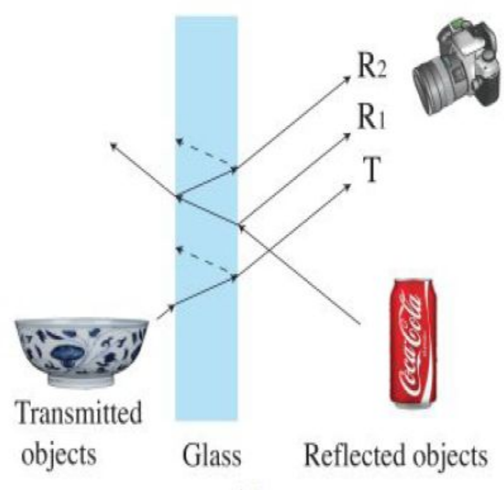
DIP REPORT
Reflection Removal Using Ghosting Cues

Parth Partani (20161034)
Sparsh Garg (20161025)

Introduction

When taking a picture through a window pane, reflections of objects are often captured. To minimize reflection artifacts, one may try to change camera position, use polarizers, or put a piece of dark cloth around the camera, but often it is impractical to make any of these adjustments. This raises the need for post-processing to remove reflection artifacts.



(a) *Input* **(b)** *Close-up of ghosting*

(c) *Recovered Transmission* **(d)** *Recovered Reflection*

Problem Statement

Image captured by camera contain both transmitted objects (T) and undesired reflected objects (R) as can be seen from above ray diagram. Our aim is to separate T and R.

Traditional Imaging model:

$$I = T + R$$

This model assumes that both T and R play symmetric role when forming output I. As both T and R are natural, separating them is ill-posed since both T and R are natural images and have the same statistical properties.

Method used in paper

Separate the reflection layer using

- The Double reflection imaging model (Ghosting).
- With patch-based image prior.

Now, how to accomplish above points?

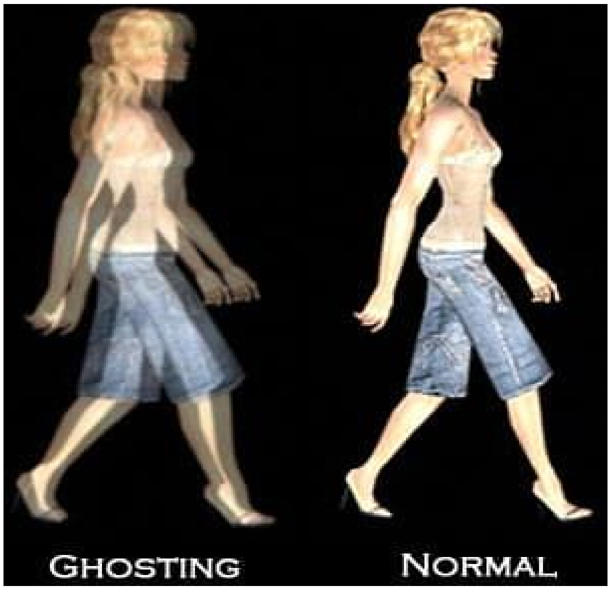
Key Idea

Break the symmetry of T and R using Ghosting.

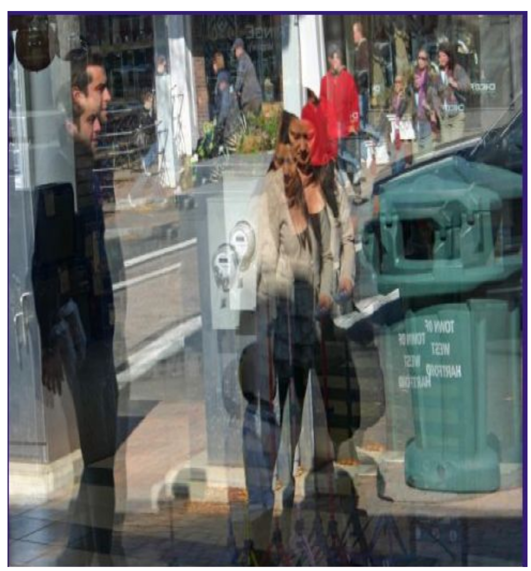
Now What's this Ghosting?

It can be described as appearance of secondary image on the main display. Ghosting can be caused when moving objects

across multiple images are combined into one image. In our problem case, it is caused by multiple reflections.



