

Layer Extraction from Superimposed Image

Husna K, Abdul Rias P

Abstract— An approach for removing reflection and distortion from a single image. This will effect in the quality of image. The distortion will happen due to the position of the camera relative to the subject, so the image will be distorted or blurred with noise. Photographs taken through glass windows often contain both the desired scene and undesired reflections. Decomposing of the single input image from the reflection is ill posed problem. Here the input single image contains reflection layer with distortion. We have proposed a new, simple Gaussian Mixture model with EPLL which performs surprisingly well on image denoising, deblurring.

Index Terms— Gaussian Mixture model, Reflection Removal

I. INTRODUCTION

A very large portion of digital image processing is devoted to image restoration. This includes research in algorithm development and routine goal oriented image processing. Image restoration is the removal or reduction of degradations that are incurred while the image is being obtained. Degradation comes from blurring as well as noise due to electronic and photometric sources. Blurring is a form of bandwidth reduction of the image caused by the imperfect image formation process such as relative motion between the camera and the original scene or by an optical system that is out of focus. When aerial photographs are produced for remote sensing purposes, blurs are introduced by atmospheric turbulence, aberrations in the optical system and relative motion between camera and ground. In addition to these blurring effects, the recorded image is corrupted by noises too. A noise is introduced in the transmission medium due to a noisy channel, errors during the measurement process and during quantization of the data for digital storage. Each element in the imaging chain such as lenses, film, digitizer, etc. Contribute to the degradation. The Image denoising is often used in the field of photography or publishing where an image was somehow degraded but needs to be improved before it can be printed. For this type of application we need to know something about the degradation process in order to develop a model for it. When we have a model for the degradation process, the inverse process can be applied to the image to restore it back to the original form.

Blur is a result of imperfect image formation process. It may be either due to the relative motion between the camera

and the object or due to an out of focus optical system. Based on this blur is classified into motion blur and out of focus blur.

Motion blur mainly occurs in a dim lighting environment where long exposure time is required. Motion blur can be modeled as the convolution of the sharp image u with the blur kernel k or point spread function (PSF). PSF refers to the extent to which an image of a point source is blurred by the motion blur.

Image processing is a technique to enhance the image and to remove artifacts from images. It is used in many fields like medical, textiles, military, printing industry etc. If it is taken through a glass window, then the picture is viewed through transparent glass or a photograph the image consist of two parts, first one is the real image of the scene beyond the glass and second is virtual image of the scene reflected by the glass window. Decomposing of the single input image from the reflection is a massive ill posed problem. The distortion will happen due to the position of the camera relative to the subject, so the image will be distorted or blurred with noise. Similarly photographs taken under low-light conditions also produce a variety of undesirable effects and artifacts. They tend to saturate nearby objects while failing to light up distant ones. Since the flash intensity falls with distance from the camera, flash produces a tunnel effect, where brightness decreases quickly with depth. Furthermore, flash is widely known for producing undesirable reflections. Direct reflection of the flash itself is caused by glossy objects in the scene. Because of the artifacts, it causes an error in image processing, due to the absence of additional knowledge about the scene. It becomes necessary to remove artifacts, before processing the image in artifacts removal process. Flash and no flash images are used to produce better flash images. Gradient projection method is used to remove reflection and highlights from an image. A gradient orientation coherence model, relates gradients in the flash and ambient images, and tries to capture the properties of image gradients that remain invariant under the change of lighting that takes place between a flash and an ambient image. Based on a gradient projection method it is possible to remove the component of image gradients that are introduced by undesirable reflections. Gradient coherence model is used as a guide, to combine the gradients of flashed image and without flashed image. Here, we address the reflection removal problem using “ghosting” effects and distortion separation – multiple reflections on glasses in the captured image. A common example is a double-pane window, which consists of two thin

glass panes separated by some distance for insulation. The glass pane at the inner side (closer to the camera) generates the first reflection, and the outer side generates the second, which is a shifted and attenuated version of the first reflection. The distance between the two reflections depends on the space between the two panes. In single-pane windows of typical thickness 3-10mm, ghosting arises from multiple reflections by the near and far surfaces of the glass.

