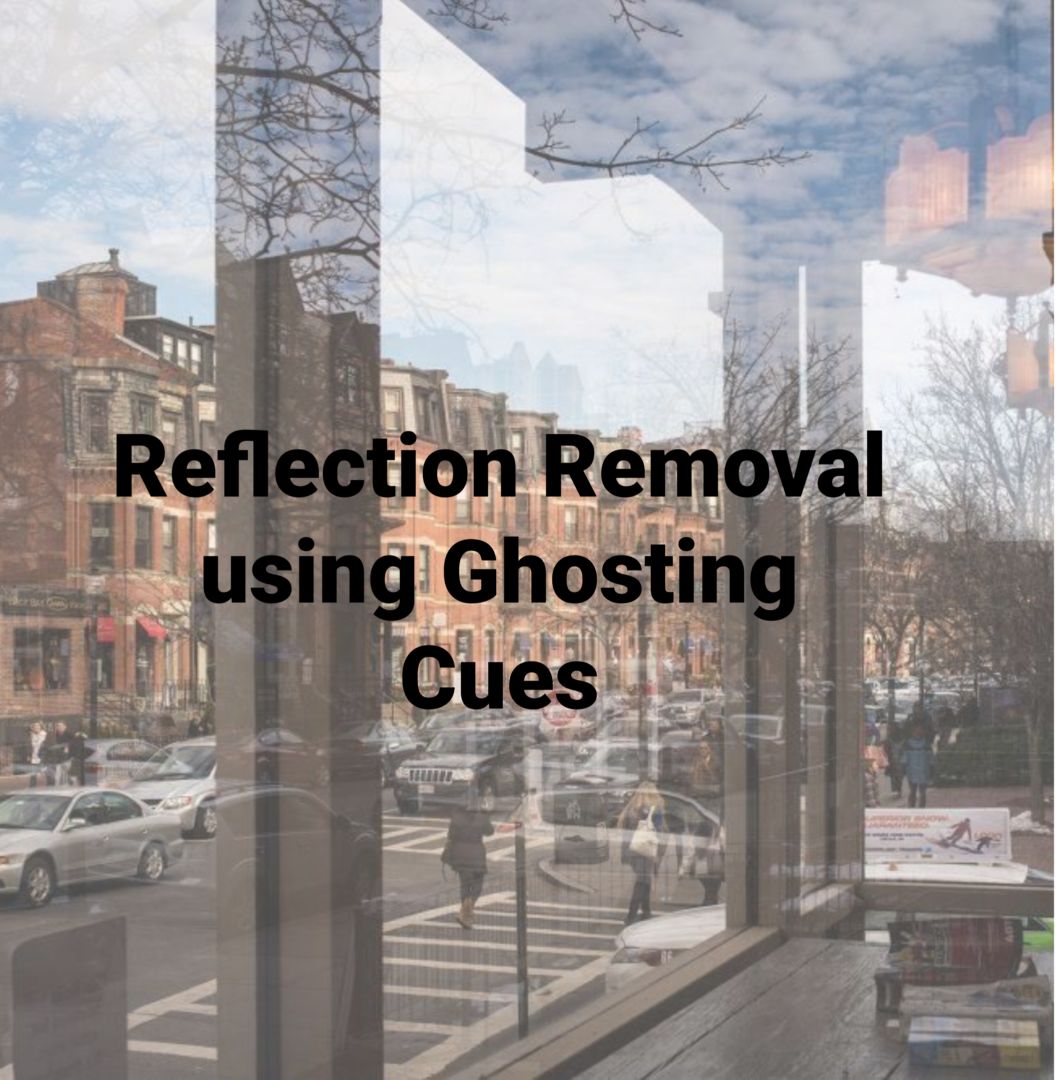


Reflection Removal using Ghosting Cues

Team : Framed

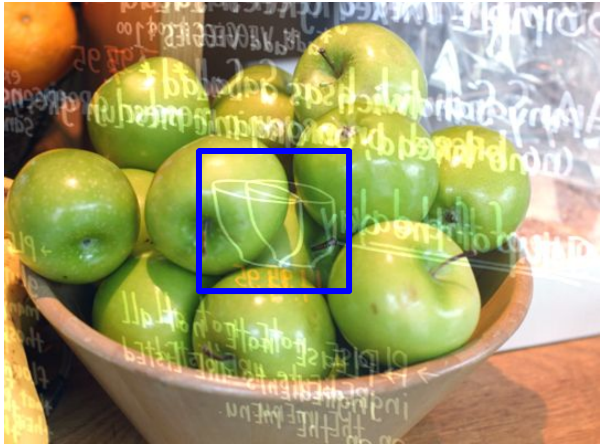
Avinash Prabhu : 2018102027 ECE
Fiza Husain : 2018101035 CSE
Mallika Subramanian : 2018101041 CSE
Tanvi Karandikar : 2018101059 CSE

TA Mentor : Meher Shashwat Nigam

Project Repository

Problem Statement Overview

The primary goal of the project is to implement an algorithm that performs post-processing on the images to remove reflection artifacts. The important property that is leveraged here is the ghosting cues that arise from double shifted reflections of the reflected scene off the glass surface.



How do these cues occur?

Many a times images that are taken through glasses surfaces have some unwanted artifacts in the final image. These arise due to the reflections of the scene on the same side of the glass as the camera.

