Mobile Image Based Color Correction Using Deblurring

Yu Wang,[a] Chang Xu,[a] Carol Boushey,[b,c] Fengqing Zhu[a] and Edward J. Delp[a]

[a]*School of Electrical and Computer Engineering and* [b]*Department of Nutrition Science, Purdue University, West Lafayette, Indiana, 47906, USA.*
[c]*Epidemiology Program, University of Hawaii Cancer Center Honolulu, Hawaii, USA*

ABSTRACT

Dietary intake, the process of determining what someone eats during the course of a day, provides valuable insights for mounting intervention programs for prevention of many chronic diseases such as obesity and cancer. The goals of the Technology Assisted Dietary Assessment (TADA) System, developed at Purdue University, is to automatically identify and quantify foods and beverages consumed by utilizing food images acquired with a mobile device. Color correction serves as a critical step to ensure accurate food identification and volume estimation. We make use of a specifically designed color checkerboard (i.e. a fiducial marker) to calibrate the imaging system so that the variations of food appearance under different lighting conditions can be determined. In this paper, we propose an image quality enhancement technique by combining image de-blurring and color correction. The contribution consists of introducing an automatic camera shake removal method using a saliency map and improving the polynomial color correction model using the LMS color space.

Keywords: color correction, image deblurring

1. INTRODUCTION

Dietary intake, the process of determining what someone eats during the course of a day, provides valuable insights for mounting intervention programs for prevention of many chronic diseases such as obesity and cancer. Accurate methods and tools to assess food and nutrient intake are essential for epidemiological and clinical research on the association between diet and health. The goals of the Technology Assisted Dietary Assessment (TADA) System, developed at Purdue University, is to automatically identify and quantify foods and beverages consumed by utilizing one image of a users food acquired with a mobile device (mobile telephone) [1]–[4]. Figure 1 illustrates the overall architecture of our proposed system. First, images acquired with the mobile telephone and metadata are sent to the backend server. Then, the image analysis will be done on the server, which consists of four main steps: preprocessing, segmentation, classification and volume estimation [5]–[7]. Every step is essential in the sense that a relatively small error may result in misclassification of the food or errors in the volume estimation. Next, the automatic segmentation results are sent back to the user, so that the user can confirm and/or modify food segments. Finally, the confirmed results from users are processed on the server. Based on the user feedback, refinement is done to the image segmentation and classification.

Color information is of great importance in our dietary assessment system and it serves as a key feature to identify foods [5], [6]. Thus, a consistent color descriptor of an object is critical. The colors of an object recorded by a camera depend mainly on three factors: illumination conditions in the scene (which are unknown in most cases), object intrinsic surface properties and various photometric parameters (e.g., exposure time, white balancing, gamma correction) [8]. A real world example is that the rendered colors of the same scene can be quite different even with the same camera from slightly different angles. Some approaches seek to overcome these problems by estimating illumination invariance color descriptors from training images, including the RGB histogram, color moments, and C-SIFT [9], [10]. In [9], a combined set of color descriptors with invariance properties surpass the performance of intensity based descriptors by 8% on category recognition.

This work was sponsored by grants from the National Institutes of Health under grant NIEH/NIH 2R01ES012459-06.
Address all correspondence to E. J. Delp: ace@ecn.purdue.edu or see www.tadaproject.org.

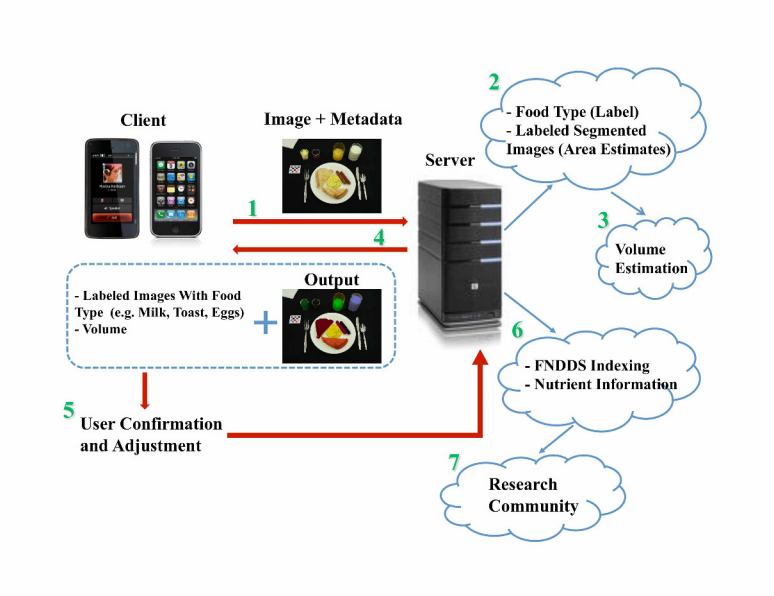


Figure 1: The architecture of TADA system.