II. LITERATURE SURVEY

Tianfan Xue, Michael Rubinstein [1] have presented a unified computational approach for taking photos through reflecting or occluding elements such as windows and fences. Rather than capturing a single image, they instruct the user to take a short image sequence while slightly moving the camera. Differences that often exist in the relative position of the background and the obstructing elements from the camera allow us to separate them based on their motions, and to recover the desired background scene as if the visual obstructions were not there.

Yi Chang Shih, Dilip Krishnan [2] proposed a method for removing reflections from single image. Takes in a single image input and searches for “ghosting” effects, multiple reflections from the window in the captured image.

Xiaojie Guo, Xiaochun Cao, and Yi Ma [3] have presented a method to separate the two layers from multiple images, which exploits the correlation of the transmitted layer across multiple images, and the sparsity and independence of the gradient fields of the two layers. A novel Augmented Lagrangian Multiplier based algorithm is designed to efficiently and effectively solve the decomposition problem.

Cheng Lu and Mark S. Drew [4] have developed a method to estimate ambient illuminants using no-flash/flash image pairs. Accurate estimation of the ambient illuminant is useful for imaging applications. To estimate the scene illumination, a version of the “illuminating illumination” method suggested by Dicarlo et al. is used. The method introduces camera flash light into the scene, and the reflected light is used to estimate the ambient illuminant. The original method needs an extra step of estimating the object surface reflectance, using a 3-dimensional linear surface model and the knowledge of the spectral responsivities of camera sensors. Here we consider the problem of estimating the ambient illuminant directly, with only flash/no-flash pairs, without information on surface reflectance and camera sensors. First, the flash image is registered with the no-flash image: the difference between the two gives a pure-flash image, as if it were taken under flash only. The no-flash and pure-flash images are represented by a physically-based model of image formation which uses assumptions of Lambertian surfaces, Planckian lights, and narrowband camera sensors.

N. Kong, Y. Tai, and J. Shin [5] have proposed a physically-based approach to separate reflection using multiple polarized images with a background scene captured behind glass. The input consists of three polarized images; each captured from the same view point but with a different

polarizer angle separated by 45 degrees. The output is the separation of the reflection and background layers from each of the input images. A main technical challenge for this problem is that the mixing coefficient for the reflection and background layers depends on the angle of incidence and the orientation of the plane of incidence, which are spatially varying over the pixels of an image. Exploiting physical properties of polarization for a double-surfaced glass medium, propose a multiscale scheme which automatically finds the optimal separation of the reflection and background layers.

Y. Li and M. S. Brown [6] have developed a method of extracting two layers from an image where one layer is smoother than the other. Layer decomposition from a single-image is inherently ill-posed and solutions require additional constraints to be enforced. The method introduces a novel strategy that regularizes the gradients of the two layers such that one has a long tail distribution and the other a short tail distribution. While imposing the long tail distribution is a common practice, the introduction of the short tail distribution on the second layer is unique. It describes an optimization scheme to solve this regularization with only a few iterations.

Qing Yan, Yi Xu [7] proposed a method, which based on the prior knowledge that edges of weak reflection are always smoother than most edges of observed objects. To filter out the edges of weak reflection, an MRF-EM (Markov Random Field and Expectation Maximization) framework is proposed. In the MRF model, a data energy function is established based on the edge smoothness metric GPS (Gradient Profile Sharpness) and a spatial smoothness energy function are formulated using a weighted Potts model. Moreover, the parameters in the data energy function are updated using the EM algorithm.

Yu Li [8] When the picture is taken which is behind the glass surface, reflections are occurs. In this paper SIFT (scale invariant feature transforms) flow is used to produce good results. For this approach images are taken with a slightly different point of view. From this set of images they got minor changes in the reflection. SIFT flow is used to align the images for pixel wise comparison across input set. Gradients with variation across the image set are assumed to belong to the reflected scenes while constant gradients are assumed to belong to the desired background scene. Then by giving appropriate labels to the gradients which belongs to reflection and background. Reflection interference is taken separated from background scene.

