-reflection increase the dimensionality of the RGB signal recorded for a single object (see discussion at the end Section IV-A-2).

Intrinsic to our ASIFT based matching is the notion of color *structural complexity* which is simply defined

11

Table 2: Average Match Percentile Results for all 1000 objects

Test	rg	m1m2m3	CN	H	Hybrid H
Viewing Angle	**98.8**	95.3	96.8	97.3	97.0
Illumination Angle	92.2	88.3	94.8	93.8	**95.1**
Illumination Color	87.3	98.2	98.0	97.2	**98.2**
Average	92.8	93.9	96.5	96.2	**96.8**

Table 3: Average Match Percentile Results for Query Image with Color Structure Complexity. Numbers in bracket indicate percentage of color-homography-compatible queries.

Test	rg	m1m2m3	CN	H
Viewing Angle (12%)	**98.6**	94.8	96.3	97.4
Illumination Angle (19.4%)	91.5	90.1	94.8	**96.3**
Illumination Color (24.4%)	92.1	99.1	99.3	**99.5**
Average	93.4	94.3	96.4	**97.7**
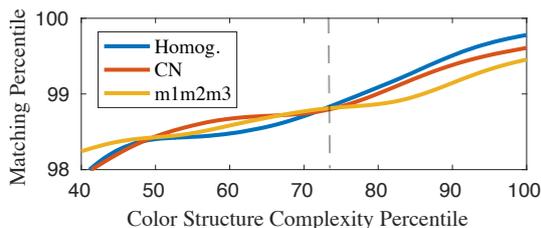
Figure 6: Plot of Color Structure Complexity Percentile and Matching Percentile. The result is filtered by a moving Gaussian convolution kernel (width = 10% of total color structure complexity span, σ = 1/6 kernel width). The dashed line indicates the threshold, over which homography-based method works better.

as the number ASIFT points found for a given chromaticity distribution. We hypothesize that, empirically, we can find a threshold of the structural complexity to determine whether color homography indexing is likely to work (*e.g.* there is sufficient structure in the chromaticity distribution to find matches using homographies in a database).

To test this hypothesis, for the "Illumination Color" test, in Figure 6, we plot the percentiles of the structural complexity (*i.e.* we rank images according to the number of ASIFT points found in their chromaticity distributions) and plot against the corresponding match percentiles. Clearly, there is a good correlation between the two measures and this indicates that as structural com-

plexity increases so the homography-based measure delivers better object recognition. We find a similar positive correlation for the m1m2m3 and comprehensive normalization approaches.

In this example, assuming we wish to find the correct match in the top 1% of a database we need to use a threshold structural complexity as defined by the 75th percentile image. For our test the 75th percentile image and above have a structural complexity of at least 520 points. Importantly, notice, that in this top quantile range, the homography match is better than either the CN and m1m2m3 methods. Note that a threshold of 75% for the color structural complexity means that only 25% of the query images, can be homographically matched.

Suppose, we seek to match only the chromaticity distributions with 520 or more ASIFT points. We report the results in Table 3. The percentage shown are the number of chromaticity distributions in the database that have more than 520 ASIFT points. Note this is less than 25% (compared to our test in Figure 6). Because under some illumination and viewing angles, there are fewer object colors so fewer ASIFT points. For images whose chromaticity distributions have sufficient structural complexity, the homography method delivers better results compared with competing techniques. In the last column of Table 2, we show the performance of a Hybrid method (Hybrid H): if the chromaticity distribution has more than 520 ASIFT points then homography-based matching is used. Otherwise, we adopt the CN

12

method. The hybrid method delivers the best results for the illuminant colour and illumination angle test conditions and the best result overall.

4.3 Color Transfer

Color transfer is an image editing process that adjusts the colors of an input image I to match the palette of a target image J. Instead of adopting often computationally costly non-linear color mappings (*e.g.* [59]), we investigate whether color transfer can also be interpreted as a simple linear color homography mapping which revisualizes an image with respect to real physical scene changes (*e.g.* from summer to autumn) and/or illumination. Our experiment is a continuation of our previous work [60] which is built on recent research [61]. The recent work [61] has demonstrated that it is possible to find a simple 3D similarity transform to linearly approximate some of the effects of color transfer.