An alternative approach to characterize the imaging properties is based on the spectral response/sensitivity of the camera. If the camera spectral sensitivity is known, then it is possible to estimate a relationship between the spectral sensitivity of the camera and the CIE color matching functions [11]–[14]. This approach, however, is not practical for common application due to the fact that the spectral sensitivity of the camera should be measured by using specialized devices, such as monochromators, or radiance meters.

1.1 Color Calibration

Our goal is to achieve color constancy under all kinds of lighting conditions so that the color of food can be used as a proper classification feature. When we look at an image acquired by a mobile phone camera, each pixel can be represented as a function f_i, where i is the color index (e.g. R, G, B). f_i is mainly dependent on three factors: the illuminant spectral power distribution $I(\lambda)$, the surface spectral reflectance $S(\lambda)$ and the sensor spectral sensitivities $V_i(\lambda)$.

$$f_i(x, y, S) = \int I(\lambda)S(\lambda)V_i(\lambda)d\lambda, i = R, G, B \tag{1}$$

The color sensor response form a vector $\boldsymbol{F}(S) = (f_R(S), f_G(S), f_B(S))$, which is also referred as the RGB tristimulus (R, G, B). Suppose that two images have been acquired from the same scene under different lighting conditions and cameras. For any pixel in the two images,

$$\begin{array}{rcl} \mathbf{RGB}_1 & = & \mathbf{F}_1(S) \\ \mathbf{RGB}_2 & = & \mathbf{F}_2(S) \end{array} \tag{2}$$

Furthermore,

$$\mathbf{RGB}_1 = \mathbf{F}_1(S) = \mathbf{F}_1(\mathbf{F}_2^{-1}(\mathbf{RGB}_2)) \tag{3}$$

We would like to be able to express explicitly such transformation, $T = \mathbf{F}_1 * \mathbf{F}_2^{-1}(\cdot)$, between an unknown illuminant and a reference. Many chromatic adaptation techniques have been proposed to address this problem [15]–[17]. Since this is an ill-posed inverse problem most of the proposed solutions lacks uniqueness and stability. It has been shown that the universal best and the universal worst technique do not exist [18]: the method that performs best for a specific image depends on the image content.

There are generally two ways of achieving color correction. The first approach changes the overall colors in an image and is often used for colors other than neutrals to appear correct or pleasing. Methods for this type

of correction are generally known as gray balance, neutral balance or white balance [19], [20]. Gray world is one of the most well-known gray balance methods [21], [22]. It is based on the assumption that given an image with sufficient amount of color variations, the average value of the R, G, and B components of the image should average to a common gray value. Another opponent technique is known as white patch, which assumes that the maximum response in an image is caused by a perfect reflectance [23]. To combine both gray world and white patch approaches, Alessandro and Carlo Gatta proposed Automatic Color Equalization (ACE) in [24]. Their method extends the Retinex model of color equalization, merging Retinex with the Gray world and the White Patch equalization methods. Recently, the use of visual information automatically extracted from the images gas been investigated. Moreno et al. [25] obtained memory colors for three different objects (grass, snow and sky) using psychophysical experiments. They then used a supervised image segmentation method to detect memory color objects to color correct the image using a weighted Von Kries method. S. Bianco and R. Schettini [26] investigated color statistics extracted from faces in a scene to estimate illuminants. However, their results are largely based on the performance of the face detector and the knowledge of the corresponding skin color.

The second approach is usually referred to as color calibration uses the image of a reference chart for each set of acquisition conditions. Wang et al. [27] used a Munsell ColorChecker as the reference target. They then picked 13 color patches to train the parameters for the correction model. Adrian Ilie and Greg Welchin proposed a two-phase calibration technique in [28], where a 24-sample GretagMacbeth [29]ColorChecker was set up in each image acquired by different cameras. The two-phase method consists of an iterative closed-loop hardware calibration and software refinement, which is argued to ensure color constancy across multiple imaging devices.

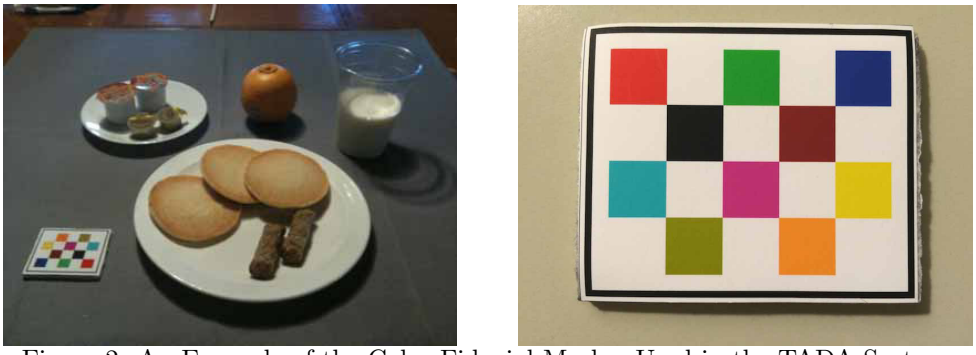


Figure 2: An Example of the Color Fiducial Marker Used in the TADA System.