Yilong Geng [9] proposes phone app to take two images, specially designed for smart phones. In this paper flash and ambient image is used. First find out position of flash and size of a flash in the flash image. To find out hot spot in flash image template matching and connected component method is used. Input images are taken from different angles so before combining them there is need to align them together. For this operation SIFT descriptor is used with RANSC. Then color transformation resolves the color inconsistency in two pictures caused by different lighting sources.

J. Kopf, F. Langguth, D. Scharstein, R. Szeliski, and M.

Goesele [10] have presented an image-based rendering algorithm for handling complex scenes that may include reflective surfaces. The key contribution lies in treating the problem in the gradient domain. This method uses a standard technique to estimate scene depth, but assign depths to image gradients rather than pixels. A novel view is obtained by rendering the horizontal and vertical gradients, from which the final result is reconstructed through Poisson integration using an approximate solution as a data term. This algorithm is able to handle general scenes including reflections and similar effects without explicitly separating the scene into reflective and transmissive parts.

B.himabindu [11] proposes a technique for removal of shadow and reflections in the images. In this paper cross projection tensor technique is used edge suppression with affine transformation on gradient fields. Affine transformation is a linear mapping method that preserves points, straight lines and planes. Sets of parallel lines remain parallel after an affine transformation. Cross projection tensor technique remove the scene texture edges of an image by transforming the gradient field. Flash and ambient image is used. Cross projection tensor is obtain from flash image and transform the gradient field of ambient image in it. Here no need for color calibration to handle color images.

K. Gai [12] has presented a method of blind separation of multiple source layers from their linear mixtures with unknown mixing coefficients and unknown layer motions. Such mixtures can occur when one takes photos through a transparent medium, like a window glass, and the camera or the medium moves between snapshots. To understand how to achieve correct separation, we study the statistics of natural images in the Labelme data set. It not only confirms the well-known sparsity of image gradients, but also discovers new joint behavior patterns of image gradients. Based on these statistical properties, he develops a sparse blind separation algorithm to estimate both layer motions and linear mixing coefficients and then recover all layers. This method can handle general parameterized motions, including translations, scalings, rotations, and other transformations. In addition, the number of layers is automatically identified, and all layers can be recovered, even in the underdetermined case where mixtures are fewer than layers.

S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski [13] have developed a system for image-based modeling and rendering of real-world scenes containing reflective and glossy surfaces. Previous approaches to image-based rendering assume that the scene can be approximated by 3D proxies that enable view interpolation using traditional back-to-front or z-buffer compositing. In this work, they introduce how these can be generalized to multiple layers that are combined in an additive fashion to model the reflection and transmission of light that occurs at specular surfaces such as glass and glossy materials. To simplify the analysis and rendering stages, create a model the world using piecewise-planar layers combined using both additive and opaque mixing of light. In this paper, introduce techniques for estimating multiple depths in the scene and

separating the reflection and transmission components into different layers.

Song,Bo; Gong,shenwen; Ren,chunjian [14] used only single image to remove artifacts from image. Due to the image de-convolution image ringing artifacts arises. Image de-convolution algorithm is used for the motion blurred images. Poisson interpolation is used to remove artifacts from blurred image. In this paper only one image is used known as source image. As gradient projection method is used it needs another image which is taken from source image by de-blurring it known as blurred image. Then gradient projection method is used to adjust the gradients with source image. Finally Poisson equation is used to reconstruct the image.

III. PROPOSED SYSTEM

In this approach we are removing the reflection and distortion from a single image using two methods. The methods used for this purpose are reflection separation method and distortion Separation method. The issue of this reflection and distortion separation arises naturally in our everyday life when a desired scene contains another scene reflected off a transparent or semi reflective medium. One of the common examples for this is when we take photographs through windows or taking the images of an object which are placed inside a glass. When we take such kind of images there will be reflection occurs. Sometimes the camera may be shaking when we take photographs so the image will be distorted or blurred with noise. Figure 1 shows block diagram of proposed system.