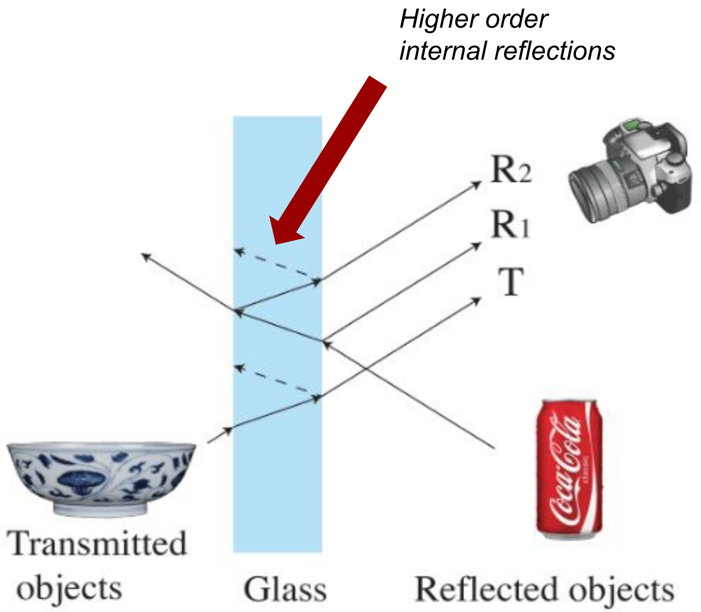
Assumptions while modelling ghosting

1. The spatial shift and relative attenuation between R1 and R2 is spatially invariant.

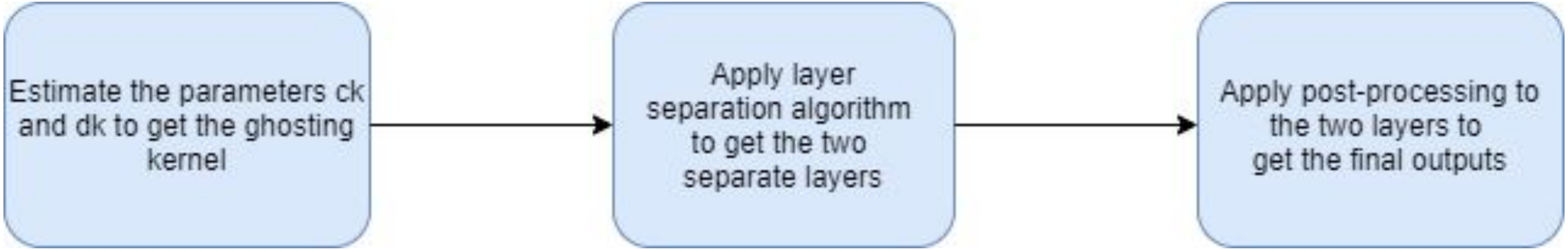
2. These assumptions hold when the reflection layer does not have large depth variations, and when the angle between camera and glass normal is not too oblique.

3. The higher-order internal reflections of both T and R can be ignored. For typical glass with refraction index around 1.5, these higher-order reflections contribute to less than 1% of the reflected or transmitted energy.

4. The glass is planar.



Solution Pipeline



STEP 1 : Estimating K

Estimating K

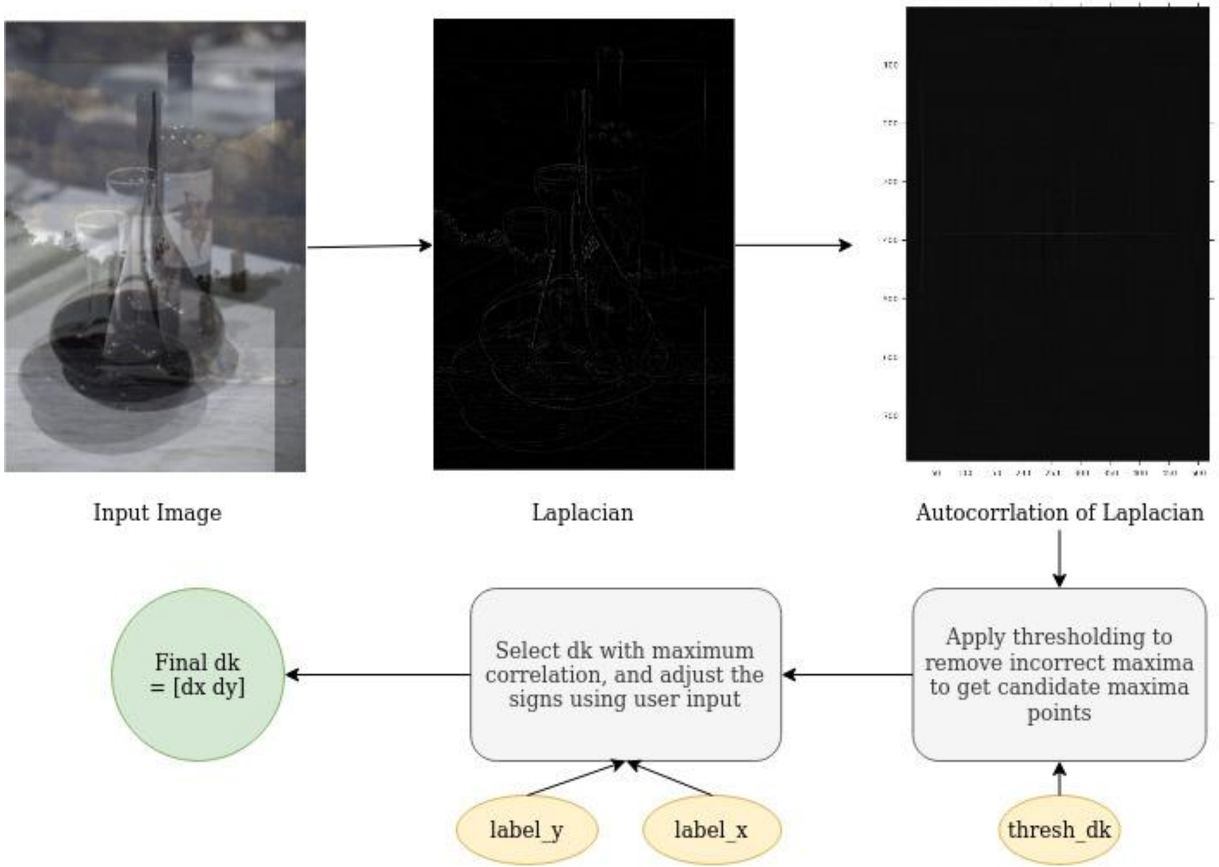
In order to remove the ghosting effects of the reflections, a ghosting kernel is used.

