Edge Suppression by Gradient Field Transformation using Cross-Projection Tensors

Amit Agrawal
University of Maryland
College Park, MD, 20742
aagrawal@umiacs.umd.edu

Ramesh Raskar
Mitsubishi Electric Research Labs
Cambridge, MA, 02139
raskar@merl.com

Rama Chellappa
University of Maryland
College Park, MD, 20742
rama@cfar.umd.edu

Abstract

We propose a new technique for edge-suppressing operations on images. We introduce cross projection tensors to achieve affine transformations of gradient fields. We use these tensors, for example, to remove edges in one image based on the edge-information in a second image. Traditionally, edge suppression is achieved by setting image gradients to zero based on thresholds. A common application is in the Retinex problem, where the illumination map is recovered by suppressing the reflectance edges, assuming it is slowly varying.

We present a class of problems where edge-suppression can be a useful tool. These problems involve analyzing images of the same scene under variable illumination. Instead of resetting gradients, the key idea in our approach is to derive local tensors using one image and to transform the gradient field of another image using them. Reconstructed image from the modified gradient field shows suppressed edges or textures at the corresponding locations. All operations are local and our approach does not require any global analysis.

We demonstrate the algorithm in the context of several applications such as (a) recovering the foreground layer under varying illumination, (b) estimating intrinsic images in non-Lambertian scenes, (c) removing shadows from color images and obtaining the illumination map, and (d) removing glass reflections.

1. Introduction

Our goal in this paper is to design edge-suppressing operations on images. Image formation depends on shape and reflectance of the objects in the scene and the scene illumination. Scene analysis involves, for example, factoring the image to recover the reflectance or illumination map. In techniques that use local per-pixel operations, a common approach is to preserve (or suppress) image gradients at known locations so that in the recovered map, corresponding edges and textures are preserved (or suppressed). For

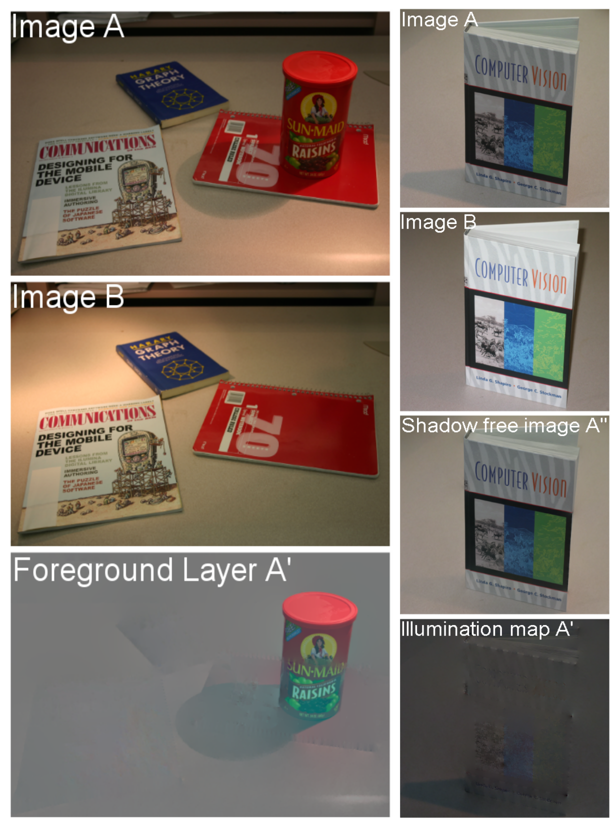

Figure 1. Edge suppression under varying illumination using affine transformation of gradient fields. (Left Column) Two images of a scene captured under different illumination, but with one having a foreground object. By removing edges in ∇A which are present in ∇B, we recover the foreground layer A'. Notice that A' is free of all scene texture edges apart from those due to the box (even inside the box shadow). (Right Column) Ambient and flash images of a book on a table. We remove the edges from ∇A using ∇B to get $\nabla A'$, which is integrated to obtain the illumination map A'. Even though the face of the book is highly textured, A' does not have effects of scene texture. Reconstruction from $\nabla A - \nabla A'$ gives the shadow free image A''.

instance, the Retinex algorithm by Land and McCann [11] assumes reflectance to be piece-wise constant (Mondrian scenes) and illumination to be smooth. Horn [9] proposed to manipulate the image gradient field under these assumptions, by setting large derivatives corresponding to the reflectance edges to zero using thresholds. By integrating the modified gradient field, one can recover the illumination map.

However, a single threshold for the entire image cannot account for illumination and reflectance variations across the image. In this paper, we propose a new method for manipulating image gradient fields based on affine transformation using projection tensors. Our approach provides a principled way of removing scene texture edges from images as compared to thresholding (or zeroing the corresponding gradients). We make no assumptions on ambient lighting, smoothness of the reflectance or the illumination map and do not use explicit shadow masks.