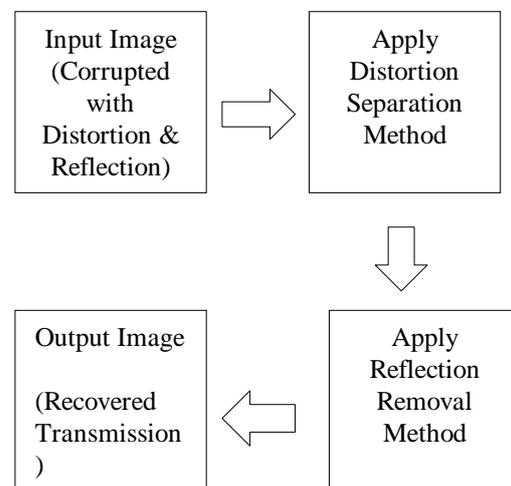


Fig.1. Block diagram of Proposed System

A. Distortion Separation Method

The basic idea behind our method is to try to maximize the Expected Patch Log Likelihood (EPLL) while still being close to the corrupted image in a way which is dependent on the corruption model. Given an image x (in vectorized form) we define the EPLL under prior p as

$$EPLL_p(x) = \sum_i \log p(P(X)) \quad (1)$$

Where P_i is a matrix which extracts the i -th patch from the image (in vectorized form) out of all overlapping patches, while $\log p(P_i x)$ is the likelihood of the i -th patch under the prior p . Assuming a patch location in the image is chosen uniformly at random, EPLL is the expected log likelihood of a patch in the image.

Let $\|Ax - y\|^2$ be the corruption model on the image where x is a vectorized image, A defines the corruption model and y is the vectorized noisy observation. The cost we propose to minimize in order to find the reconstructed image using the patch prior p is

$$f_p(x|y) = \lambda/2(\|Ax - y\|^2) - EPLL_p(x) \quad (2)$$

In corruption model, here it is quite general for denoising, and deblurring. The choice of the matrix A is determined by the application, for denoising- A is an identity matrix, and λ is the noise precision, for deblurring- A is a convolution matrix with a known kernel.

Direct optimization of this cost function is very hard, so an alternative method called half quadratic splitting is used, we introduce a set of patches $\{z^i\}_1^N$, one for each overlapping patch $P_i x$ in the image, yielding the following cost function ,

$$c_{p,\beta}(x, \{z^i\} | y) = \lambda/2(\|Ax - y\|^2) + \sum_i \frac{\beta}{2} (\|P_i x - z^i\|^2) - \log p(z^i) \quad (3)$$

Note that as $\beta \rightarrow \infty$ we restrict the patches $P_i x$ to be equal to the auxiliary variables z^i and the solutions of Equation 3 and Equation 2 converge. For a fixed value of β , optimizing Equation 3 can be done in an iterative manner, first solving for x while keeping z^i constant, then solving for z^i given the newly found x and keeping it constant.

Optimizing Equation 3 for a fixed β value requires two steps, (1) Solving for x given $\{z^i\}$, This can be solved in closed form. Taking the derivative of 3 w.r.t to the vector x , setting to 0 and solving the resulting equation yields,

$$\hat{x} = (\lambda A^T A + \beta \sum_j P_j^T P_j)^{-1} (\lambda A^T y + \beta \sum_j P_j^T z^j) \quad (4)$$

Where the sum over j is for all overlapping patches in the image and all the corresponding auxiliary variables $\{z^i\}$. (2) Solving for $\{z^i\}$, given x —The exact solution to this depends on the prior p in use - but for any prior it means solving a MAP problem of estimating the most likely patch under the prior, given the corrupted measurement $P_i x$ and parameter β . We repeat the process for several iterations, solve for Z given x and solve for x given the new Z , both given the current value of β . Then, increase β and continue to the next iteration. These two steps improve the cost $c_{p,\beta}$ from Equation 3, and for large β we also improve the original cost function f_p from Equation 2. We note that it is not necessary to find the optimum of each of the above steps, any approximate method (such as an approximate MAP estimation procedure) which still improves the cost of each sub-problem will still optimize the original cost function.