From extensive studies the current TADA system has adopted the concept of using a reference target [2], [30]. Feedback from the participants in our studies indicated that it would be easy to use a credit card-sized fiducial marker due to the convenient incorporation into their current lifestyles [1], [31]. Thus, we decided to use a compact checkerboard pattern to for color calibration (see Figure 2). The color checkerboard was designed to have the dimension of $7 \times 6\,cm^2$. The color patches were chosen to cover the full color spectrum.

1.2 Single Image Deblurring

In our previous work [30], an image quality measurement method as well as a nonlinear color correction model using the CIELAB color space was proposed. However, an underlying problem persists that a user might still send a blurry image to our image analysis system even our image quality "checker" suggests retaking the image. There are cases where the user may be reluctant to retake the image of his/her meal or the image on the small mobile telephone screen appears to be good enough. Often, the checkerboard in a blurred image cannot be correctly detected and consequently color correction will be skipped by our system. This imposes a critical problem for the image analysis steps, i.e. food segmentation and identification. This then becomes a Blind Deconvolution (BD) problem with the unknown blur represented as a Point Spread Function (PSF).

Blind deconvolution is the process of recovering a sharp version of a blurry image. It is also well known to be ill-posed, small perturbations of the data produce large deviations in the resulting solution [32]. Mathematically, the general model for a linear degradation caused by blurring and additive noise is given by

$$y = h \otimes x + n \tag{4}$$

where x is a visually sharp image or original image, n is known noise and h is a nonnegative blur kernel, whose support is small compared to the image size. When the noise is ignorable, the objective of blind restoration is to estimate x and h. Often, the model above is also represented in terms of a matrix formulation, that is,

$$\vec{y} = \mathbf{H}\vec{x} \tag{5}$$

where the vectors $\vec{x}$ and $\vec{y}$ represent the original image and the observed image respectively by stacking the image matrix into a vector. $\mathbf{H}$ is a Block Toeplitz with Toeplitz Blocks (BTTB) matrix.

While numerous approaches have been described [33]–[36], the results are still far from perfect. Recent methods have characterized x using natural image statistics [37]–[40]. These techniques exhibit some common principles. A. Levin et al. [41] argued the failure of the MAP approach and suggested that the key component making blind deconvolution possible is not the choice of the prior, but the estimator. In [42], D. Krishnan et al. pointed out that there is a major drawback in many common forms of image priors because the minimum of the resulting cost function does not correspond to the true sharp solution. They proposed a new image regulation method using the ratio of the l_1 norm to the l_2 norm on the high frequencies of an image.

Based on Krishnan's approach [42], we introduce an image deblurring scheme using a saliency map. The idea behind using visual saliency is that we want to reduce the processing time by analyzing a sub-image. The sub-image should contain enough features to estimate the blur kernel. In our application, it is plausible to assume that the blur in a image is uniform and only comes from slight camera movement, such as camera shift or in-plane rotation. There are mainly two reasons for such assumption. First, we have implemented an image quality check on the mobile telephone, which should prevent users from taking blurry images or images without the fiducial marker present. Second, in our image dataset, most of the food images are acquired in a stationary environment with reasonably adequate lighting condition and the fiducial marker is always detected as a salient region, if present. Since the checkerboard region contains plenty of corner or edge features as well as a wide range of colors, the estimated blur kernel is consistent to what is analyzed from the entire image. The results show that our saliency based image deblurring is robust and fast.

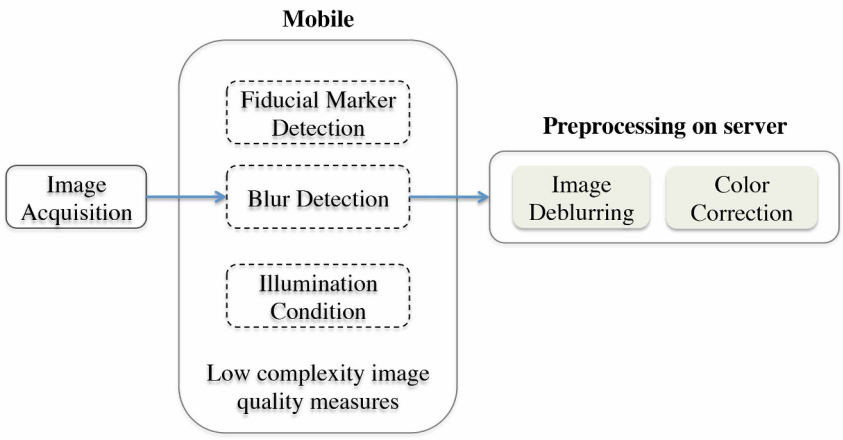


Figure 3: TADA mobile quality measure system.