Scene analysis from a single image is a challenging task. We use more than one image under variable illumination for recovering the maps. We show how to remove edges to handle foreground objects, shadows, and glass reflections in an image, using a second image of the same scene taken under different illumination condition. Our approach is based on obtaining the projection tensor from one image and using it to transform the gradient field of the other image. We present techniques that work under natural as well as active illumination variations. For natural illumination, we use an approach proposed by Weiss [26], which uses multiple images for estimating intrinsic images and improve on the estimation of illumination maps. For active illumination, we use the attached flash unit in digital cameras to introduce additional illumination in the scene. These additional images are used to extract reliable information about the scene texture edges, thus avoiding hard thresholds and assumptions on smoothness of reflectance or illumination images.

1.1. Contributions

- We propose a new technique which use *cross projection tensors* derived from local edge structures in one image to suppress edges in a second image.
- We present a class of problems where edge-suppression can be a useful tool. We show applications in traditional problems such as recovering the reflectance or the illumination map and demonstrate usefulness in other problems such as recovering reflection or foreground layers.

1.2. Related work

Intrinsic images were proposed as a useful mid-level scene description by Barrow and Tenenbaum [3]. The observed image is considered to be the product of a reflectance image and an illumination image [9, 26]. Decomposing a given image into intrinsic images is an ill-posed problem. Funt *et al*. [8] extended the Retinex problem to color images, again using thresholds but correcting for the non-zero curl of the modified gradient field. Impressive results were shown by Finlayson *et al*. [7] for removing shadows from a single color image, by projecting the 2D log-chromaticities along an invariant direction. However, their approach requires imaging under Planckian lights (daylight is a close approximation). In addition, they have an explicit shadow mask for zeroing the edges corresponding to shadows. Recently, Weiss [26] proposed to use multiple images of a scene under changing illumination for estimating intrinsic images. A probabilistic approach, based on maximum-likelihood (ML) estimation was proposed in [26], assuming the scene to be Lambertian. However, for **non-Lambertian scenes**, the estimated reflectance image does not accurately represent the scene reflectance and some portion of the scene reflectance will be included in the illumination images [13]. Matsushita *et al*. [13] proposed to remove the scene texture edges from the illumination images using a manually specified threshold. Our approach provides a natural way of removing such effects of scene texture by removing the scene texture edges present in the reflectance image from the intensity images, thus avoiding the thresholding altogether.

Our approach can also be used to remove complex scene structures such as **reflection layers** due to glass. While photographing through glass, flash images (images under flash illumination) usually have undesirable reflections of objects in front of the glass. We show how to recover such reflection layers. Agrawal *et al*. [1] proposed a *gradient projection* technique to remove reflections by taking the projection of the flash image intensity gradient onto the ambient image intensity gradient. We show that the gradient projection algorithm is a special case of our approach, and introduces color artifacts which can be removed by our method. Other methods for reflection removal include changing polarization or focus [14, 19] and Independent Component Analysis (ICA) [6].

Background subtraction is used to segment moving regions in image sequences taken from a static camera [5, 20]. There exists vast literature on background modeling using adaptive/non-adaptive Gaussian mixture models and its variants. See review by Piccardi [16] and references therein. Layer separation in presence of motion has been discussed in [18, 21]. We show how mutual edge-suppression can be effectively used for **foreground extraction** of opaque layers. Our gradient-based approach relies on local structure rather than absolute intensities and can handle significant illumination variations across images.

Local structure tensors and **diffusion tensors** derived from them have been used for spatio-temporal image processing and optical flow [10], and PDE based image reg-

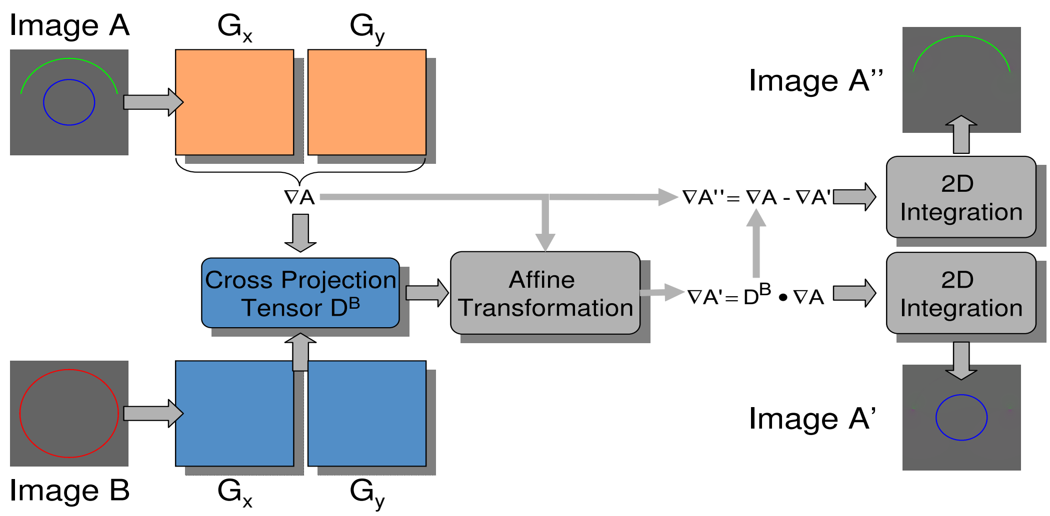

Figure 2. Suppressing edges in image A using image B by affine transformation of gradient field using cross projection tensors. The cross projection tensor D^B is obtained using images. The gradient field ∇A is transformed using D^B to give $\nabla A'$, removing those edges from A which are present in B. Reconstruction from $\nabla A'$ gives image A', with corresponding edges suppressed. Reconstruction from the difference gradient field $(\nabla A - \nabla A')$ gives image A'', which preserves those edges in A which are also present in B.