There are three important reasons for visualizing color transfer as possibly being a color homography. First, if true, it would indicate a surprising result. Specifically, color transfer, though defined quite generally, tends to generate images which have a real-world physical interpretation. Second, the homography formulation is simpler compared with some of the transfer algorithms. This simplicity means that spatial artifacts are less likely to be introduced (a fact borne out by our experiments). Lastly, the homography calculation is fast. As we will show below, we can calculate an expensive color transfer on a thumbnail, calculate the homography and then apply the result to a full resolution image.

We start with the color transfer output Image O and try to approximate its color transfer result by color homography. An example is shown in Figure 2c where Image O' is the color transfer approximation result. In the discussion that follows it is useful to think of the image I (or O) being a simple $n \times 3$ matrix of RGB pixels, which can be reconstituted into an image grid for display purposes. Mathematically, we write:

$$O' = \mathrm{H}(I, O) \approx O \qquad (11)$$

To solve Equation 11 using our ALS algorithm, we respectively convert the images I and O images to the corresponding $n \times 3$ matrices A and B. Further, the ALS computed output DAH mapped back to an image is denoted O'. Here, the homography matrix H can be understood as a global chromaticity shift and distortion,

the matrix D can be interpreted as shading change factors which simulate the change of surface reflectance or position of illuminant.

The visual results of color transfer approximations of four color transfer methods [18, 62–64] are shown in Figure 7. In our experiments, the number of ALS iterations is set as 10. As can be seen, the global 3D Similarity mapping [61] does not perfectly reproduce the shading adjustments of color transfer. Our homography-based method offers a closer color transfer approximation.

We can also quantify this visual closeness by calculating the error between the color transfer approximation result and the original color transfer result. We adopt three error metrics:

PSNR (Peak Signal-to-Noise Ratio). PSNR is the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. Acceptable values for wireless image transmission quality loss are considered to be over 20 dB [65].

SSIM (Structural SIMilarity) [66]. SSIM is a model that considers image degradation as perceived change in structural information. SSIM can be used to assess the artifacts of color transfer. A SSIM value "1" indicates a perfect match.

HI (Histogram Intersection) of rgb Chromaticities. This is a score for measuring the similarity between two rgb chromaticity distributions, which is also shading-independent. A similarity score "1" indicates a perfect match.

We show the average color transfer approximation result in Table 4 (see the supplementary material for the complete table and their visual results). The quantitative test is based on 7 classic color transfer image pairs and 4 color transfer methods [18, 62–64]. Our color homography transfer produces the best results overall for the PSNR and SSIM tests. For the chromaticity mapping test (HI), our color homography approximation performs better for Pouli and Reinhard [63, 64].

We note that both the PSNR and SSIM metrics operate on the images output from our method. But, the histogram intersections are on the chromaticities of the original and matched images. The chromaticity mapping takes all RGBs and scales them so they sum to 1. This has a dramatic effect especially for dark pixel values. Yet, these values are precisely those that are de-weighted in the ALS method (which operates by least-squares). They are also the pixels in the image that we cannot see.

13

| Source Image | Original Color Transfer | 3D Similarity [61] | Shading Homography |

Figure 7: Visual result of color transfer approximations (in the order of [18], [62], [63], [64]). The images in Column 4 (Homography) are generally more similar to those in Column 2 than those shown in Column 3 (3D Similarity).

Table 4: Errors of color transfer approximation.