Observation in our case : Window reflection often appear multiple times causing ghosting for R. A common example is a double-pane window, which consists of two thin glass panes separated by some distance. 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.



Above image shows combination of T and ghosting of R due to multiple reflections.

Modelling Ghosting

Ghosting provides a critical cue to separate the reflection and transmission layers, since it breaks the symmetry between the two layers. We 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:

$$\mathbf{I = T + R \otimes k}$$

Assumptions made:

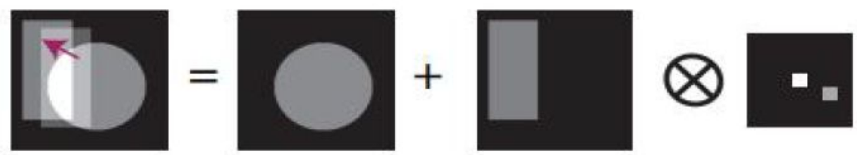
- We model the kernel k as a two-pulse kernel, parameterized by the distance and the relative intensity between the primary and secondary reflections. We ignored higher order reflections as they carry minimal energy.

Under these assumptions, the ghosting kernel k consists of two non-zero values. k is parameterized by a two-dimensional spatial shift d_k and relative attenuation factor c_k.

Given an image X, the result of convolving it with the kernel k gives an output Y , whose value at pixel i is:

$$Y_i = X_i + c_k X_i - d_k$$

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



Final Goal:

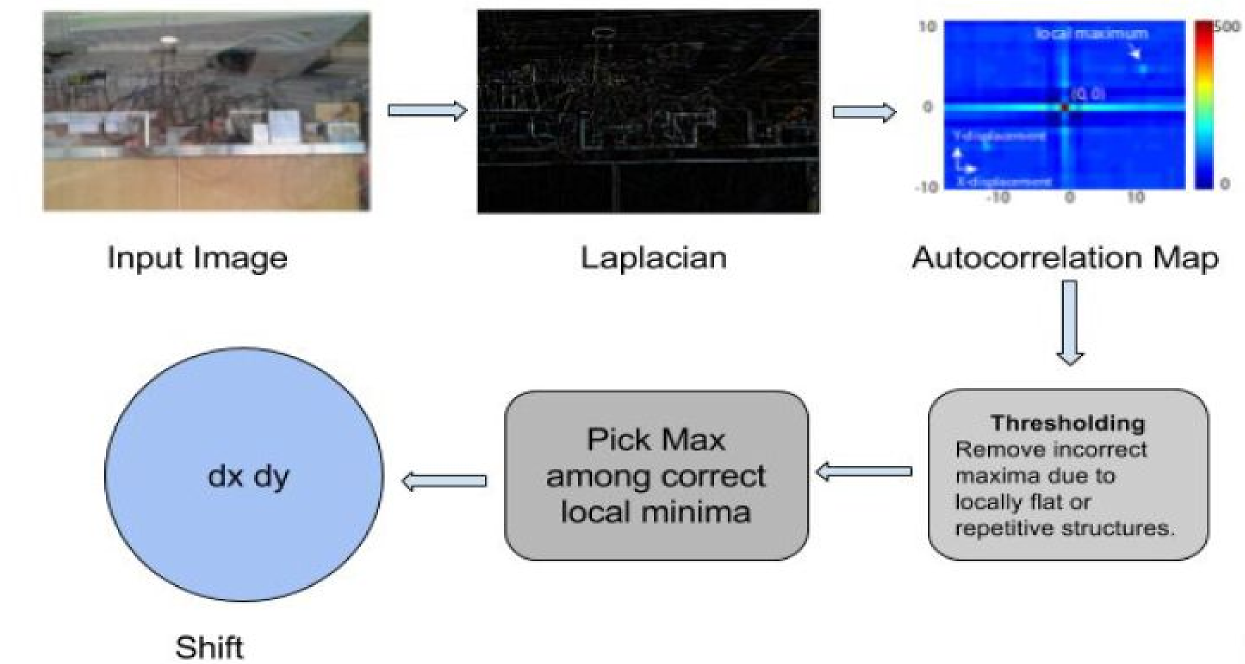
Given the input image I, our goal is to recover the kernel k, transmission layer T, and reflection layer R.

Estimation of kernel

First we'll find d_k.