ularization [2, 22, 23, 25]. These approaches are based on modifying the image intensities using the non-linear diffusion equation

$$I_t = \mathtt{div}(D\nabla I), \tag{1}$$

where $\mathtt{div}$ denotes the divergence operator, ∇I is the image gradient and D denotes the diffusion tensor. In comparison, our approach is a gradient domain approach based on transforming the gradient field ∇I using D. Recently, **gradient domain algorithms** have been used for Poisson image editing [15], day night image fusion [17], and seamless image stitching [12]. Our approach is inspired by these algorithms.

2. Affine transformation on gradient fields

Let $I(x, y)$ be an intensity image and $\nabla I = \begin{bmatrix} g_x \\ g_y \end{bmatrix}$ denote the gradient vector of I at each pixel. The smoothed structure tensor $\mathbf{G}_\sigma$ is defined as [22]

$$\mathbf{G}_\sigma = (\nabla I \nabla I^T) * K_\sigma = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} * K_\sigma, \tag{2}$$

where $*$ denotes convolution and K_σ is a normalized 2D Gaussian kernel of variance σ. The matrix $\mathbf{G}_\sigma$ can be decomposed as

$$\mathbf{G}_\sigma = V\Sigma V^T = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix}, \tag{3}$$

where $\mathbf{v}_1, \mathbf{v}_2$ denote the eigen-vectors corresponding to the eigen-values λ_1, λ_2 respectively and $\lambda_2 \le \lambda_1$. The eigen-values and eigen-vectors of $\mathbf{G}_\sigma$ give information about the local intensity structures in the image [2]. For homogeneous regions, $\lambda_1 = \lambda_2 = 0$. If $\lambda_2 = 0$ and $\lambda_1 > 0$, it signifies the presence of an intensity edge. The eigen-vector $\mathbf{v}_1$ (corresponding to the higher eigen-value λ_1) corresponds to the direction of the edge.

For the problem of image restoration based on diffusion process, Weickert [24, 25] proposed a generalization of the divergence based equation given by (1), where D is a field of *diffusion tensors*. At each pixel, $D(x, y)$ is a 2×2 symmetric, positive definite matrix. Weickert proposed to design the diffusion tensors D by selecting its eigen-vectors $\mathbf{u}_1, \mathbf{u}_2$ and eigen-values μ_1, μ_2 based on the eigen-values and eigen-vectors of $\mathbf{G}_\sigma$. D is then obtained as

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix}.$$

Several designs for obtaining D have been proposed for coherence enhancing diffusion [2, 25], edge enhancing diffusion [2], color image restoration, in-painting, and magnification [23]. Usually, D is obtained from the given image I. All these approaches modify the image intensities using the diffusion equation (1). In this paper, we show how to obtain *projection tensors* and discuss the properties and applications of affine transformation of the gradient field ∇I of an image using them.

2.1. Self-Projection Tensors

We first discuss how to remove edges from a single image by estimating projection tensors from the image itself. The idea is to project the image gradient vector onto its own orthogonal direction and hence the name self-projection tensors. This analysis will lead us to our main idea of cross-projection tensors: to estimate these tensors from a second image and apply them to the given image to suppress edges.