	Ngu. [18]	Pitie [62]	Pouli [63]	Rein. [64]
PSNR (Peak Signal-to-Noise Ratio)				
3D Similarity [61]	26.85	26.04	26.92	28.49
Homography	**34.05**	**31.91**	**37.28**	**36.31**
SSIM (Structural SIMilarity)				
3D Similarity [61]	0.91	0.85	0.84	0.88
Homography	**0.94**	**0.91**	**0.97**	**0.98**
HI (Histogram Intersection)				
3D Similarity [61]	**0.87**	**0.75**	0.75	0.81
Homography	0.78	0.71	**0.76**	**0.87**

Original Approximation | 50% Downsampling | 2⁻⁴ Downsampling

PSNR: 32.68; SSIM: 0.96; HI: 0.88 | PSNR: 32.49; SSIM: 0.96; HI: 0.88 | PSNR: 32.43; SSIM: 0.96; HI: 0.88

Total Time: 1,824s | Total Time: 1,172s

Figure 8: Color transfer approximation for [18] from downsampled images. The sizes of source images and target images (*i.e.* I and J) are reduced by the corresponding factors. The original color transfer (MAT-LAB) takes about 3.630s. The 3 evaluation measurements and the total estimation time with down-sampling (MATLAB) are shown over the images. As it is shown, image down-sampling barely affects the color transfer approximation quality. And, it takes less time to color transfer an image by using our down-sampling trick.

To reduce the computational cost, it is also possible to estimate the shading homography (*i.e.* H and D) from the down-sampled images. In addition to the original ALS, we also upsample the smaller shading matrix returned by ALS (from the downsampled input images) by using Joint Bilateral Upsampling [67] (guided by the chromaticity-transferred result AH). We find that image down-sampling barely affects the color homography mapping quality. In Figure 8, we show an example corresponding to the result of Row 1, Figure 7. This can be useful because we can run a computationally costly color transfer algorithm on thumbnail images, extract the color transfer effect from the downsampled results, and apply the extracted effect to a full-resolution source image. For some computationally costly color transfer algorithms (*e.g.* [18, 62, 63]), this offers a similar quality color transfer result but requires a reduced amount of processing time.

Color homography is also useful for color transfer enhancement. Figure 9 shows an example where the original color transfer result contains some obvious artifacts. These artifacts are usually caused by sharp image gradient changes. Since a color homography transform barely modifies the original image gradient, the decomposed shading component absorbs most of the artifacts (*e.g.* original shading in Figure 9). By spatially smoothing the original noisy shading component with a bilateral filter [68], we can remove the gradient artifacts in the shading component (*e.g.* smoothed shading in Figure 9). The original color transfer result is improved by

Source Image | Original Color Transfer | Original Approximation

Direct RGB Smoothing | Shading Smoothing

Figure 9: Color transfer enhancement. A source image is color transferred by [63] with some noticeable imperfections (on the clouds). This issue is fixed by approximating the original color transfer effect with a *shading-smoothed* shading homography. Visually, the enhanced result also preserves more texture details compared with the result obtained by directly smoothing the original approximation RGB.

applying the modified shading homography (*i.e.* original color homography + smoothed shading).

When we proved the *Uniqueness* theorem – that the ALS procedure could be used to find the homography for the 4 point case – we introduced the idea of perturbing the data to find the unique homography. The data was perturbed when, albeit very rarely, the ALS procedure did not converge to the unique homography (see discussion after the proof of Theorem 2 in Section 3). For the color transfer problem we explored perturbing the input image by a random homography to investigate whether, overall, a better color transfer (a better fit) is found. We did not find that this was the case. This result is perhaps not surprising since for the 4 point case when the ALS procedure converged to the wrong answer the fitting error was, empirically, very good (typically less than 1 percent).