- Laplacian filter is applied on input Image I and autocorrelation map of output ($\nabla^2 I$) is made.
- The shifted copies of the reflection layer create a local maximum at d_k on the autocorrelation map.
- To detect d_k, apply a local maximum filter in each 5-by-5 neighborhood.
- Discard local maxima in neighborhoods where the first and second maxima are closer than a pre-defined threshold.
- Also remove local maxima within 4 pixels of the origin.
- Finally, of the remaining local maxima, we select the largest one as the ghosting distance d_k.

Estimation of d_k



After getting d_k, we'll find c_k.

- Corners are detected using Harris corner method and a patch of 5x5 is taken

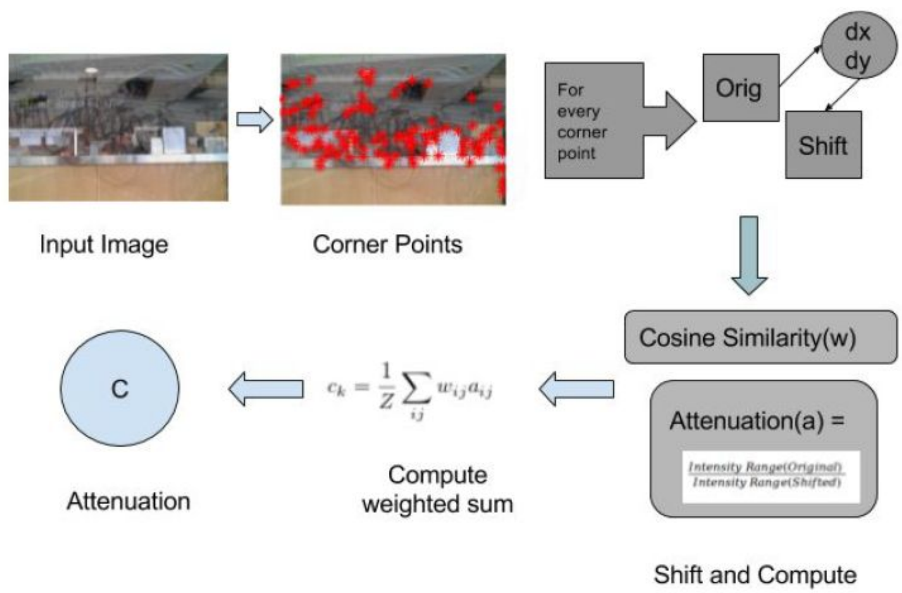
  around each corner feature.

- Correlation of this patch is calculated with a patch at spatial offset d_k. If these patches are strongly correlated then

  we estimate attenuation between these pairs.

- Finally we sum over all such pairs to give an estimate for final attenuation constant c_k.

Estimation of c_k



Also there can be ambiguity on the sign of d_k, and we resolve this by choosing d_k such that $c_k < 1$, using the property that the second reflection has lower energy than the first.

☆☆Above method for estimation of k can fail on images with strong globally repetitive texture.

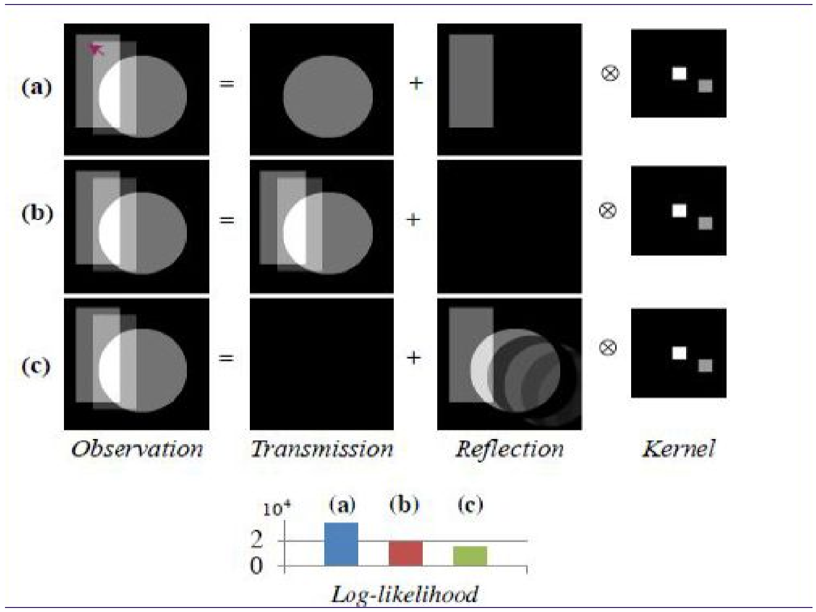
Reconstructing T and R for a given kernel