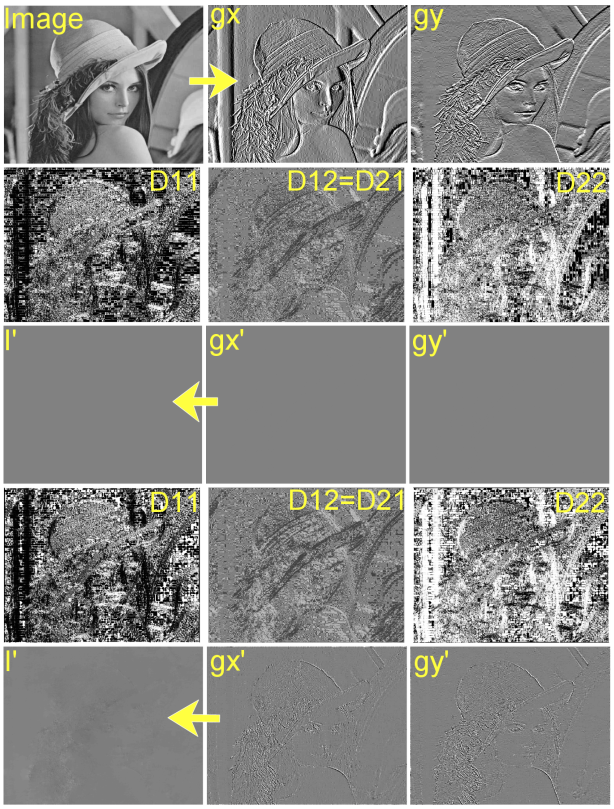

Figure 3. Affine transformation on image gradient field using D^{self} for different σ. (Top row) Lena image and the corresponding g_x and g_y. (Second row) Components D_{11}, D_{12} and D_{22} of D^{self} with $\sigma = 0$. (Third row) Transformed gradients g'_x, g'_y, and the image I' reconstructed from them. g'_x, g'_y and I' are zero all over. (Last two rows) Components of the projection tensor, modified gradient field and the reconstructed image corresponding to D^{self} using $\sigma = 0.5$. Even if $\sigma > 0$, all dominant edges are removed. A non-zero σ incorporates spatial information over the neighborhood for better estimation of cross projection tensors in presence of noise. D_{11} and D_{22} are shown between $[0, 1]$. D_{12} is shown between $[-1, 1]$.

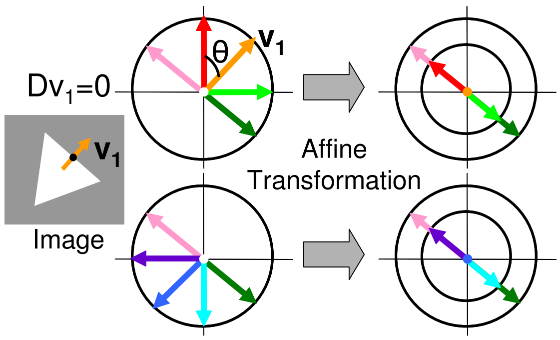

Figure 4. Visualizing affine transformation on gradient vectors. At each pixel in an image, $\mathbf{v}_1$ corresponds to the direction of the dominant edge. After affine transformation using D^{self}, any vector gets projected to the direction orthogonal to the local gradient vector $\mathbf{v}_1$.

2.1.1 Gradient Projection

In [1], Agrawal *et al.* proposed the technique of gradient projection (GP) to remove artifacts from flash image using a no-flash ambient image. They project the flash image gradient onto the direction of the ambient image gradient to remove spurious edges from flash image due to glass reflections. They use the idea that the direction of the image gradient remains stable under illumination changes [4]. We first show that taking a projection can also be defined by an affine transformation of the gradient field.

As discussed in previous section, the eigen-vector $\mathbf{v}_1$ of the structure tensor matrix $\mathbf{G}_\sigma$ correspond to the direction of the edge. Suppose we define the self-projection tensor D^{self} as

$$\begin{aligned} &\mathbf{u}_1 = \mathbf{v}_1 \quad \mathbf{u}_2 = \mathbf{v}_2, \quad \mu_1 = 0 \quad \mu_2 = 1, \\ &D^{self} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} \end{aligned} \tag{4}$$

It is easy to see that an affine transformation of the image gradient using D^{self} will remove the local edge.

$$\begin{aligned} D^{self}\mathbf{v}_1 &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} \mathbf{v}_1 \\ &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{5}$$

Figure 4 shows the effect of transforming gradient vectors using D^{self}. All vectors are projected to the direction *orthogonal* to the local gradient vector $\mathbf{v}_1$. Thus, we can establish the following relationship.

Transforming a vector using D^{self} is equivalent to projecting on the orthogonal direction of the local gradient vector.

Figure 3 shows the reconstructed images obtained by integrating the transformed gradient field of the Lena image using D^{self}. To handle noise, it is useful to have a larger spatial support by using $\sigma > 0$ for reliable estimation of the direction of the local edge. In that case, although the estimated $\mathbf{v}_1$ may not lie in the null space of D^{self}, the affine transformation can still remove the dominant edges from the gradient field ∇I. Figure 3 shows that with $\sigma = 0$, the reconstructed image I' is zero everywhere. With $\sigma = 0.5$, I' has most of its edges removed.

The gradient projection approach as described in [1] cannot handle homogeneous regions and introduces color artifacts (see Figure 6). This is because it does not include neighborhood support for gradient direction estimation, which is unstable in presence of noise and low frequency regions. In addition, the projection is done for each channel separately which leads to color artifacts. In the next section, we show how to estimate cross-projection tensors. Our approach combines information spatially (using $\sigma > 0$)

and across channels to handle noise and have no color artifacts.

3. Cross-Projection Tensors

We now show how to remove the scene texture edges from an image by transforming its gradient field using cross projection tensors obtained from a second image of the same scene (see Figure 2). The final image is obtained by a 2D integration of the modified gradient field.