This kernel has two variables c_k *and* d_k. The formation model for the observed image ***I***, given the transmission ***T***, reflection ***R*** and ghosting kernel ***k***, is:

- d_k represents the spatial shift : it depends on glass thickness, camera focal length etc.
- c_k is the attenuation factor that is affected by wave optics.

$$I = T + R \otimes k + n$$

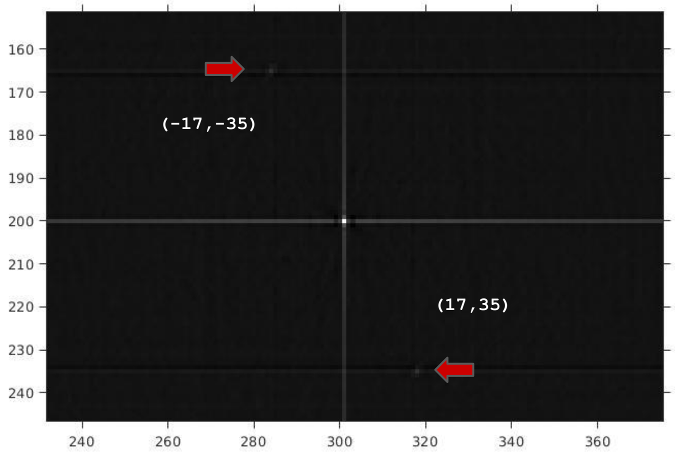
Estimating dk



Choosing Right Dk

Original Input Image

(Zoomed in Version)



Autocorrelation $\nabla^2 I$

Consequence of Incorrect d_k

(-18,39)

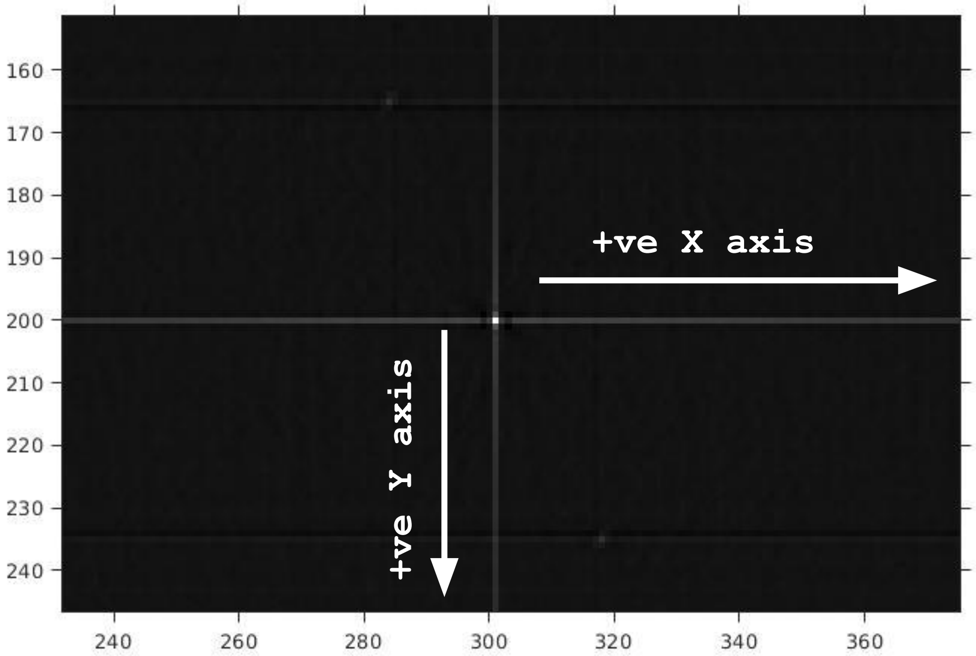
Incorrect *Reflection Layer*

(18,-39)