The choice of β values by two approaches - the first is optimizing the values on a set of training images. The second option, which is relevant in denoising, is to try to estimate β

from the current image estimate at every step, this is done by estimating the amount of noise σ present in the image \hat{x} , and setting $\beta = 1/\sigma^2$. We use weiss noise estimation method.

In denoising, we have additive white Gaussian noise corrupting the image, so we set the matrix A from Equation 4 to be the identity matrix A , and set λ to be related to the standard deviation of the noise ($=\frac{1}{\sigma^2}$). This means that the solution for x at each optimization step is just a weighted average between the noisy image y and the average of pixels as they appear in the auxiliary overlapping patches. The solution for Z is just a MAP estimate with prior p and noise level $\sqrt{1/\beta}$. If we initialize x with the noisy image y , setting $\lambda = 0$ and $\beta = \frac{1}{\sigma^2}$ results in simple patch averaging when iterating a single step. The big difference, however, is that in our method, because we iterate the solution and $\lambda \neq 0$, at each iteration we use the current estimated image, averaging it with the noisy one and obtaining a new set of Z patches, solving for them and then obtaining a new estimate for x , repeating the process, while increasing β . For image deblurring (non-blind) A is a convolution matrix with a known kernel. We learn a finite Gaussian mixture model over the pixels of natural image patches. GMM model performance with EPLL give better image restoration of captured image.

B. Layer Separation Method

Consider taking an image through a double-pane window which consists of two thin glass panes separated by some distance. The glass pane closer to the camera generates the first reflection, and the opposite side generates the second, which is a shifted and attenuated version of the first reflection. The distance between the two reflections depends on the space between the two panes. In single-pane windows of typical thickness 3-10mm, ghosting arises from multiple reflections by the near and far surfaces of the glass. Ghosting provides a critical cue to separate the reflection and transmission layers, since it breaks the symmetry between the two layers. Here, model the ghosting as convolution of the reflection layer R with a kernel k . Then the observed image I can be modeled as an additive mixture of the ghosted reflection and transmission layers by R and T respectively,

$$I = T + R \otimes k \quad (5)$$

The kernel k represented as a two-pulse kernel, parameterized by the distance and the relative intensity between the primary and secondary reflections. Given the input image I , our goal is to recover the kernel k , transmission layer T , and reflection layer R . Assumed that $I = T + R$ in their imaging models. Solving this ill-posed problem requires either very effective image priors, or auxiliary data such as multiple images captured with motion or polarizers, or user input.

These parameters may be estimated with a simple algorithm that relies on the auto-correlation of I . Then, given the observation I and kernel k , we recover T and R by using two forms of regularization: a patch based GMM method, and non-negativity constraints on T and R , which help in the

regularization of low frequencies. Then formulate the recovery problem as follows,

$$\min_{T,R} \frac{1}{\sigma^2} \|I - T - R \otimes k\|^2 - \sum_i \log(GMM(P_i T)) - \sum_i \log(GMM(P_i R)) \quad \text{s.t. } T \geq 0, R \geq 0 \quad (6)$$

Where $GMM(P_i I)$ refers to the Gaussian Mixture Model based patch method, and $P_i I$ is a linear operator that extracts patch i from image I . We minimize the above function using alternating minimization with half-quadratic optimization and a constrained L-BFGS algorithm.