Let A and B denote the two images. Let $\mathbf{G}_\sigma^A$ and $\mathbf{G}_\sigma^B$ denote the smoothed structure tensors for image A and B respectively. The eigen-values and eigen-vectors of $\mathbf{G}_\sigma^A$ and $\mathbf{G}_\sigma^B$ will be denoted by superscripts A and B respectively. The technique for obtaining the cross projection tensor D^B is explained now. Note that by transforming ∇A with D^B, we wish to (a) remove all edges from A which are present in B, and (b) retain all edges in A which are not in B. To obtain D^B, we propose the following rules:

- $\mathbf{u}_1 = \mathbf{v}_1^B, \quad \mathbf{u}_2 = \mathbf{v}_2^B$.
- If B is homogeneous ($\lambda_1^B = 0$)
  - If A is also homogeneous ($\lambda_1^A = 0$), set $\mu_1 = \mu_2 = 0$. This results in $D(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ for that pixel.
  - If A is not homogeneous ($\lambda_1^A > 0$), set $\mu_1 = \mu_2 = 1$. This results in $D(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and edges which are in A but not in B can be retained.
- Else, if there is an edge in B ($\lambda_1^B > 0$), remove that edge by setting $\mu_1 = 0$, $\mu_2 = 1$.

In practice, due to noise and gradient estimation using finite differences, a small non-zero value (*e.g.*, 1) is used as a threshold to decide for homogeneity. One might think that the above homogeneity threshold needs to vary across the image, if the image has spatially varying illumination. Since we take into account the direction of the edge, we do not need spatially adaptive thresholds. Figure 1 shows such an example. Also note that there are no other thresholds in our scheme.

3.1. Combining information across color channels

The above formulation can be used for gray scale images. A naive way of handling color images would be to obtain the cross projection tensor for each channel and transform the gradient field in each channel separately. However, this scheme introduces color artifacts in the final reconstructed image as the projection tensor does not utilize information across channels. To this end, we obtain a common cross projection tensor for all channels by estimating a common $\mathbf{G}_\sigma$ matrix as [22]

$$\mathbf{G}_\sigma = (\sum_{i=1}^{3}(\nabla I_i \nabla I_i^T)) * K_\sigma, \tag{6}$$

where i denote the color channel.

4. Applications

We show applications on recovering the foreground layer under varying illumination, estimating intrinsic images for non-Lambertian scenes, removing shadows from color images, recovering the illumination map, and removing glass reflections from images[1]. We use $\sigma = 0.4$ in all the experiments.

4.1. Recovering the foreground layer under varying illumination

Background subtraction and foreground layer recovery is a challenging problem in the presence of significant illumination variations. Consider the pair of images in the first column of Figure 1. Image A was captured with a foreground object (raisin box) illuminated from a table lamp on the right. Image B was captured with the table lamp on the left, but without the object. Notice the spatially non-uniform illumination in the images. Intensity based measures such as frame differencing cannot discount such illumination variations across images. Using normalized cross-correlation can handle varying illumination, but only in textured regions.

We compute the cross projection tensor D^B at each pixel using the background image B and transform the gradient field ∇A using D^B to obtain $\nabla A'$. This suppresses all the texture edges corresponding to the background. The foreground layer is obtained by integrating $\nabla A'$. The recovered foreground layer is free of the background texture, even inside the shadow of the box. Notice that a part of the foreground (red box) is similar in color to the background (red book). A color based differencing approach will fail at such regions. In addition, homogeneous regions on the foreground objects will leave holes for any local pixel intensity based approach. Our method is able to "fill-in" such regions by propagating information from edges during the integration of the modified gradient field. However, edges of the foreground object which align with the background edges (i.e., share the same gradient vector direction) are treated as part of the background and suppressed. Fortunately, as is well known, the likelihood of alignment of 1D features, such as edges, on two different objects is low. Nevertheless, in overlapping high frequency regions the likelihood is increased and some foreground edges may be lost. Notice how the top of the text "SUN-MAID" on the red box is smeared in A', as it overlaps with the red book binding in the image B.