2. IMAGE DEBLURRING

In this section, we explain our proposed deblurring technique in detail. Figure 3 illustrates the workflow from mobile quality measure to image preprocessing on the server. The user first takes a food image that contains the color fiducial marker under an unknown illumination. Then, several image quality checks are initiated before the user can send the food image to our backend server. The examination includes checkerboard detection [43], blur detection and a coarse illumination condition check. Due to the limited computational resources and the need for quick feedback on the mobile device, simple processing approaches are used. If the image does not pass the blur detection on the mobile phone, image deblurring will be triggered. Both image deblurring and

color correction we describe in this paper are implemented on the backend server to complete the preprocessing. We want to restore the blurry images to the maximum extent so that they can be color corrected for further analysis. As shown in Equation 4, we observe the resulting blurry image y and the goal is to recover the unknown sharp image x as well as the blur kernel h. Based on the deblurring technique proposed in [42], we introduce a faster and robust deblurring method using visual saliency. Our method dramatically reduces the computational time without sacrificing restoration quality. As shown in Figure 4, the proposed deblurring method consists of

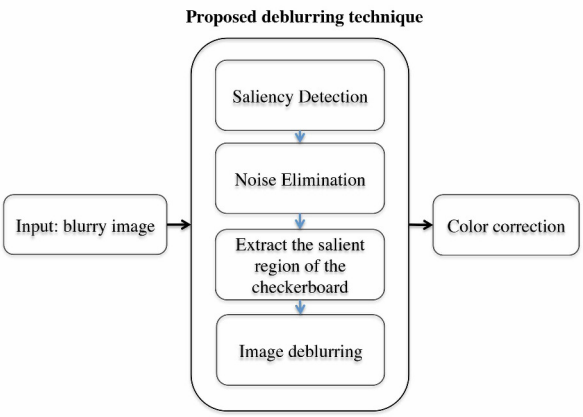


Figure 4: Proposed image deblurring technique.

four steps. Given an input blurry image, the saliency map of the image is first computed. A saliency map is a multi-scale feature map which contains local spatial discontinuities in the modalities of color, intensity and orientation. We adopted the idea of image signature for faster saliency detection, which was first introduced by X. Hou etc. in [44]. If we denote the grayscale blurry image as x, which is the mixture of foreground and background, its image signature is defined as

$$ImageSignature = sign(DCT(x)) \tag{6}$$

where DCT represents Discrete Cosine Transformation. Consider the reconstructed image $\tilde{x} = IDCT[ImageSignature]$, the saliency map s is computed by smoothing the squared reconstructed image $\tilde{x}$,

$$s = g * (\tilde{x} \circ \tilde{x})$$

where g is a Gaussian kernel, $*$ is convolution symbol and $\circ$ represents entry-wise product operator. Then we use a Flood Fill technique [45] to eliminate the noisy saliency regions, especially small blobs. If we consider

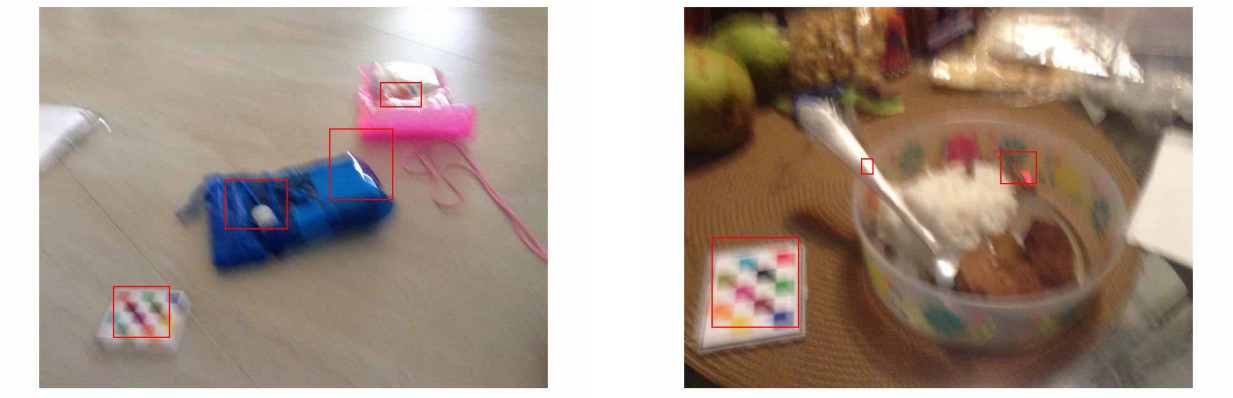


Figure 5: Examples of saliency regions.

a sub-image containing the checkerboard, it possesses many of visually recognizable features, such as corners,

edges and contrast color patches, so it almost always is detected as a salient region. However, it is likely that more than one region will be detected. Figure 5 shows two examples of the initial saliency map. Our goal is to extract the area containing the checkerboard and use that to estimate the blur kernel. This can be achieved by analyzing the histogram of each salient region. For one salient region, the histogram of each channel is split into 8 bins, which can be represented by an 8 dimensional vector. The element with the largest value is discarded and 7 other elements are then normalized. By combining 3 channels, we get a 21 dimensional feature vector. The cross-correlation between such feature vector and the reference checkerboard histogram is computed to find the optimal match.