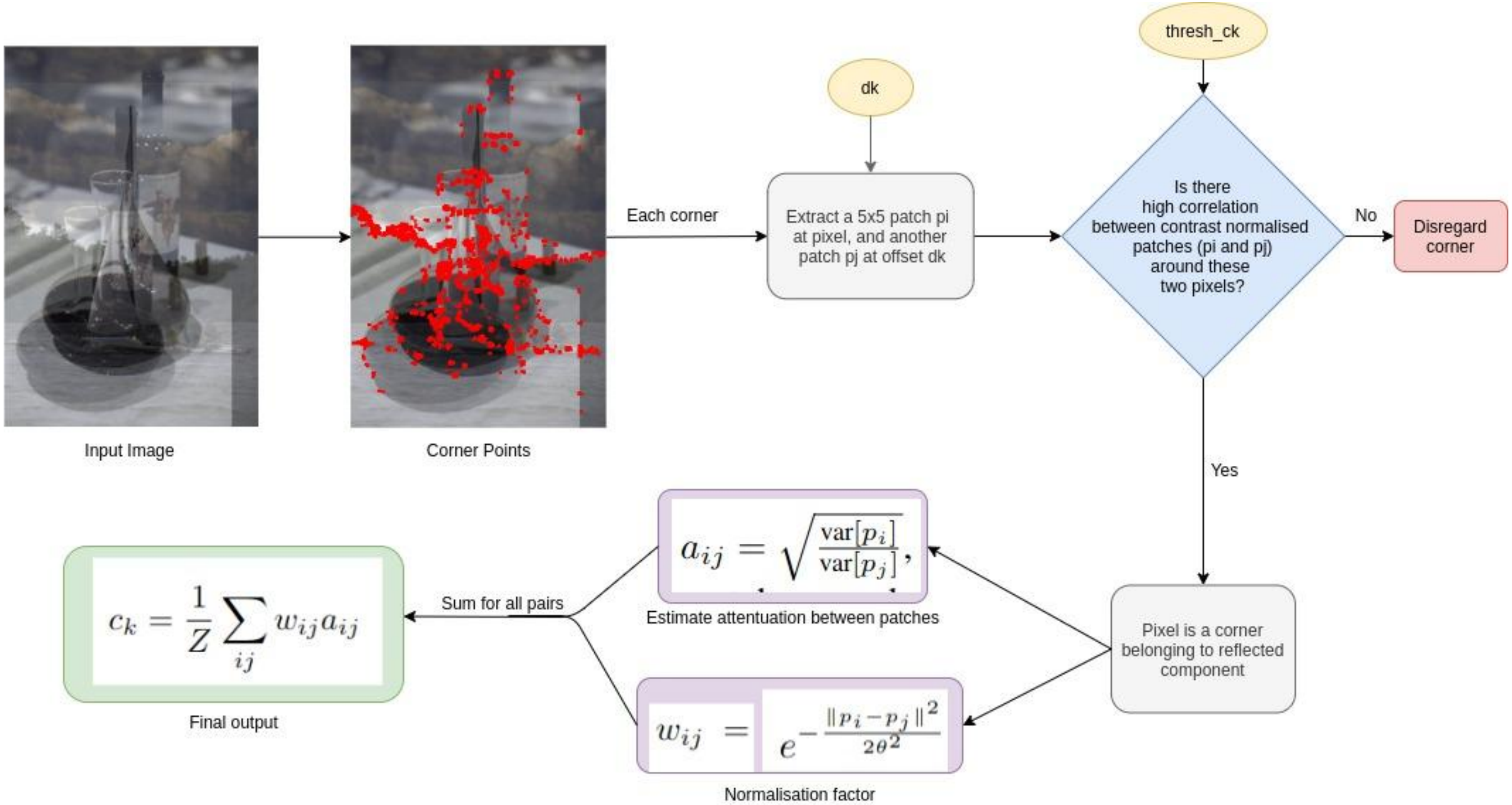
Correct Reflection Layer

Resultant Reflection Removal

Sign Convention followed



Estimating ck



Choosing right Ck

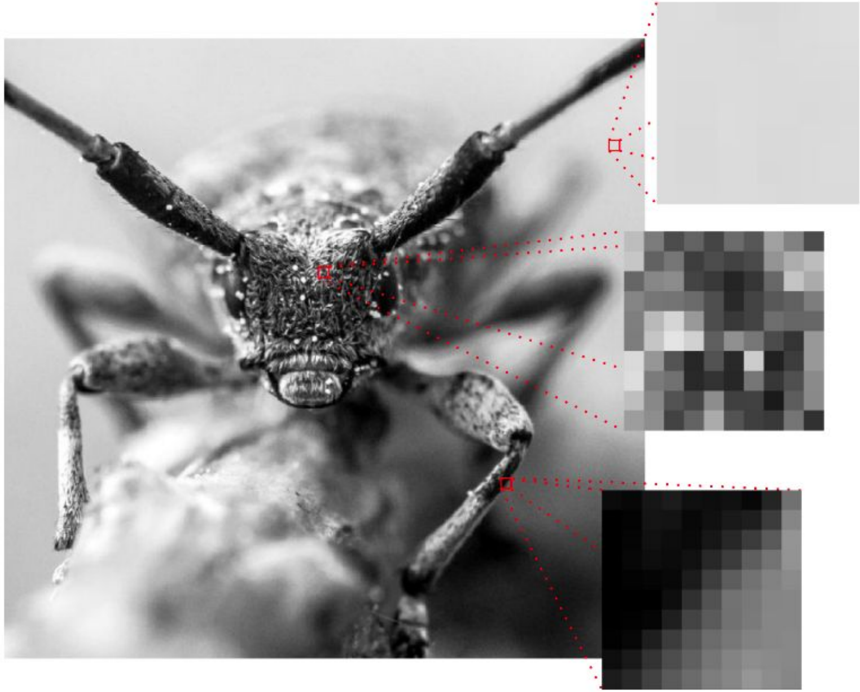
- Of all the pairs of 5x5 patches that are extracted, we assume that if there is a strong correlation between two patches, the edges are due to the reflection layer.