5 Raw vs Rendered Images

That a color homography is a tool for describing color change across viewing conditions appears, at first glance, to hold only for RAW linear images. Empirically, our work on color transfer also indicates we can apply the homography method to non-linear (rendered) images. An example of RAW-to-rendered mapping approximation by our shading homography is shown in Figure 10 where the approximation result is visually close to the rendered – actual camera output – image
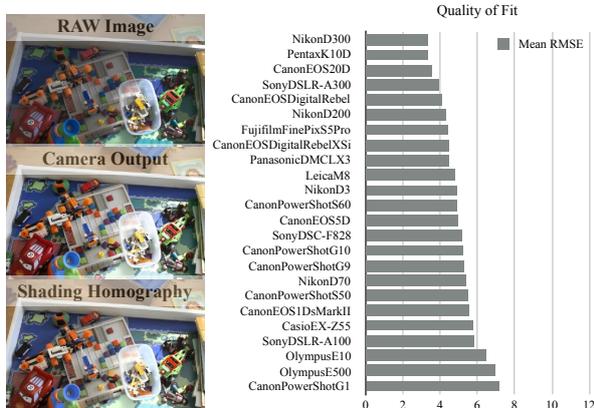
Figure 10: RAW-to-JPEG Approximation. Left: The non-linear mapping from a RAW image to its rendered camera output image can be well approximated by a shading homography. Right: RMSE between estimation and ground truth using our shading homography model of the 24 RAW-capable cameras in the Middlebury dataset [69].

where the RAW and rendered pair is drawn from the Middlebury dataset [69]. But, why should a shading homography relate a RAW input to a camera output?

To a first approximation that an $n \times 3$ matrix of pixels, A, are captured under RAW (*i.e.* linear conditions) and that a corresponding sRGB image B, to a tolerable approximation [70] – for real cameras – can be represented as

$$B \approx f(AH) \qquad (12)$$

where H is a 3×3 matrix and $f()$ is monotonically increasing function (a camera curve). This formalism is similar to the homography formalism. Indeed the form of f is far from arbitrary. If $\underline{b}_j^\intercal = \underline{a}_j^\intercal H$ (the color corrected j^{th} pixel value), then one of the requirements for $f()$ is that

$$f(\underline{b}_j) \approx \gamma_j \underline{b}_j \qquad (13)$$

where γ_j is a scale factor. That is, the function $f()$ changes the magnitude of the vector but not its direction. If this was not the case then, as the same object is seen at different brightness levels in the same image, its color will change significantly. If we map RAW to rendered using a homgraphy then

$$f(\underline{b}_j) = \gamma_j \underline{b}_j \qquad (14)$$

In [69], for a series of 24 cameras, several RAW and sRGB images (with intensities in $[0, 255]$, for different

illuminants), of a 140 patch Macbeth SG color checker are captured. Chakrabarti *et al.* analyzed the RMSE error between the predicted sRGB images and those actually recorded where $f()$ in Equation 12 is a monotonically increasing 5th order polynomial. We repeat the same experiment for our homography-based approach and RMSE errors are reported in Figure 10 where for convenience we sort the cameras from lowest to highest error. The reader is referred to Figure 4 in [69] for comparison.

Broadly, the errors we found by homography fitting are on the same order as those found in the antecedent work [69]. However, some camera data is fit with lower error (*e.g.* the Canon Powershot P1 has a mean RMSE of 7.1 for the homography-based method but 12 using [69]). In contrast for the Olympus E500, the homography error is larger (6.9 compared with 2). However, overall the range of fitting error of the homography based method and [69] is about the same (both have an overall RMSE of ~ 5).

6   Conclusion

In this paper, we demonstrated the surprising result that colors across a change in viewing condition (changing light color, shading and camera) are well related by a homography. We apply color homographies to color correction (mapping RAW RGBs to display counterparts), color object recognition and color transfer. Our homography-based color correction algorithm delivers improved color fidelity compared with the state-of-the-art. Matching chromaticity distributions using homographies delivers leading color-based object recognition. Re-interpreting color transfer as color homography mapping inspires a new direction for natural color transfer algorithm development.