After we extract the checkerboard region, we use Krishnan's approach [42] to estimate the blur kernel. Krishnan's method is described as follows:

1. Use derivative high-pass filters on the blurry image y, creating a high-frequency image g
2. Blind multi-scale estimation of blur matrix h from g using a coarse-to-fine pyramid of image resolutions. At each scale, update sharp high-frequency image g and h using l_1/l_2 regularization. Use bilinear interpolation to up-sample the current kernel to finer level as initialization.
3. Image recovery using non-blind algorithm [46].

3. COLOR CORRECTION

In this section, we describe the selection of the color space we use for color correction and then compare our method to our previous work [30]. Color correction is done on the backend server, which mainly consists of color extraction from the checkerboard and color mapping to the D65 reference and matching the acquired image to the reference lighting condition [30]. In this paper we investigate polynomial models using three different color spaces and compare them to our previous work.

3.1 Color Space Models

A color space is a mathematical model used to describe how colors can be interpreted as tuples, typically of three or four elements. sRGB color space is commonly used in mobile cameras and displays [8]. The transformation between sRGB color space and linear RGB space is defined as follows:

$$C_{sRGB} = \begin{cases} 12.92C_{linear}, C_{linear} \leq 0.0031308 \\ 1.055{C_{linear}}^{1/\gamma} - 0.055, C_{linear} > 0.0031308 \end{cases}$$

where C represents R,G or B channel and γ is the gamma correction value. In our previous work [30], the checkerboard image captured using a mobile telephone camera under the D65 illumination was used as the reference. We implemented a different approach for measuring the color patch. The checkerboard shown in Figure 2 is placed inside a "SpectraLight II" illumination booth and we use a spectral radiometer to measure each color patch on the checkerboard. The output of spectral radiometer is in XYZ color space. The conversion from linear RGB color space to CIEXYZ color space is defined as [8],

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix}$$

The inverse conversion from CIEXYZ to linear RGB color space is,

$$\begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = M^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Here, we propose to use LMS color space for color correction. LMS color space is derived from the human visual system. Humans have three distinct types of color receptors, which are referred to as *long,medium* and

short cones [47]. Though the LMS color space is not commonly used in color specification, it is often used for chromatic adaptation. It has simple and positive color matching function for each channel. It is computationally simpler compared to other nonlinear color spaces, such as the CIELAB color space. It is also proportional to the illuminant energy. The coordinates in the XYZ system are related to LMS through the following transformation,

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.2430 & 0.8560 & -0.044 \\ -0.3910 & 1.1650 & 0.0870 \\ 0.0100 & -0.0080 & 0.5630 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{7}$$

3.2 Proposed Color Correction

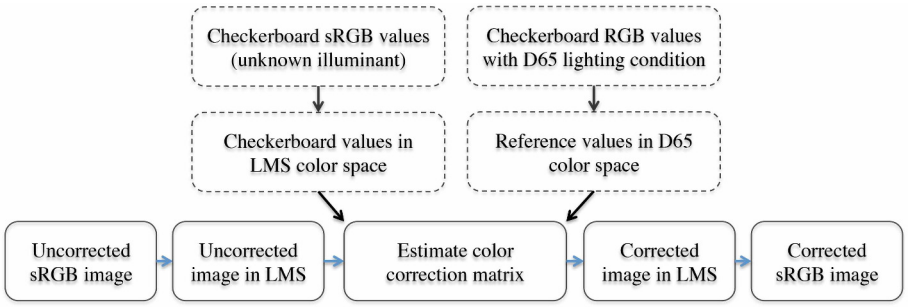


Figure 6: Diagram of our proposed color correction method.

Figure 6 shows a diagram of our proposed color correction method. The uncorrected image is first converted into LMS color space. Then, we implement and optimize the polynomial transforms to find the correction matrix. For each color patch, the tristimulus values in LMS color space can be represented as a vector $V : (L_i, M_i, S_i)^T (i = 1, 2, \ldots, 11)$, we have 11 colors on the checkerboard including black and white. Similarly, the reference checkerboard has 11 corresponding color values, denoted as $R : (RL_i, RM_i, RS_i)^T (i = 1, 2, \ldots, 11)$. We use the following vector $X : [L, M, S, LM, LS, MS, 1]^T$ to estimate the color correction matrix. The transformation model can be represented as,

$$\begin{cases} CL_i = a_{11}L_i + a_{12}M_i + a_{13}S_i + a_{14}LM_i + a_{15}LS_i + a_{16}MS_i + a_{17} \\ CM_i = a_{21}L_i + a_{22}M_i + a_{23}S_i + a_{24}LM_i + a_{25}LS_i + a_{26}MS_i + a_{27} \\ CS_i = a_{31}L_i + a_{32}M_i + a_{33}S_i + a_{34}LM_i + a_{35}LS_i + a_{36}MS_i + a_{37} \end{cases}$$

where CL_i,CM_i and CS_i are the corrected tristimulus. The equation can also be rewritten in matrix form as,

$$\begin{bmatrix} CL \\ CM \\ CS \end{bmatrix} = A^T X = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \\ a_{15} & a_{25} & a_{35} \\ a_{16} & a_{26} & a_{36} \\ a_{17} & a_{27} & a_{37} \end{bmatrix}^T \begin{bmatrix} L \\ M \\ S \\ LM \\ LS \\ MS \\ 1 \end{bmatrix} \tag{8}$$

where A is the color correction matrix and X is the polynomial combination. Thus, we want to find a matrix A, which minimizes the overall error between the corrected image and the reference across all 11 color patches. We further formulate this problem as finding constrained least square solution. By using the notation above, we have

$$A = \arg\min_X \sum_{i=1}^{11} \|AX - R\|^2 \tag{9}$$

Equation 9 can be solved using Levenberg-Marquardt methods [48]. Finally, we correct the image in the LMS color space using A and convert it back to the sRGB space for display.