- The correlation values of :
    - corr > 0.7 : indicate *high* correlation
    - 0.5 < corr < 0.7 : *moderate* correlation
    - corr < 0.4 : *weak* or *no* correlation

- Hence, consider or discard patches based on this approach

STEP 2 : Layer Separation

Image Representation Models for Natural Images



Natural images are full off redundancy, meaning that we could discard some of the information but still be able to represent the same image well. This inherent property of natural images gives us a chance to find image representation models that only need a small number of variables to represent the given image data.

What are image priors and why GMM?

Image Priors is prior information on set of your images. Priors can be put into math form and can be merged into the processing (filtering, segmentation, etc) and reduce the feasible set of solutions through optimization algorithms.

- We use the Estimated Patch Prior Log Likelihood to determine the patch priors for the Transmission and Reflection layers. Such a cost function is non convex because of the presence of GMM priors.

- With GMM, any continuous probability density can be approximated to some arbitrary accuracy by using a sufficiently large number of single Gaussians and by adjusting their means and covariances as well as the mixture weights of the linear combination.

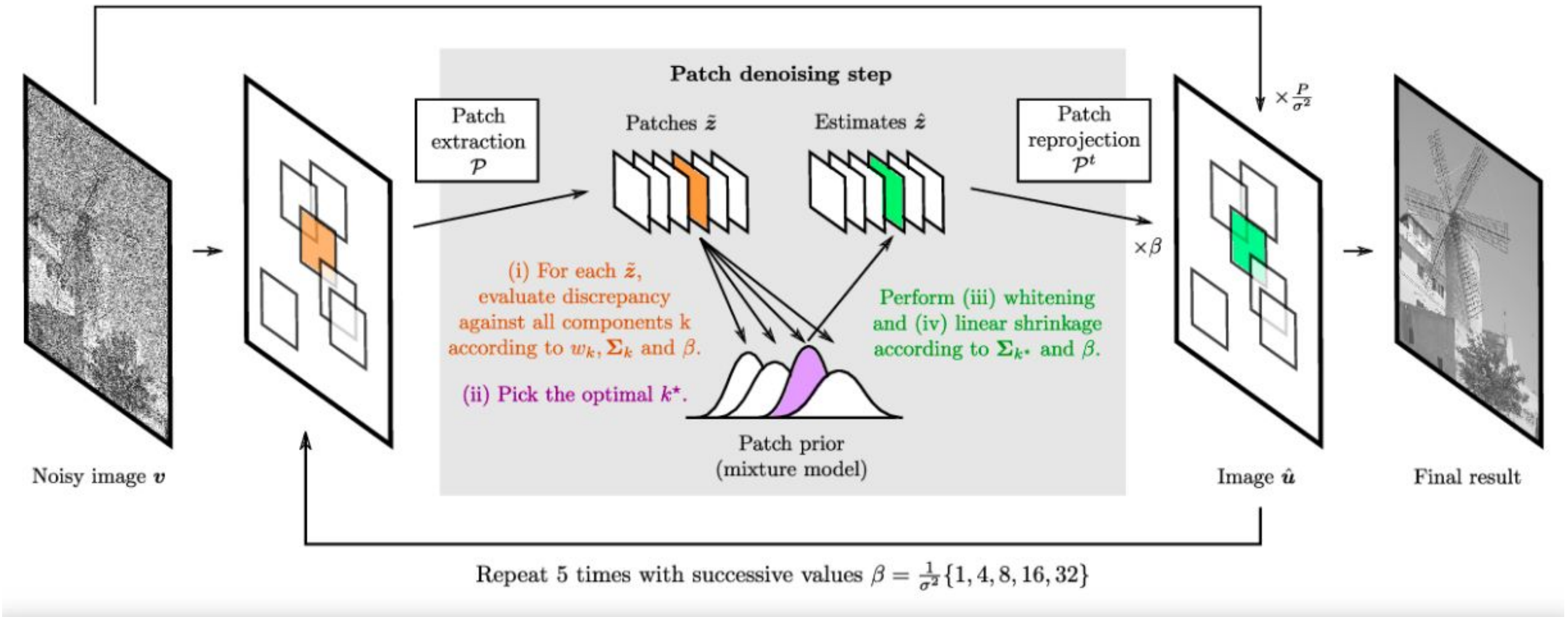


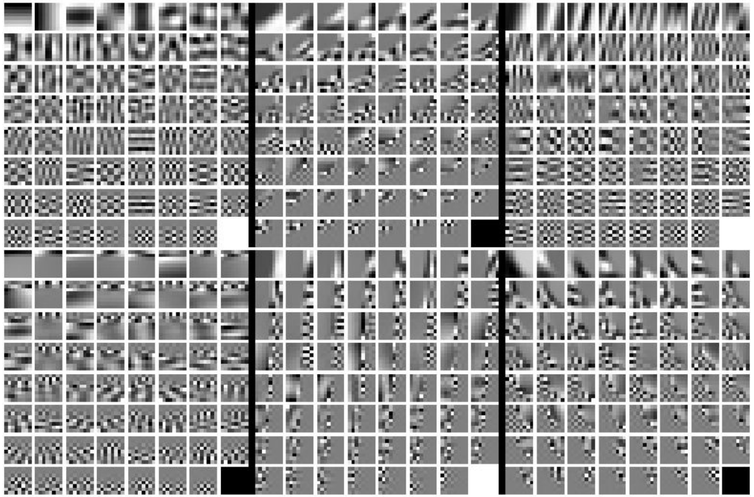
Illustration of EPLL framework for image denoising with a GMM prior

https://pdfs.semanticscholar.org/c37b/cb285b0e8a273457cbe002483ccf0daa056f.pdf?_ga=2.15386870.1505779178.1605668887-1212042111.1604820108

Why we use GMM?