[1] Matlab code and images are available at http://www.umiacs.umd.edu/~aagrawal/

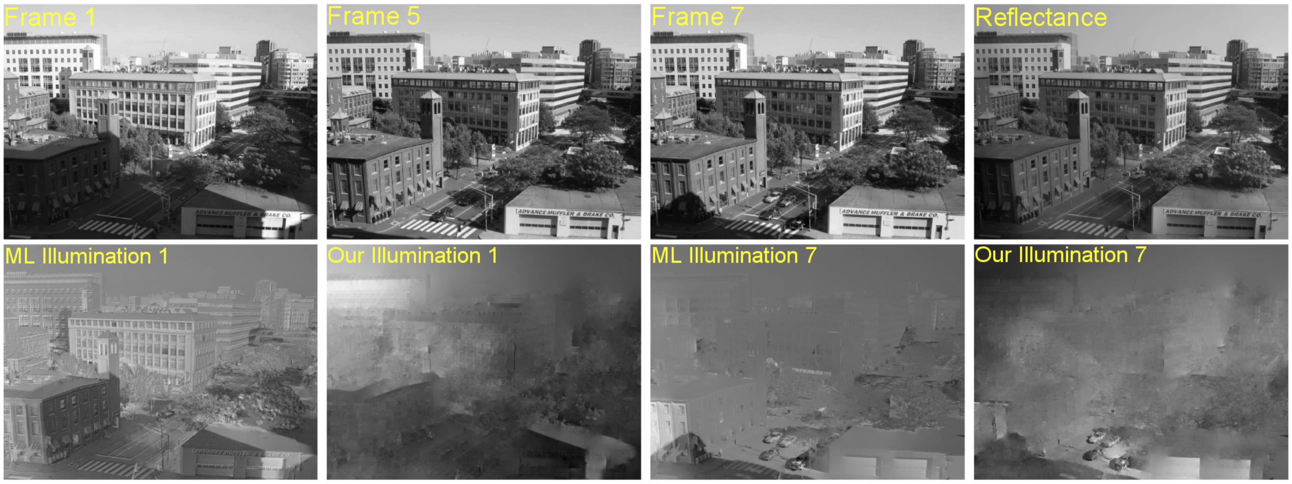

Figure 5. Recovering intrinsic images for an outdoor non-Lambertian scene. (Top row) Input images of an outdoor scene taken at different times of the day and the ML reflectance image. (Bottom row) Estimated illumination images using ML estimation and our approach. The scene texture edges (white stripes on the road) are visible in the ML illumination images. These are removed in our result while all the shadows are maintained. However, we make the usual assumption that illumination and reflectance edges do not coincide. All such illumination edges cannot be recovered. All illumination images have been shown with logarithmic non-linearity following [26].

4.2. Recovering illumination images in non-Lambertian scenes

Weiss [26] proposed to decompose a set of N intensity images $I(x,y,t)_{t=0}^{N-1}$ obtained from a fixed view-point under changing illumination into a single reflectance image $R(x,y)$ and the corresponding illumination images $L(x,y,t)$ as:

$$I(x,y,t) = R(x,y)L(x,y,t) \quad t = 0 \ldots N-1. \tag{7}$$

Taking the logarithm of both sides, we get

$$i(x,y,t) = r(x,y) + l(x,y,t) \quad t = 0 \ldots N-1. \tag{8}$$

The method in [26] uses a prior that when derivatives filters f_n are applied to l, the output tends to be sparse. Assuming the filter outputs to be Laplacian distributed, the maximum-likelihood (ML) estimate of the filtered reflectance image $\widehat{r_n} = r * f_n$ is given by the median of the filtered images $i_n = i * f_n$ along the temporal axis. The filtered illumination images l_n can then be obtained as

$$l_n(x,y,t) = i_n(x,y,t) - \widehat{r_n}(x,y) \quad t = 0 \ldots N-1. \tag{9}$$

However, if the scene is not Lambertian, l_n will have some effect of scene texture edges. Matsushita *et al.* [13] proposed to remove the scene texture edges from l_n using a threshold T by setting

$$l_n(x,y,t) = \begin{cases} 0 & \text{if } |\widehat{r_n}(x,y)| > T, \\ l_n(x,y,t) & \text{otherwise.} \end{cases} \tag{10}$$

However, the threshold was manually specified in [13] and is difficult to generalize to different scenes. Our approach provides an elegant way of estimating the illumination images l by avoiding the two-step process which involves thresholding. We first estimate r using Weiss's method. For each image i, we then find the cross projection tensor D^r using r and i, and transform the gradient field ∇i using D^r. This will remove all the edges from i which are present in r. Thus,

$$\nabla l(x,y,t) = D^r \cdot \nabla i(x,y,t). \tag{11}$$

The illumination images $l(x,y,t)$ are obtained by integrating the resulting gradient field $\nabla l(x,y,t)$ for each t.

Figure 5 shows results on images of an outdoor scene taken under different times of the day. Notice that the ML illumination images contains the effect of scene texture, especially white lines on the road surface. Using our approach, all such scene texture edges can be successfully removed from the illumination images while preserving shadows.

4.3. Removing shadows from color images

We use a flash image F of the scene to remove shadows from the ambient (no-flash) image A. The flash and the ambient images were captured in quick succession using the remote capture utility with the camera mounted on a tripod. We obtain the cross projection tensor D^F using F and transform the gradient field ∇A using it. Figure 1 shows an example on a highly textured book. Notice that the recovered shadow free image A'' has no color artifacts and the recovered illumination map A' is free of strong texture edges on the face of the book. Figure 6 shows a challenging scenario where the hat on the mannequin casts shadows on the mannequin's face and neck. Usually, the ambient and flash images have different color tone due to ambient lighting being yellow-reddish and flash illumination being bluish. Our algorithm requires no pre-processing or color calibration and