4. EXPERIMENTAL RESULTS

To evaluate our visual saliency based image deblurring technique, we manually choose 25 blurry images from the TADA free-living study. The free-living study contains a number of 315 meal images, which were acquired under natural eating conditions by 11 participants. Two examples are shown in Figure 5. All food images acquired in

(a) Checkerboard can be detected. (b) Checkerboard cannot be detected.

Figure 7: Examples of deblurred images.

this study were taken under natural eating conditions by our participants. We also acquired another 25 images of plastic food using Samsung Galaxy Nexus. When acquiring those images, we tried to simulate the real life situation by deliberately moving the camera slightly to create blur effect. Now, we have a total of 50 images as testing data. The TADA checkerboard was included in all the images, but none could be detected due to blurriness. After applying our method to the testing images, 31 out of 50 were correctly detected to have the TADA checkerboard. Therefore, color correction can be applied to them.

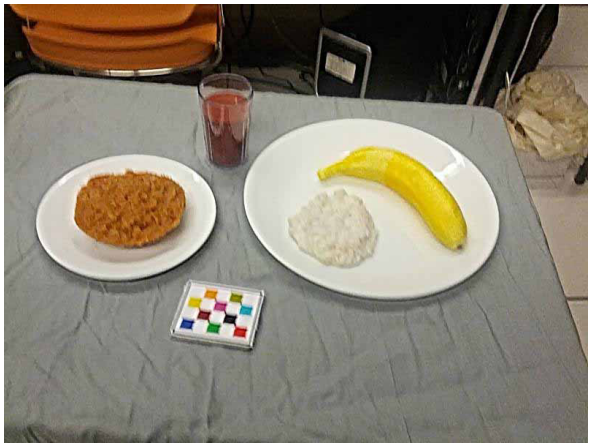
Figure 7 illustrates two deblurred images corresponding to the original ones in Figure 5. The one on the left was detected to have the TADA checkerboard after image deblurring, even though it seems visually unpleasant in Figure 7. The image on the right failed to be detected with a checkerboard pattern due to the heavy rotation and shift of the camera. A plastic food image example is shown in Figure 8. The fiducial marker was correctly located in the deblurred image in this case. In addition to the robust image restoration, our method achieved approximately $\frac{1}{6}$ of computational time compared to Krishnan's algorithm without using saliency map. As for the two cases in Figure 7(from left to right), Krishnan's algorithm took 51.45s and 74.42s respectively and the proposed method consumed 8.13s and 11.51s including saliency detection as well as deblurring. All the experiments were conducted on OSX Yosemite with 2.6G quad core i7 CPU and 16G RAM. The images were scaled to 800×600 to speed up the process.

To evaluate the performance of the proposed color correction methods, we use GretagMacbeth Colorchecker [49] as the testing target. Both the TADA fiducial marker and GretagMacbeth Colorchecker were placed inside a SpectraLight II illumination booth. We acquired several images of the two checkerboards using four different illuminants, i.e. simulated daylight (CIE D65, 6500 K), horizon daylight (simulated early morning sunrise or afternoon sunset, 2300 K), CIE A (incandescent home lighting, 2856 K), and commercial fluorescent (cool white, 4000 K). All the images were acquired using iPhone 5 camera.

The captured images with non-D65 illuminations were corrected using the method described in Section 3. Here, we compared our proposed method with the similar approach using the CIELAB and sRGB color spaces. Let $(R_i, G_i, B_i)(i = 1, 2, \ldots, 24)$ denote the reference color of each patch in the GretagMacbeth Colorchecker and $(R_i^c, G_i^c, B_i^c)(i = 1, 2, \ldots, 24)$ be the corrected values of corresponding patch under various lighting conditions. The average Euclidean distance for all 24 pairs is defined as:

$$\Delta = \frac{1}{24} \sum_{i=1}^{24} \left\| (R_i^c, G_i^c, B_i^c)^T - (R_i, G_i, B_i)^T \right\| \tag{10}$$

(a) The original image.



(b) The deblurred image.

Figure 8: Example of blurry and deblurred plastic food images.