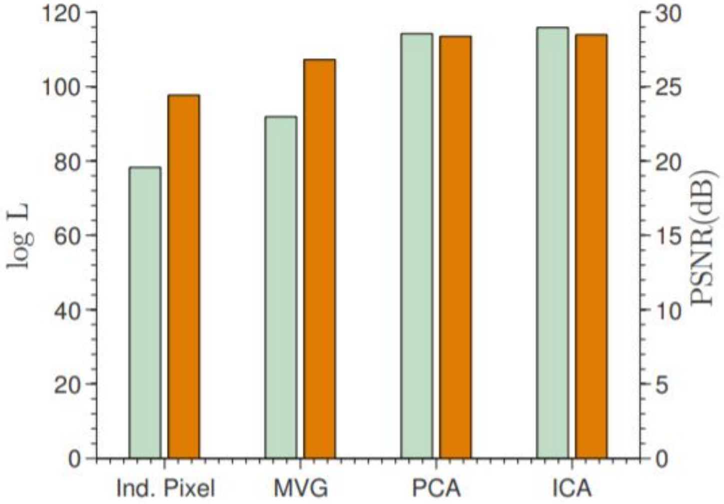
GMM framework is based on the assumption that the accumulation of similar patches in a neighborhood are derived from a multivariate Gaussian probability distribution with a specific covariance and mean.

Natural images are believed to be some set that has similar statistical structure(i.e. They are highly non-random) to which our visual system is adapted to.



Eigenvectors of 6 randomly selected covariance matrices from the learned GMM model, sorted by eigenvalue from largest to smallest. Note the richness of the structures - some of the eigenvectors look like PCA components, while others model texture boundaries, edges and other structures at different orientations.

Models other than GMM

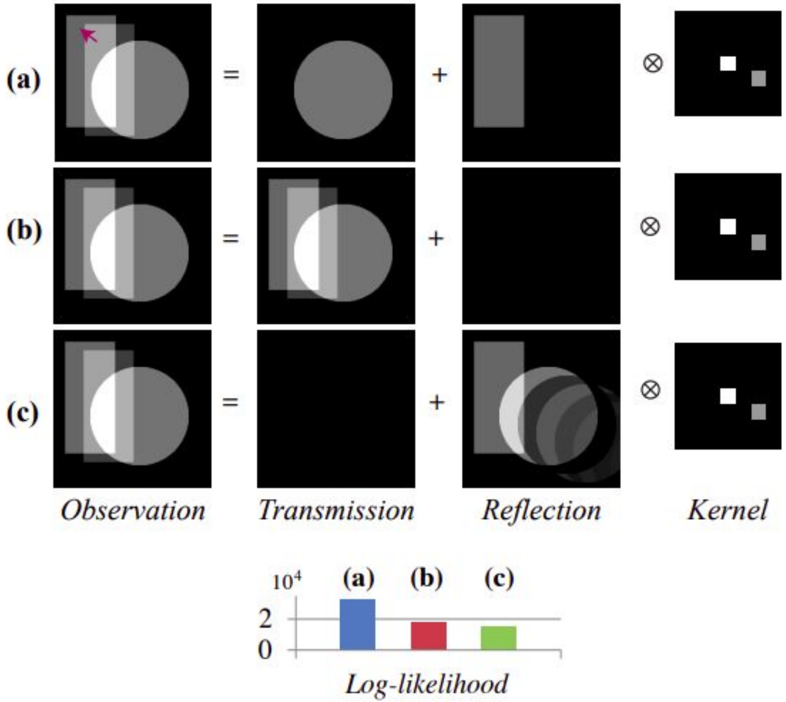


- GMM outperforms ICA when total number of components are less than 10 and the performance is almost identical for larger number of components

- The principal components distribution, conditioned on the observed variables, is multivariate normal whereas in GMM, hidden components are a mixture distribution, sum of several multivariate normal therefore, can accommodate non-linear.

https://people.csail.mit.edu/polina/papers/Kim_MSc_Thesis.pdf
https://www.hindawi.com/journals/mpe/2020/1202307/
https://people.csail.mit.edu/danielzoran/EPLLICCVCameraReady.pdf

Log-likelihood example



In (a), the ground-truth decomposition is sparsest, e.g., in the gradient domain, and therefore the most “natural”. This decomposition has the highest likelihood of this decomposition under the GMM model. The two extreme decompositions include ghosting artifacts and are less sparse. They are less natural, and their likelihoods are lower.