Toy Example

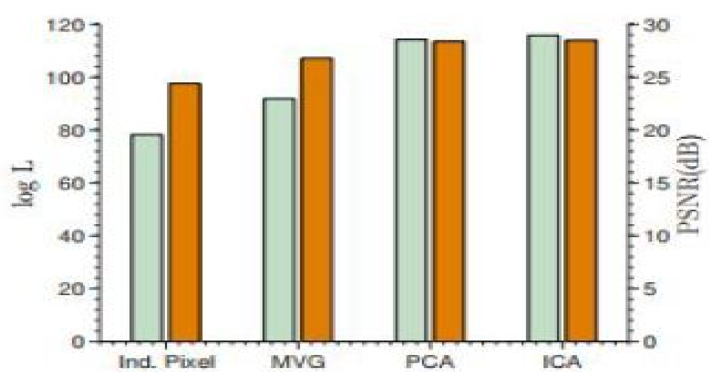
A synthetic example with a circle as the transmission layer and a rectangle as the reflection layer.

We compare the log likelihoods of the various possible decompositions under a GMM Model.

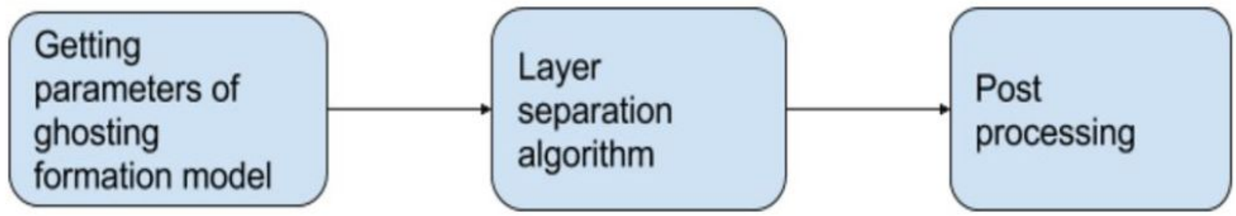
The log likelihood of the (a) is the highest, (implying is most “Natural”) which is indeed the ground truth.



We use GMM as it performs well for image restoration when compared to other methods.



The BIG PICTURE



Till now we have estimated parameters for ghosting formation model *i.e.* kernel

Layer Separation Algorithm

Our formation model for the observed image I, given the transmission T, reflection R and ghosting kernel k, is:

$$\mathbf{I = T + R \otimes k + n}$$

Here n is additive i.i.d. Gaussian noise with variance σ^2. Kernel parameters , we have already calculated.

Given k, the above formation model leads to a data (log-likelihood) term for reconstruction of T and R:

$$L(T,R) = \frac{1}{\sigma^2}\|I - T - R \otimes k\|_2^2$$

Minimizing L(T, R) for the unknowns T and R is ill-posed. Additional priors are needed to regularize the inference. The best-performing prior is patch-based prior based on Gaussian Mixture Models (GMM).

The GMM prior captures covariance structure and pixel dependencies over patches of size 8×8, thereby giving superior reconstructions to simple gradient-based filters, which assume independence between filter responses of individual pixels. Therefore the following cost has to be minimised.