Table 1 shows the mean error between the reference and corrected images for different methods. The entry *Total* in the table is simply the summation of R, G and B channel errors. The column of LMS demonstrates the error of the method we proposed and the results from the similar technique using sRGB and CIELAB color space verified our choice of choose LMS as the color correction space. Even though our method does not produce consistently the smallest error for each channel, the overall RGB error is approximately a 10% improvement compared with the correction method using CIELAB color space. This implies that even though the CIELAB color space is uniform with respect to the Human Visual System (HSV), it is not necessarily the best choice when it comes to the linear color correction model, since the model expects each channel to be correlated when combining the polynomial terms. In our experiment, we used the fixed gamma value of 2.2 based on te iPhone 5's camera specification. By examining some of the most popular smart phones on the market, we concluded that such gamma is plausible (see Table 2).

Table 1: Errors (Δ) Between the Reference Image and Corrected Images

Lighting	Error	LAB	sRGB	LMS
Incandescent	Red	7.98	7.44	7.76
	Green	9.54	8.86	7.59
	Blue	10.56	9.58	7.56
	Total	28.08	25.88	22.91
Horizon Light	Red	7.53	6.34	3.30
	Green	3.84	3.98	3.85
	Blue	11.85	11.55	9.37
	Total	23.22	21.87	16.52
Coolwhite	Red	3.54	3.44	3.22
	Green	4.18	4.15	4.31
	Blue	3.88	3.65	2.89
	Total	11.60	11.24	10.52

Table 2: Gamma correction values of popular smart phones

Mobile Phones	iPhone 6	iPhone 5	iPhone 4s	Samsung Galaxy S5	Samsung Galaxy S4
Gamma	2.23	2.22	2.1	2.25	2.16

5. CONCLUSION AND FUTURE WORK

In this paper we presented an improved polynomial color correction model using LMS color space and proposed a visual saliency based image deblurring method. We have shown that image deblurring using a saliency map is

plausible in the TADA system because we have enough prior knowledge of food images. The proposed technique recovers more than 60 percent of the images in our dataset which could not be color corrected previously. It is approximately 5 times faster than Krishnan's method [42]. The color correction model in LMS color space demonstrates more accuracy compared to other color correction models using CIELAB or sRGB space. In the future we would like to explore color correction using multiple lighting sources and generalize the saliency based method for image deblurring or segmentation.

REFERENCES

[1] B. Six, T. Schap, F. Zhu, A. Mariappan, M. Bosch, E. Delp, D. Ebert, D. Kerr, and C. Boushey, "Evidence-based development of a mobile telephone food record," *Journal of American Dietetic Association*, vol. 110, pp. 74–79, January 2010.

[2] F. Zhu, M. Bosch, I. Woo, S. Kim, C. Boushey, D. Ebert, and E. Delp, "The use of mobile devices in aiding dietary assessment and evaluation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 4, pp. 756–766, August 2010.

[3] M. Bosch, F. Zhu, T. Schap, C. Boushey, D. Kerr, N. Khanna, and E. Delp, "An integrated image-based food database system with application in dietary assessment," *Presentation at the 2010 mHealth Summit*, November 2010, Washington, DC.

[4] Z. Ahmad, N. Khanna, D. Kerr, C. Boushey, and E. Delp, "Image enhancement and quality measures for dietary assessment using mobile devices," *Proceedings of the IS&T/SPIE Conference on Mobile Devices and Multimedia: Enabling Technologies, Algorithms and Applications*, vol. 9030, pp. 903 007–1–9, February 2014, San Francisco, CA.

[5] F. Zhu, M. Bosch, N. Khanna, C. Boushey, and E. Delp, "Multiple hypotheses image segmentation and classification with application to dietary assessment," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 377–388, January 2015.

[6] Y. He, C. Xu, N. Khanna, C. Boushey, and E. Delp, "Analysis of food images: Features and classification," *Proceedings of the IEEE International Conference on Image Processing*, October 2014, Paris, France.

[7] C. Xu, Y. He, N. Khanna, C. Boushey, and E. Delp, "Model-based food volume estimation using 3D pose." *Proceeding of the IEEE International Conference on Image Processing*, pp. 2534–2538, September 2013, Melbourne, Australia.

[8] G. Sharma, *Digital Color Imaging Handbook*. CRC Press, 2002, Boca Raton, FL.

[9] K. Van De Sande, T. Gevers, and C. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1582–1596, September 2010.

[10] F. Mindru, T. Tuytelaars, L. Van Gool, and T. Moons, "Moment invariants for recognition under changing viewpoint and illumination," *Computer Vision and Image Understanding*, vol. 94, no. 1-3, pp. 3–27, April/May/June 2004.

[11] G. Hong, R. Luo, and P. Rhodes, "A study of digital camera colorimetric characterisation based on polynomial modelling," *Color Research and Application*, vol. 26, no. 1, pp. 76–84, February 2001.

[12] R. Bala, G. Sharma, V. Monga, and J.-P. Van de Capelle, "Two-dimensional transforms for device color correction and calibration," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1172–1186, August 2005.

[13] V. Cheung, S. Westland, and M. Thomson, "Accurate estimation of the nonlinearity of input/output response for color cameras," *Color Research & Application*, vol. 29, no. 6, pp. 406–412, December 2004.