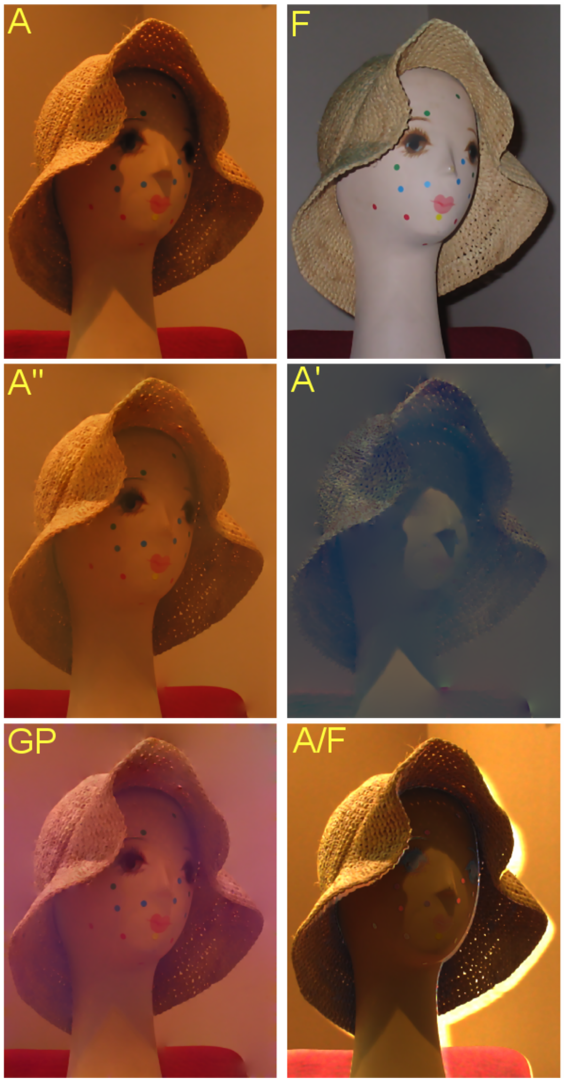

Figure 6. Removing cast shadows. (Top row) Ambient and flash images of a mannequin. The hat casts shadows on the mannequin's face and neck in the ambient image A. The flash image F is taken with a short exposure time. (Second row) Recovered shadow free image A'' and the illumination map A'. (Last row) Result using gradient projection has visible color artifacts. One cannot obtain the illumination map by taking the ratio A/F (shown on right) which is confounded by shadows due to flash and uneven lighting on the hat. Notice that the white balance in the flash and ambient images are different. Our result does not have any color artifacts.

has no color artifacts as compared to the result using gradient projection. One might think that the ratio image $\frac{A}{F}$ could give the illumination map of the scene. However, the ratio image (shown in Figure 6) does not represent the illumination map due to the effects of flash shadows (at depth discontinuities) and lighting variations (on top of the hat) due to the flash. The illumination map obtained by our approach better represents the diffuse ambient illumination.

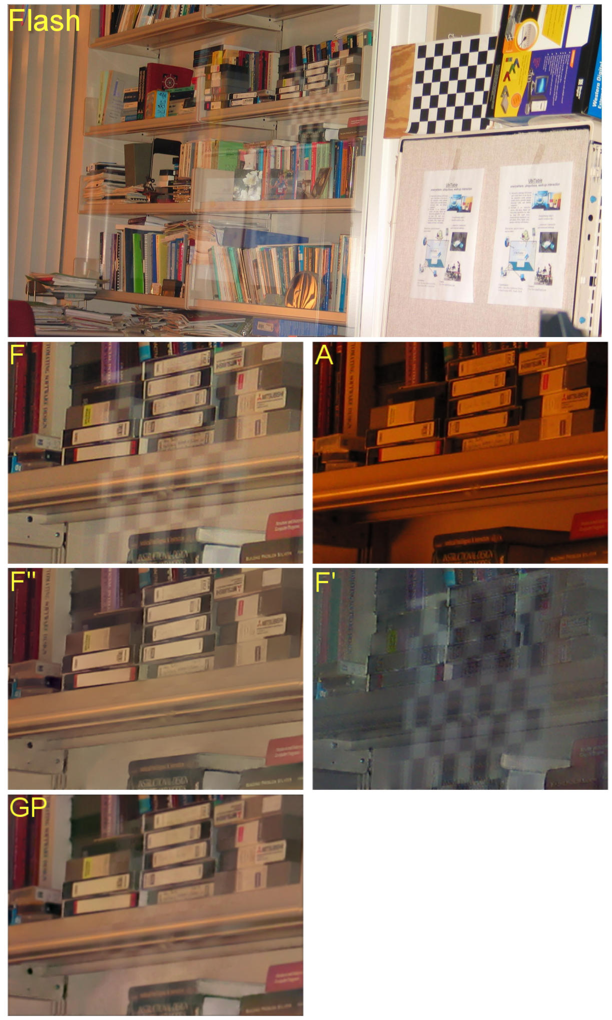

Figure 7. Removing glass reflections from a flash image using an ambient image. (Top row) Flash image F of an office scene through a glass window. The checkerboard outside the office results in reflections on the glass window. (Second row) Zoomed in flash and ambient images. (Third row) Recovered reflection layer F' and the reflection free image F''. (Last row) Result using gradient projection has a slight tinge of the reflection layer remaining along with a brownish hue (on top of books in the lower shelf).