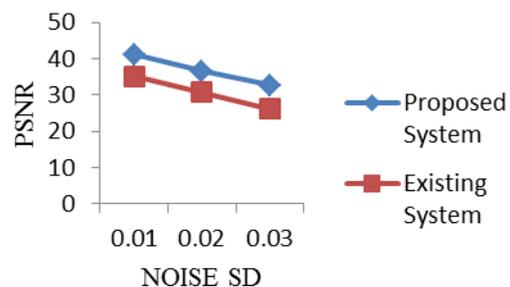
IV. EXPERIMENTAL RESULTS

Patch based models are easier to learn and to work with than whole image models. We have shown that patch models which give high likelihood values for patches sampled from natural images perform better in patch and image restoration tasks. Given these results, we have proposed a method which allows the use of patch models for whole image restoration, motivated by the idea that patches in the restored image should be likely under the prior. We have shown that this work improves the results of whole image restoration considerably when compared to simple patch averaging, used by most present day methods. The GMM model performance with EPLL gives better image restoration. Finally, we have proposed a new, simple Gaussian Mixture prior which performs surprisingly well on image denoising, deblurring. This algorithm applies on both synthetic and real world inputs, demonstrating how it helps in the recovery of transmission and reflection layers without distortion. It shows that ghosting cues help in the recovery of high quality results from just a single input image. An example of the resulting image is shown in Figure 2. MATLAB implementation takes one hour to process an input RGB image of size 200 x 300. Results are superior both in PSNR and quality of image.

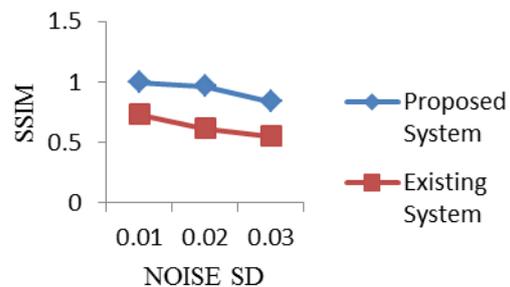
Fig.3. shows the plot of PSNR, SSIM value of proposed method and existing method to compare the performance. From the plots it can be seen that the PSNR value of input image and recovered image is high in the proposed method comparing to the existing method. When comparing to the existing method, this method gives good PSNR value and SSIM of transmitted image and, thereby giving a high accuracy in recovered image. This model assumes white Gaussian noise, also we can see that, when noise standard deviation of input image increases, PSNR value and SSIM decreases. For blur kernel size of input image increases, PSNR value and SSIM decreases. Finally, this experiment can be repeated using different images.



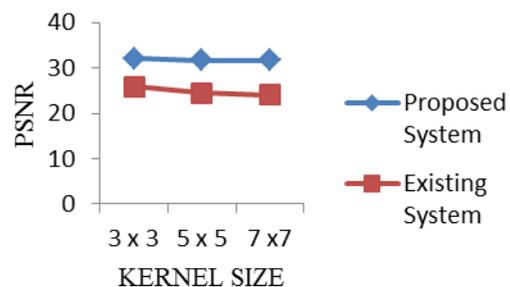
Fig.2. Results on real world input



(a)



(b)



(c)

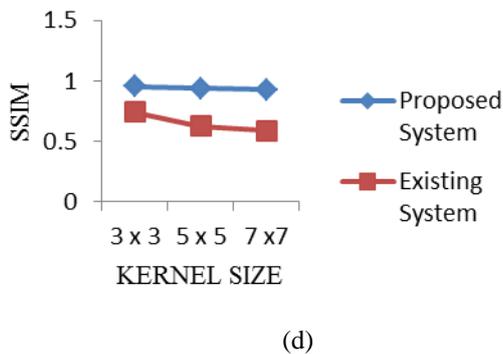


Fig.3. Plots of various parameters of existing and proposed method. (a) PSNR of input (noise) image and recovered Transmission. (b)SSIM value of input (noise) image and recovered Transmission. (c)PSNR of input (blur) image and recovered Transmission.(d) SSIM of input (blur) image and recovered Transmission.

V. CONCLUSION

The contribution of this work is removing reflection and distortion from a single image. Here a reliable method is used for removing the reflection and distortion from single images. For this we use two algorithms for removing both the contents from the images, and present an algorithm which uses patch-based GMM priors to achieve high-quality separation of reflection and transmission images on real-world and synthetic data, the input image consists both of this. The output is high quality reflection and distortion separated images. Ghosting cues break the symmetry between reflection and transmission but tend to be less effective for low frequencies. Better low-frequency regularization techniques would be an interesting research direction. Using this method the quality of image can be measured by PSNR and SSIM.