Appendix

6.1   Numerical ALS convergence to the wrong answer

The matrices A and B defined below are – when rows are interpreted as "rays" as a consequence of the Planar Homography theorem [2] – a homography apart. Equivalently, there exists a 4×4 diagonal matrix D and a 3×3

linear transform H such that $B = DAH$. Suppose,

$$A = \begin{bmatrix} 5 & 5 & 8 \\ 6 & 9 & 2 \\ 1 & 7 & 3 \\ 4 & 7 & 10 \end{bmatrix} \quad B = \begin{bmatrix} 78 & 76 & 96 \\ 107 & 82 & 71 \\ 531 & 270 & 423 \\ 87 & 74 & 111 \end{bmatrix} \quad (15)$$

Let us find a matrix $K = DAH \approx B$ using ALS to find D and H. On convergence, to 2 decimal places, we find K:

$$K = \begin{bmatrix} 78.31 & 76.23 & 96.41 \\ 109.33 & 56.28 & 85.89 \\ 532.63 & 276.44 & 421.57 \\ 87.37 & 73.94 & 111.62 \end{bmatrix} \quad (16)$$

The % error ($\|K - B\| / \|B\|$) between K and B is almost 4%. According to our Theorem 2, when we converge to the wrong answer the null-space vector of $B - K$ should be sparse (1, 2, or 3 of the elements should be non-zero but not all 4). For this example, the unit length vector $\underline{v}$ which is orthogonal to $A - B$, again to 2 decimal places is: $\underline{v} = \begin{bmatrix} -0.74 & 0 & 0 & 0.68 \end{bmatrix}^{\mathsf{T}}$ (as a numerical check, $\underline{v}^{\mathsf{T}}(A - B) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, to 2 decimal places).

6.2 Closed-Form Homography

Here we present a new closed form solution for solving for the Homography matrix H that maps 4 points in one view to corresponding points in a second view. Let us denote the paired matching points in the 4×2 matrices $\mathcal{A}$ and $\mathcal{B}$ (the x and y coordinates are in the first two columns). Moving to homogeneous coordinates – we add a vector of 1's to each matrix – to make 4×3 matrices A and B. Let the operator $\mathbf{diag}(\underline{v})$ return a diagonal matrix with components of $\underline{v}$ along its diagonal. Let $M_{1:3}$ denote the first 3 rows of a matrix and that M_k is the kth row vector.

Theorem 3 (Closed-Form Homography). *The Homography matrix H relating the 4×3 matrices A and B (for the 4×2 matched $\mathcal{A}$ and $\mathcal{B}$, where no three points in either matrix are collinear) can be written in closed-form as:* $H = [A_{1:3}]^{-1}\mathbf{diag}(A_4[A_{1:3}]^{-1})^{-1}\mathbf{diag}(B_4[B_{1:3}]^{-1})B_{1:3}.$

Proof. From the non-collinearity assumption $A_{1:3}$ and $B_{1:3}$ are full rank 3×3 invertible matrices. We define X:

$$X = A[A_{1:3}]^{-1} = \begin{bmatrix} \mathcal{I} \\ A_4[A_{1:3}]^{-1} \end{bmatrix}$$

$$X\mathbf{diag}(A_4[A_{1:3}]^{-1})^{-1}\mathbf{diag}(B_4[B_{1:3}]^{-1}) =$$

$$\begin{bmatrix} \mathbf{diag}(A_4[A_{1:3}]^{-1})^{-1}\mathbf{diag}(B_4[B_{1:3}]^{-1}) \\ B_4[B_{1:3}]^{-1} \end{bmatrix}$$

Let D denote a 4×4 diagonal matrix such that the upper left 3×3 sub-matrix (first 3 rows and columns) is

$$\mathbf{diag}(B_4[B_{1:3}]^{-1})^{-1}\mathbf{diag}(A_4[A_{1:3}]^{-1})$$

and the 4th component along the diagonal of D is 1. We define Y:

$$Y = DX\mathbf{diag}(A_4[A_{1:3}]^{-1})^{-1}\mathbf{diag}(B_4[B_{1:3}]^{-1})$$
$$= \begin{bmatrix} \mathcal{I} \\ B_4[B_{1:3}]^{-1} \end{bmatrix}$$

which follows that $YB_{1:3} = B$ *i.e.* the matrix B. Substitute for Y and X and post-multiply by $B_{1:3}$, we can write:

$$A[A_{1:3}]^{-1}\mathbf{diag}(A_4[A_{1:3}]^{-1})^{-1}\mathbf{diag}(B_4[B_{1:3}]^{-1})B_{1:3} = DB.$$