4.4. Removing glass reflections

While photographing through glass in low light environments, an ambient image is usually of low quality and has low contrast. Using a flash improves the contrast, but it may result in reflections of objects in front of the glass. Figure 7 shows such an example, where the camera is looking into an office scene through a glass window. The flash image has undesirable reflections of the checkerboard outside the glass window. We use the ambient image A to obtain the cross projection tensor D^A and transform the gradient field ∇F of the flash image F using it. The reflection layer is obtained by integrating $\nabla F'$ and the reflection free flash image is obtained by integrating $\nabla F - \nabla F'$. For this example, we repeat the affine transformation 5 times as the reflection layer has strong edges. In comparison, one can see a slight tinge of reflection remaining in the gradient projection result.

5. Conclusions

We have presented an approach for edge-suppressing operations on an image, based on affine transformation of gradient fields using cross projection tensor derived from another image. Our approach is local and requires no global analysis. In recovering the illumination map, we make the usual assumption that the scene texture edges do not coincide with the illumination edges. Hence, all such illumination edges cannot be recovered. Similarly, while extracting foreground layer, edges of the foreground object which exactly align with the background edges cannot be recovered. This may be handled by incorporating additional global information in designing the cross projection tensors, which remains an area of future work. In addition, image saturation, specular objects, and black objects will create problems due to the lack of reliable information. We used a fixed variance σ for estimating the structure tensor $\mathbf{G}_\sigma$, but an adaptive neighborhood scheme might improve results.

We showed applications on extracting foreground layer, removing shadows and glass reflections from images, recovering the illumination map, and estimating intrinsic images in non-Lambertian scenes. Our approach is conceptually simple and can easily handle color images without the need for any color calibration or white balancing. We are optimistic that this framework can be used for edge manipulations that go beyond edge suppression and preservation and for applications in image understanding, editing and special effects.

References

[1] A. Agrawal, R. Raskar, S. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Trans. Graph.*, 24(3):828–835, 2005.

[2] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002.

[3] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*, pages 3–26, 1978.

[4] H. Chen, P. Belhumeur, and D. Jacobs. In search of illumination invariants. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 254–261, 2000.

[5] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. In *Proc. European Conf. Computer Vision*, pages 751–767, 2000.

[6] H. Farid and E. Adelson. Separating reflections and lighting using independent components analysis. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 1262–1267, June 1999.

[7] G. D. Finlayson, S. D. Hordley, and M. S. Drew. Removing shadows from images. In *ECCV*, pages 823–836, 2002.

[8] B. Funt, M. Drew, and M. Brockington. Recovering shading from color images. In *ECCV*, pages 124–132, May 1992.

[9] B. Horn. Determining lightness from an image. *Comput. Graphics, Image Processing*, 3:277–299, 1974.

[10] B. Jahne. *Spatio-Temporal Image Processing, Theory and Scientific Applications*, volume 751 of *Lecture Notes in Computer Vision*. Springer-Verlag, Berlin, Germany, 1993.

[11] E. Land and J. McCann. Lightness and retinex theory. *J. Opt. Soc. of America*, 61:1–11, 1971.

[12] A. Levin, A. Zomet, S. Peleg, and Y. Weiss. Seamless image stitching in the gradient domain. In *ECCV (4)*, pages 377–389, 2004.

[13] Y. Matsushita, K. Nishino, K. Ikeuchi, and M. Sakauchi. Illumination normalization with time-dependent intrinsic images for video surveillance. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(10):1336–1347, 2004.

[14] S. Nayar, X. Fang, and T. Boult. Separation of reflection components using color and polarization. *Int'l J. Computer Vision*, 21(3):163–186, Feb. 1997.

[15] P. Perez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Trans. Graph.*, 22(3):313–318, 2003.

[16] M. Piccardi. Background subtraction techniques: a review. In *Proc. IEEE SMC Intl. Conf. Systems, Man and Cybernetics*, Oct. 2004.

[17] R. Raskar, A. Ilie, and J. Yu. Image Fusion for Context Enhancement and Video Surrealism. In *Proc. of NPAR*, 2004.

[18] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *ECCV (4)*, pages 328–341, 2004.

[19] Y. Schechner, N. Kiryati, and R. Basri. Separation of transparent layers using focus. *Int'l J. Computer Vision*, 39(1):25–39, Aug. 2000.

[20] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In *Proc. Conf. Computer Vision and Pattern Recognition*, volume 2, page 252, 1999.

[21] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 246–243, June 2000.

[22] D. Tschumperle. *PDE's Based Regularization of Multivalued Images and Applications*. PhD thesis, Universite de Nice - Sophia Antipolis, dec 2002.

[23] D. Tschumperle and R. Deriche. Vector-valued image regularization with pdes : A common framework for different applications. In *Proc. Conf. Computer Vision and Pattern Recognition*, volume 1, 2003.

[24] J. Weickert. A review of nonlinear diffusion filtering. In *Scale-Space Theory in Computer Vision, Lecture Notes in Comp. Science (Springer, Berlin)*, 1997.

[25] J. Weickert. Coherence-enhancing diffusion filtering. *Int'l J. Computer Vision*, 31:111–127, 1999.

[26] Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV*, pages 68–75, 2001.