REFERENCES

- [1] Tianfan Xue, Michael Rubinstein, Ce Liu, and William T. Freeman. A computational approach for obstruction-free photography. *ACM Trans. Graph.*, 34(4):79:1–79:11, July 2015
- [2] YiChang Shih, D. Krishnan, F. Durand, and W.T. Freeman. Reflection removal using ghosting cues. In *Computer Vision and Pattern Recognition (CVPR)*, 2015 IEEE Conference on, pages 3193–3201, June 2015.
- [3] Xiaojie Guo, Xiaochun Cao, and Yi Ma. Robust separation of reflection from multiple images. In *Computer Vision and Pattern Recognition (CVPR)*, 2014 IEEE Conference on, pages 2195–2202, June 2014.
- [4] Cheng Lu and Mark S. Drew; School of Computing Science, “Practical Scene Illuminant Estimation via Flash/No-Flash Pairs” Simon Fraser University, Vancouver, British Columbia, Canada V5A 1S6 {clu,mark}@cs.sfu.ca2014
- [5] N. Kong, Y. Tai, and J. Shin. A physically-based approach to reflection separation: From physical modeling to constrained optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 36(2):209–221, 2014.
- [6] Y. Li and M. S. Brown. Single image layer separation using relative smoothness. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014
- [7] Yan, Qing and Xu, Yi and Yang, Xiaokang. Separation of weak reflection from a single superimposed image using gradient profile sharpness. In *IEEE International Conference*, 2013

- [8] Y. Li and M. S. Brown. Exploiting reflection change for automatic reflection removal. In *IEEE International Conference on Computer Vision (ICCV)*, 2013
- [9] Yilong Geng “Department of Electrical Engineering Stanford University Zizhen Jiang Department of Electrical Engineering Stanford University “Reflection Removal on Mobile Devices”.2013
- [10] J. Kopf, F. Langguth, D. Scharstein, R. Szeliski, and M. Goesele. Image-based rendering in the gradient domain. *ACM Transactions on Graphics (TOG)*, 32(6):199, 2013.
- [11] B.himabindu (Asst. professor, Department of E.C.E, Chalapathi Institute of Technology, Guntur, A.P, India). “Removal of Shadows and Reflections in the Images By Using Cross-Projection Tensors” *IOSR Journal of Engineering (IOSRJEN) ISSN: 2250-3021 Volume 2, Issue 8 (August 2012), PP 34-40 www.iosrjen.org www.iosrjen.org 34|Page*
- [12] K. Gai, Z. Shi and C. Zhang, “Blind Separation of Superimposed Moving Images Using Image Statistics,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 19-32, Jan. 2012
- [13] S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski. Image-based rendering for scenes with reflections. *ACM Transaction on Graphics (SIGGRAPH)*, 31(4):100, 2012.
- [14] Song,Bo; Gong,shenwen; Ren,chunjian “Removing artifacts using gradient projection from a single image”. *MIPPR 2011: Pattern Recognition and Computer Vision*. Edited by Roberts, Jonathan; Ma, Jie. *Proceedings of the SPIE*, Volume 8004, article id. 80041C, 6 pp. 2011.
- [15] Matlab,Design,Implementation, Available : Wikipedia
- [16] Wikipedia. http://en.wikipedia.org/wiki/Insulated_glazing.

Mrs. Husna k, completed B.Tech in Electronics And Communication Engineering from Calicut university in 2011. Currently pursuing M.Tech in Electronics and Communication Engineering from APJ Abdul Kalam Technological University. Her research interests include Image Processing and Embedded Systems.

Prof. Abdul Rias P, Completed B.Tech in Electronics And Communication Engineering from Calicut university .Completed M.tech in Applied Electronics from Anna University. Currently working as an Assistant Professor in Electronics and Communication Engg.dept. of Al Ameen Engineering College, Palakkad, Kerala State, India.