Dividing by the first two columns of DB by the third we end up with the original point set $\mathcal{B}$. $\square$

Acknowledgment

17

References

[1] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007. 1

[2] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004. 1, 2, 4, 6, 8, 16

[3] E. Land, "The retinex theory of color vision," *Scientific American*, pp. 108–129, 1977. 1, 5

[4] D. Forsyth, "A novel algorithm for color constancy," *International Journal of Computer Vision*, vol. 5, pp. 5–36, 1990. 1, 5

[5] P. Hubel, J. Holm, and G. Finlayson, "Illuminant estimation and colour correction," in *The Colour in Multimedia Conference*, 1998, pp. 97–105. 1, 5

[6] M. J. Vrhel and H. J. Trussell, "The mathematics of color calibration," in *International Conference on Image Processing*, 1998, pp. 181–185.

[7] M. Drew and B.V. Funt, "Natural metamers," *Graphical Models and Image Processing*, vol. 56, pp. 139–151, 1992. 1, 5

[8] B. Wandell, "The synthesis and analysis of color images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 9, pp. 2–13, 1987. 1

[9] R. Lenz, L. V. Tran, and P. Meer, "Moment based normalization of color images," in *Multimedia Signal Processing, 1999 IEEE 3rd Workshop on*, 1999, pp. 103–108. 1, 5

[10] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, "Proposal for a standard default color space for the internet:srgb," in *Color and Imaging Conference*, 1995, pp. 238–245. 3, 9

[11] S. Bianco and C. Cusano, "Color target localization under varying illumination conditions," in *Computational Color Imaging*, ser. Lecture Notes in Computer Science, R. Schettini, S. Tominaga, and A. Trmeau, Eds. Springer Berlin Heidelberg, 2011, vol. 6626, pp. 245–255. 2, 9

[12] A. Minagawa, Y. Katsuyama, H. Takebe, and Y. Hotta, "A color chart detection method for automatic color correction," in *International Conference on Pattern Recognition*, Nov 2012, pp. 1912–1915. 2, 9

[13] B. Funt and P. Bastani, "Irradiance-independent camera color calibration," *Color Research & Application*, vol. 39, no. 6, pp. 540–548, 2014. 2, 5, 9, 10

[14] G. D. Finlayson, M. Mohammadzadeh Darrodi, and M. Mackiewicz, "The alternating least squares technique for non-uniform intensity color correction," *Color Research & Application*, vol. 40, no. 3, pp. 232–242, 2015. 2, 5, 6, 10

[15] M. Swain and D. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 11, pp. 11–32, 1991. 2, 11

[16] W. Niblack and R. Barber, "The QBIC project: Querying images by content using color, texture and shape," in *Storage and Retrieval for Image and Video Databases I, volume 1908 of SPIE Proceedings Series*, 1993. 2

[17] B. Funt and G. Finlayson, "Color constant color indexing," *Transaction on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 522–529, May 1995. 2

[18] R. M. H. Nguyen, S. J. Kim, and M. S. Brown, "Illuminant aware gamut-based color transfer," *Computer Graphics Forum*, vol. 33, no. 7, pp. 319–328, Oct. 2014. 4, 13, 14, 15

[19] J. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The amsterdam library of object images," *International Journal of Computer Vision*, vol. 61, no. 1, pp. 103–112, 2005. 4, 10

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004. 4, 8

[21] G. Yu and J.-M. Morel, "ASIFT: An Algorithm for Fully Affine Invariant Comparison," *Image Processing On Line*, vol. 1, 2011. 4, 8, 9

[22] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000. 4

[23] S. Gefen, Y. Fan, L. Bertrand, and J. Nissanov, "Symmetry-based 3d brain reconstruction." in *International Symposium on Biomedical Imaging*. IEEE, 2004, pp. 744–747. 4

[24] C. T. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision." in *International Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 1999, pp. 1125–1131. 4

[25] P. H. S. Torr and A. Zisserman, "Feature based methods for structure and motion estimation," in *International Workshop on Vision Algorithms: Theory and Practice*, ser. International Conference on Computer Vision. London, UK, UK: Springer-Verlag, 2000, pp. 278–294. 4

[26] L. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *Journal of the Optical Society of America A*, vol. 3, pp. 1673–1683, 1986. 5

[27] J. Parkkinen, J. Hallikanen, and T. Jaaskelainen, "Characteristic spectra of munsell colors," *Journal of the Optical Society of America A*, vol. 6, pp. 318–322, 1989.