Our Deghosting :

Optimization

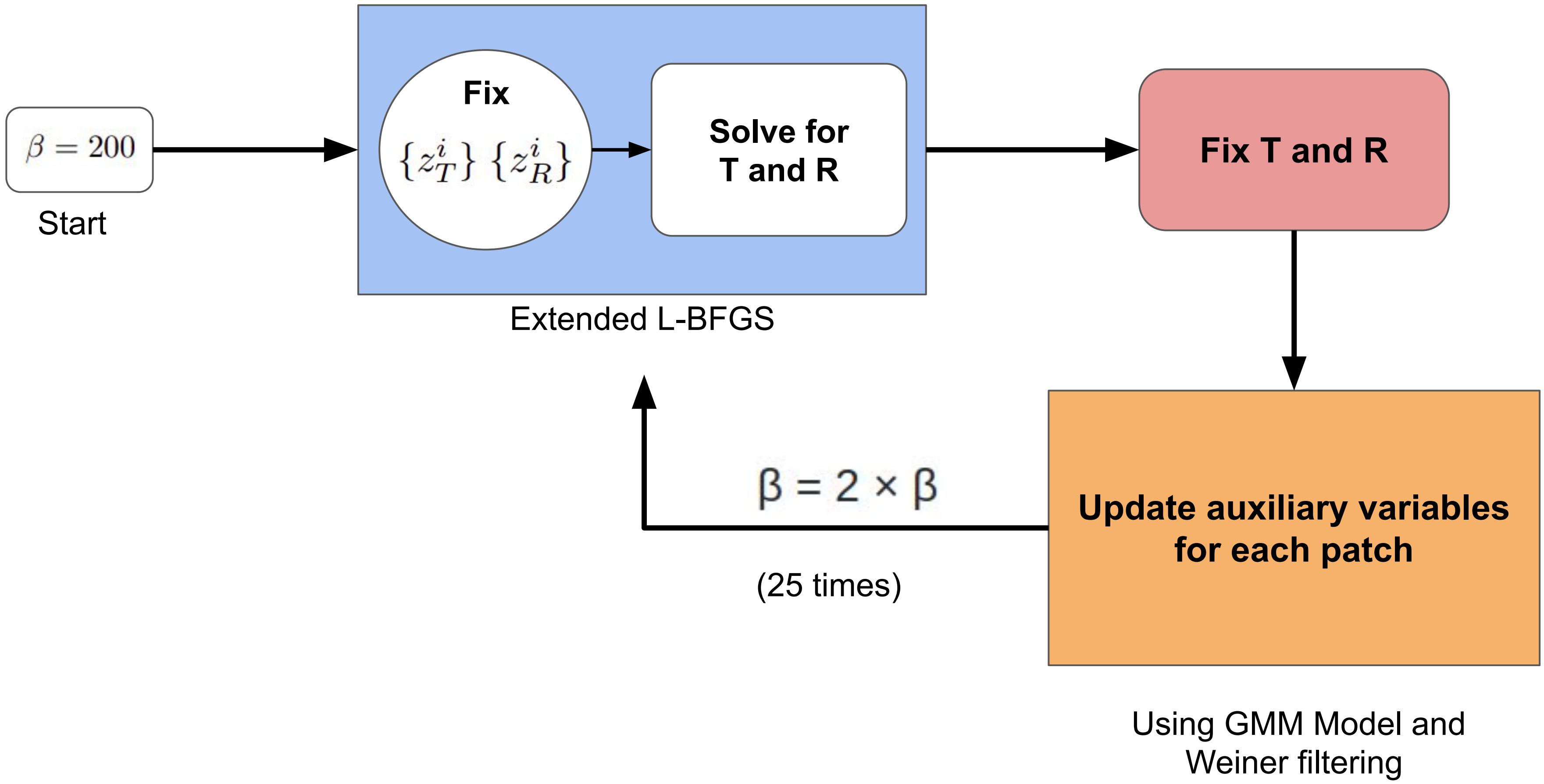
Cost Function that needs to be minimized:

$$\min_{T,R,z_T,z_R} \frac{1}{\sigma^2}\|I - T - R\otimes k\|_2^2 + \frac{\beta}{2}\sum_i \left(\|P_iT - z_T^i\|^2 + \|P_iR - z_R^i\|^2\right) - \sum_i \log(\mathrm{GMM}(z_T^i)) - \sum_i \log(\mathrm{GMM}(z_R^i))$$

$$\text{s.t. } 0 \leq T, R \leq 1$$

Non-negativity constraints on T and R to regularize low frequencies

https://dl.acm.org/doi/pdf/10.1145/279232.279236?casa_token=TmPLw-JB7f4AAAAA:FtsTU2I2k9jIJYy0Ddmv6S-AfUfbt2j_ksO0UI5qEavLlzYB49XaOEAlRQevBjvU8zaTfsIFdxujeqo



$\beta = 200$
Start
Fix
$\{z_T^i\}$ $\{z_R^i\}$
Solve for
T and R
Extended L-BFGS
Fix T and R
Update auxiliary variables
for each patch
Using GMM Model and
Weiner filtering
$\beta = 2 \times \beta$
(25 times)

STEP 3 :Post Processing

Post Processing

Using the mean and variance of the original image, we adjust the contrast of the transmitted layer to get as close as possible to the original image.