$$-\sum_i \log(\mathrm{GMM}(P_i T)) - \sum_i \log(\mathrm{GMM}(P_i R))$$

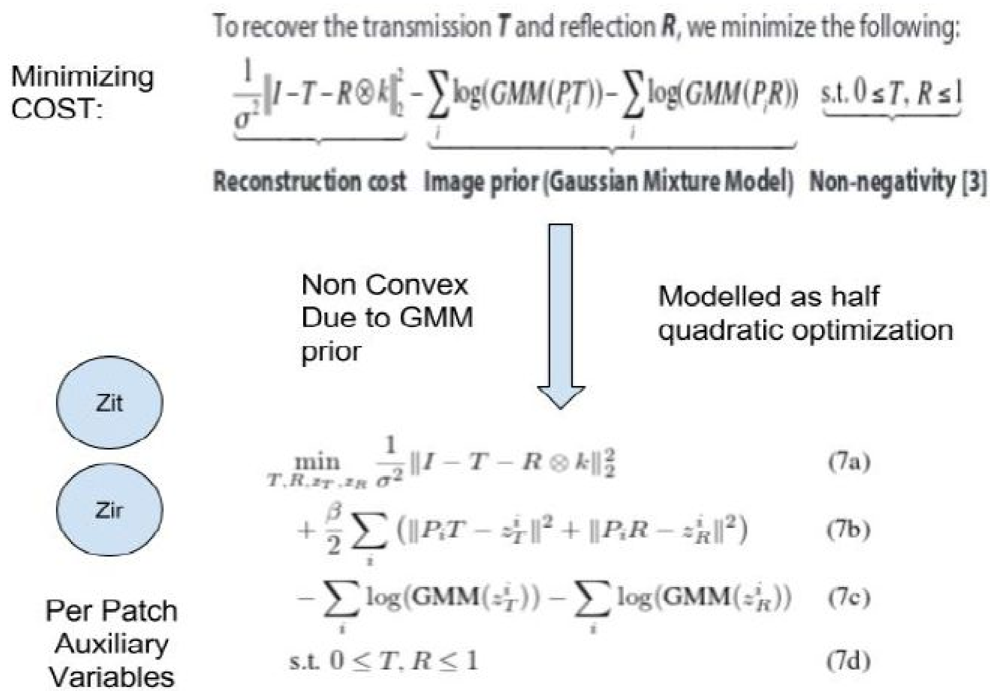
where $\mathrm{GMM}(P_i X) = \sum_{j=1}^{K} \pi_j N(P_i X; 0, \Sigma_j)$. The cost sums over all overlapping patches $P_i T$ in T, and $P_i R$ in R; where P_i is the linear operator that extracts the i th patch from T or R.A pre-trained zero-mean GMM model with 200 mixture components, and patch size 8 × 8. The mixture weights are given by $\{\pi_j\}$, and the covariance matrices by $\{\Sigma_j\}$. So our final cost function with non-negativity constraint on T and R is:

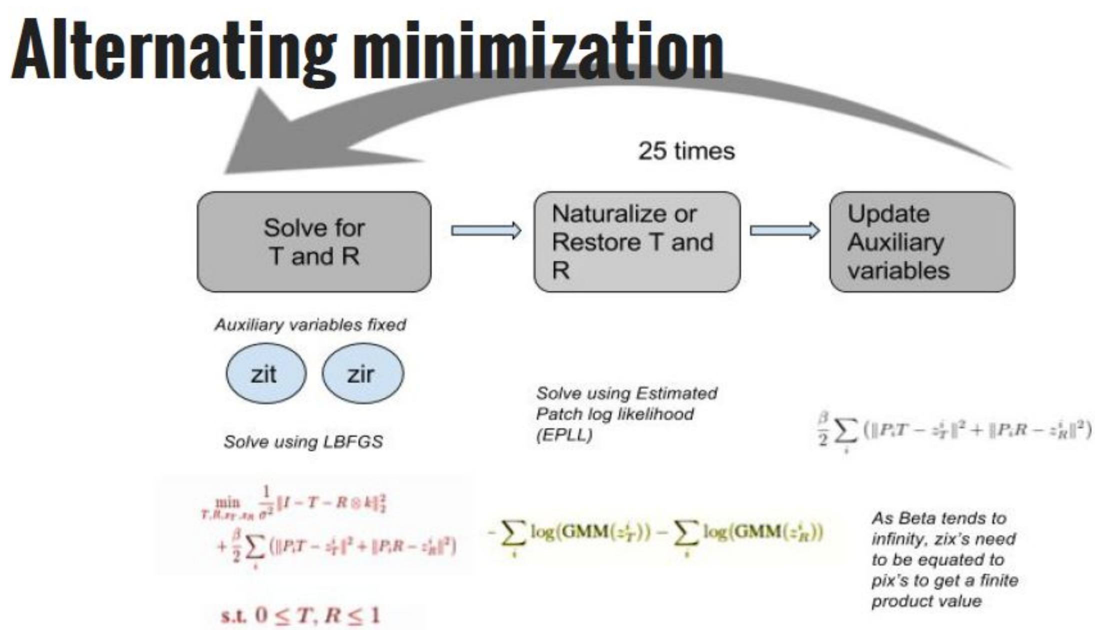
$$\begin{aligned}\min_{T,R} \quad & \tfrac{1}{\sigma^2}\|I - T - R \otimes k\|_2^2 - \textstyle\sum_i \log(\mathrm{GMM}(P_i T)) \\ & - \textstyle\sum_i \log(\mathrm{GMM}(P_i R)), \ \text{s.t. } 0 \le T, R \le 1\end{aligned}$$

This cost, we calculated above is non-convex due to the use of GMM prior. We use an optimization scheme based on halfquadratic regularization. We introduce auxiliary variables z^T_i and z^R_i for each patch $P_i T$ and $P_i R$, respectively.