[28] M. Vrhel and H. Trussel, "Color correction using principal components," *Color Research and Application*, vol. 17, pp. 328–338, 1992. 5

[29] G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd ed. Wiley, New York, 1982. 5, 9

[30] D. Marimont and B. Wandell, "Linear models of surface and illuminant spectra," *Journal of the Optical Society of America A*, vol. 9, no. 11, pp. 1905–1913, 92. 5

[31] B. Funt and H. Jiang, "Nondiagonal color correction," in *International Conference on Image Processing*, vol. 1, Sept 2003, pp. I–481–4 vol.1. 5

[32] G. D. Finlayson, B. V. Funt, and H. Jiang, "Predicting cone quantum catches under illumination change," in *Color and Imaging Conference*, 2003, pp. 170–174. 5

[33] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources," *Transactions on Image Processing*, vol. 21, no. 2, pp. 697–707, 2012. 5

[34] I. Boyadzhiev, K. Bala, S. Paris, and F. Durand, "User-guided white balance for mixed lighting conditions." *ACM Trans. Graph.*, vol. 31, no. 6, p. 200, 2012.

[35] S. Beigpour, C. Riess, J. Van de Weijer, and E. Angelopoulou, "Multi-illuminant estimation with conditional random fields," *Transactions on Image Processing*, vol. 23, no. 1, pp. 83–96, 2014. 5

[36] R. Kondepudy and G. Healey, "Use of invariants for recognition of 3-dimensional color textures," *Journal of the Optical Society of America A*, vol. 11, no. 11, pp. 3037–3049, November 1994. 5

[37] J. Berens and G. Finlayson, "Log-opponent chromaticity coding of colour space," in *International Conference on Pattern Recognition*, vol. 1, 2000, pp. 206–211 vol.1. 5

[38] G. Finlayson, M. Drew, and B. Funt, "Spectral sharpening: sensor transformations for improved color constancy," *Journal of the Optical Society of America A*, vol. 11, no. 5, pp. 1553–1563, May 1994. 5, 11

[39] H. Y. Chong, S. J. Gortler, and T. Zickler, "The von kries hypothesis and a basis for color constancy," in *International Conference on Computer Vision*, 2007, pp. 1–8. 5, 11

[40] G. Finlayson, B. Schiele, and J. Crowley, "Comprehensive colour image normalization," in *European Conference on Computer Vision*, European Vision Society. Springer Verlag, 1998, pp. 475–489. 5, 11

[41] T. Gevers and A. W. M. Smeulders, "Color-based object recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 3, pp. 453–464, 1999. 5, 11

[42] D. Berwick and S. Lee, "A chromaticity space for specularity-, illumination color- and illumination pose-invariant 3-d object recognition," in *International Conference on Computer Vision*, 1998, p. Session 2.3. 5

[43] P. Hubel, G. Fnlayson, and S. Hordley, "White point estimation using color by convolution," 2003, united States Patent No. 20030194125. 5

[44] J. T. Barron, "Convolutional color constancy," in *International Conference on Computer Vision*. IEEE, 2015, pp. 379–387. 5

[45] Y. Ohno and J. Hardis, "Four-color matrix method for correction of tristimulus colorimeters," in *Color and Imaging Conference*, 1997, pp. 301–305. 5

[46] G. Finalyson, H. Gong, and R. Fisher, "Color homography," in *Progress in Colour Studies*, 2016. 6