[14] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy algorithms. II. experiments with image data," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 985–996, September 2002.

[15] S. Bianco and R. Schettini, "Two new Von Kries based chromatic adaptation transforms found by numerical optimization," *Color Research & Application*, vol. 35, no. 3, pp. 184–192, June 2010.

[16] B. Funt and H. Jiang, "Nondiagonal color correction," *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, pp. 481–484, September 2003, Barcelona, Spain.

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

[18] S. Bianco, F. Gasparini, and R. Schettini, "Consensus-based framework for illuminant chromaticity estimation," *Journal of Electronic Imaging*, vol. 17, no. 2, pp. 023 013–023 013–9, May 2008.

[19] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Applications*, vol. 21, no. 5, pp. 34–41, September/October 2001.

[20] S. Srivastava, C. Xu, and E. Delp, "White synthesis with user input for color balancing on mobile camera systems," *Proceedings of the IS&T/SPIE Conference on Multimedia on Mobile Devices*, vol. 8304, pp. 83 040F–1–8, Jan 2012, San Francisco, CA.

[21] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 1–26, July 1980.

[22] G. Finlayson, S. Hordley, and P. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1209–1221, November 2001.

[23] J. Von Kries, "Sources of color science," *Chromatic adaptation*, pp. 109–119, 1970, Cambridge, MA.

[24] A. Rizzi, C. Gatta, and D. Marini, "A new algorithm for unsupervised global and local color correction," *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1663–1677, July 2003.

[25] A. Moreno, B. Fernando, B. Kani, S. Saha, and S. Karaoglu, "Color correction: a novel weighted Von Kries model based on memory colors," *Computational Color Imaging*. Springer, April 2011, pp. 165–175, Milan, Italy.

[26] S. Bianco and R. Schettini, "Color constancy using faces," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 65–72, June 2012, Providence, RI.

[27] X. Wang and D. Zhang, "An optimized tongue image color correction scheme," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 6, pp. 1355–1364, September 2010.

[28] A. Ilie and G. Welch, "Ensuring color consistency across multiple cameras," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, pp. 1268–1275, June 2005, San Diego, CA.

[29] A. Munsell and M. Color, "Munsell soil color charts," 2000, New Windsor, NY.

[30] C. Xu, F. Zhu, N. Khanna, C. Boushey, and E. Delp, "Image enhancement and quality measures for dietary assessment using mobile devices," *Proceedings of the IS&T/SPIE Conference on Computational Imaging X*, vol. 8296, pp. 82 960Q–1–10, January 2012, San Francisco, CA.

[31] T. Schap, B. Six, E. Delp, D. Ebert, D. Kerr, and C. Boushey, "Adolescents in the United States can identify familiar foods at the time of consumption and when prompted with an image 14 h postprandial, but poorly estimate portions," *Public Health Nutrition*, vol. 14, no. 7, pp. 1184–1191, February 2011.

[32] P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*. CRC press, 2007, Boca Raton, FL.

[33] G. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Optics Letters*, vol. 13, no. 7, pp. 547–549, July 1988.

[34] A. Katsaggelos and K.-T. Lay, "Maximum likelihood blur identification and image restoration using the EM algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 729–733, March 1991.

[35] C. Likas and N. Galatsanos, "A variational approach for bayesian blind image deconvolution," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2222–2233, August 2004.

[36] R. Lane and R. Bates, "Automatic multidimensional deconvolution," *Journal of the Optical Society of America A*, vol. 4, no. 1, pp. 180–188, January 1987.

[37] J. Miskin and D. MacKay, "Ensemble learning for blind image separation and deconvolution," *Advances in Independent Component Analysis*. Springer, 2000, pp. 123–141, London, UK.

[38] A. Levin, "Blind motion deblurring using image statistics," *Processings of Advances in Neural Information Processing Systems*, pp. 841–848, December 2006, Vancouver, Canada.

[39] M. Bronstein, A. Bronstein, M. Zibulevsky, and Y. Zeevi, "Blind deconvolution of images using optimal sparse representations," *IEEE Transactions on Image Processing*, vol. 14, no. 6, pp. 726–736, June 2005.

[40] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, July 2006.

[41] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1964–1971, June 2009, Miami Beach, FL.

[42] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 233–240, June 2011, Providence, RI.

[43] C. Xu, N. Khanna, C. Boushey, and E. Delp, "Low complexity image quality measures for dietary assessment using mobile devices," *Proceedings of the IEEE International Symposium on Multimedia*, pp. 351–356, December 2011, Dana Point, CA.

[44] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, July 2012.

[45] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, July 1991.

[46] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," *Proceedings of Advances in Neural Information Processing Systems*, pp. 1033–1041, December 2009, Vancouver, Canada.

[47] B. Wandell, *Foundations of Vision*. Sinauer Associates, Inc., 1995, sunderland, MA.

[48] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial & Applied Mathematics*, vol. 11, no. 2, pp. 431–441, June 1963.

[49] D. Pascale, "RGB coordinates of the Macbeth color checker," *The BabelColor Company*, pp. 1–16, 2006.