Output Transmission Layer

Original Image

Processed Output Transmission Layer

RESULTS

Separately Collected Outputs

Tests Conducted

Hyperparams :

***ck** = 0.8212 (thresh=0.85)*
***dk** = [-17,-35] (signx = -1, signy = -1, thresh = 0)*

Input Image

Reflection Layer

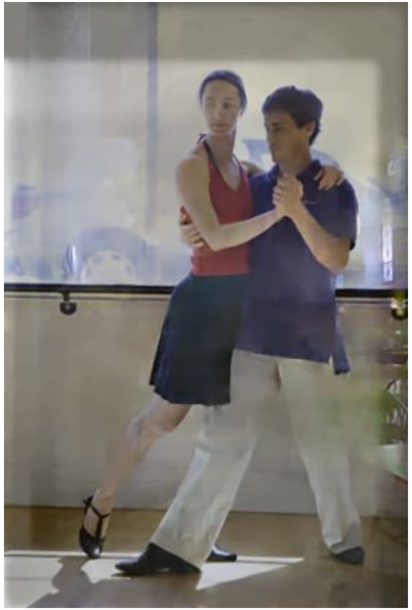
Transmission Layer

Tests Conducted

Hyperparams : ***ck*** *= 0.8268 (thresh=0.9)*
dk *= [18,-39] (signx = 1, signy = -1, thresh = 0)*

Input Image

Reflection Layer



Transmission Layer

Tests Conducted

Hyperparams :

ck *= 0.8242 (thresh=0.73)*
dk *= [-14,-25] (signx = -1, signy = -1, thresh = 0)*

Input Image

Reflection Layer

Transmission Layer

Tests Conducted

Hyperparams : ***ck*** *= 0.5026 (thresh=0.1)*
dk *= [-15,-37] (signx = -1, signy = -1, thresh = 0)*

Input Image

Reflection Layer

Transmission Layer

Tests Conducted

Hyperparams :

***ck** = 0.6623 (thresh=0.85)*
***dk** = [-14,-13] (signx = -1, signy = -1, thresh = 0)*

Input Image

Reflection Layer

Transmission Layer

Tests Conducted

Hyperparams :

***ck** = 0.8268 (thresh=0.9)*
***dk** = [18,-39] (signx = 1, signy = -1, thresh = 0)*

Input Image

Reflection Layer

Transmission Layer

RESULTS

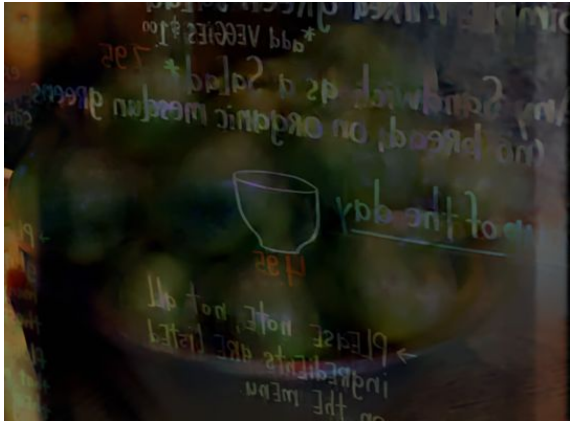
Paper Image Outputs

Paper Image

Hyperparams :

ck *= 0.802189 (thresh=0.7)*
dk *= [-30,0] (signx = 0, signy = 0, thresh = 0)*

Input Image



Reflection Layer

Transmission Layer

Paper Image

Hyperparams : ***ck*** *= 0.593473(thresh=0.9)*
dk *= [-31,27] (signx = 0, signy = 0, thresh = 0)*

Input Image

Reflection Layer

Transmission Layer

Caveats & Findings

- Ambiguity in determining *“Predefined Threshold”* to discard maximas where first and second maxima are close

- Choosing appropriate signs for candidate d_k as both values result in $c_k < 1$

- *“Highly correlated”* patches not elaborately explained. Hence discrepancy in choosing the right patches for c_k computation

Limitations

1. Requires thick glass windows, and large angles between camera viewing angle and glass surface for sufficient ghosting.

2. Sensitive to strong repetitive textures in the transmission layer, as that can be mistaken to be ghosting.

3. We assume spatially-invariant ghosting

    a. The reflection layer does not have large depth variations

    b. The angle between camera and glass normal is not too oblique

Work Distribution

Avinash

- Estimate c_k

Fiza

- GMM
- Construct Kernel

Mallika

- Estimate d_k

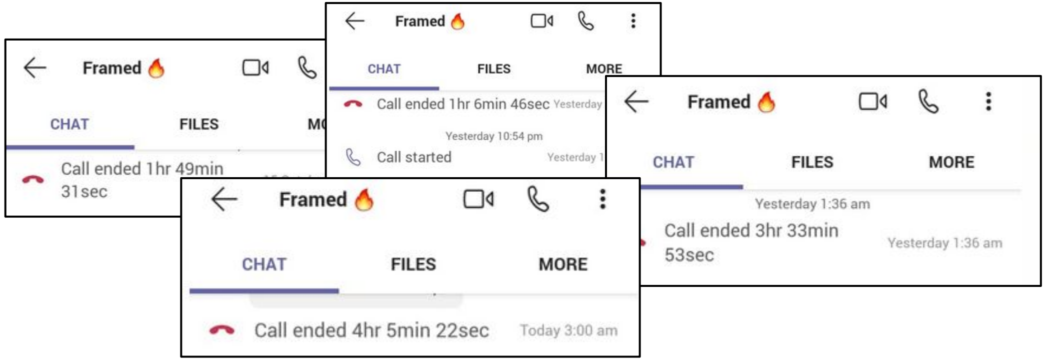
Tanvi

- Integration & running on ADA

Team Work :

All members of the team equally contributed towards

- ➔ Understanding the paper in detail
- ➔ Debugging and rectifying problems
- ➔ Brainstorming to think of solutions



THANK YOU!