[47] G. D. Finalyson, H. Gong, and R. B. Fisher, "Color homography color correction," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2016. 6

[48] T. Zhang and G. H. Golub, "Rank-one approximation to high order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 2, pp. 534–550, 2001. 6, 7

[49] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *British Machine Vision Conference*. BMVA Press, 2002, pp. 36.1–36.10, doi:10.5244/C.16.36. 8

[50] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European Conference on Computer Vision*. Springer, 2006, pp. 404–417. 8

[51] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey vision conference*, vol. 15. Citeseer, 1988, p. 50. 8

19

[52] C. McCamy, H. Marcus, and J. Davidson, "A color-rendition chart," *Journal of Applied Photographic Engineering*, pp. 95–99, 1976. 9

[53] G. D. Finlayson, M. Mackiewicz, and A. Hurlbert, "Color correction using root-polynomial regression," *Transactions on Image Processing*, vol. 24, no. 5, pp. 1460–1470, 2015. 10

[54] C. L. Novak and S. A. Shafer, "Method for estimating scene parameters from color histograms," *Journal of the Optical Society of America*, vol. 11, no. 11, pp. 3020–3036, 1994. 10

[55] G. Healey, "Estimating spectral reflectance using highlights," *Image and vision computing*, vol. 9, no. 5, pp. 333–337, 1991. 10

[56] B. V. Funt and M. S. Drew, "Color space analysis of mutual illumination," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1319–1326, 1993. 10

[57] R. Bajcsy, S. W. Lee, and A. Leonardis, "Color image segmentation with detection of highlights and local illumination induced by inter-reflections," in *International Conference on Pattern Recognition, 1990. Proceedings*, vol. 1. IEEE, 1990, pp. 785–790.

[58] S. Tominaga and T. Okayama, "A method for analyzing color images containing highlight and interreflection effects," *Systems and computers in Japan*, vol. 28, no. 13, pp. 8–18, 1997. 10

[59] Y. Hwang, J.-Y. Lee, I. S. Kweon, and S. J. Kim, "Color transfer using probabilistic moving least squares," in *Conference on Computer Vision and Pattern Recognition*. IEEE, 2014, pp. 3342–3349. 13

[60] H. Gong, G. D. Finlayson, and R. B. Fisher, "Recoding color transfer as a color homography," in *British Machine Vision Conference*. BMVA, 2016. 13

[61] F. Pitié and A. Kokaram, "The linear monge-kantorovitch linear colour mapping for example-based colour transfer," in *European Conference on Visual Media Production*. IET, 2007. 13, 14

[62] F. Pitié, A. C. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer," *Computer Vision And Image Understanding*, vol. 107, no. 1-2, pp. 123–137, Jul. 2007. 13, 14, 15

[63] T. Pouli and E. Reinhard, "Progressive histogram reshaping for creative color transfer and tone reproduction," in *International Symposium on Non-Photorealistic Animation and Rendering*. New York, NY, USA: ACM, 2010, pp. 81–90. 13, 14, 15

[64] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics And Applications*, vol. 21, no. 5, pp. 34–41, Sep. 2001. 13, 14

[65] N. Thomos, N. V. Boulgouris, and M. G. Strintzis, "Optimized transmission of jpeg2000 streams over wireless channels," *Transactions on Image Processing*, vol. 15, no. 1, pp. 54–67, 2006. 13

[66] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. 13

[67] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," in *ACM Transactions on Graphics*, vol. 26, no. 3. ACM, 2007, p. 96. 15

[68] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," in *European Conference on Computer Vision*. Springer, 2006, pp. 568–580. 15

[69] A. Chakrabarti, D. Scharstein, and T. Zickler, "An empirical camera model for internet color vision." in *British Machine Vision Conference*, vol. 1, no. 2. BMVA, 2009, p. 4. 16

[70] S. J. Kim, H. T. Lin, Z. Lu, S. Susstrunk, S. Lin, and M. S. BrownHai, "A new in-camera imaging model for color computer vision and its application," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012. 16