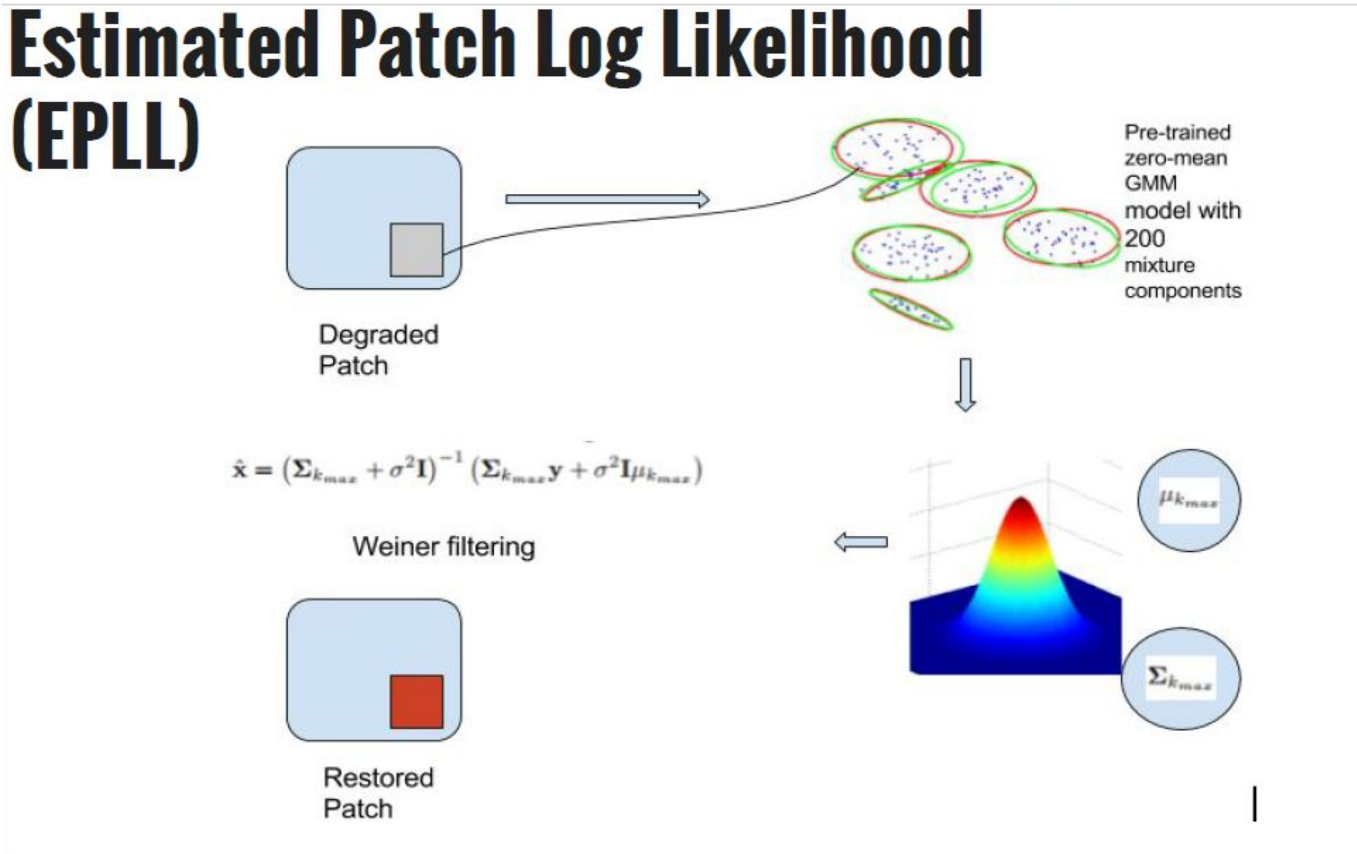
Optimization technique can be understood from below two diagram easily.

Optimization :





EPLL used for restoring T and R:



Results

Input Image 1

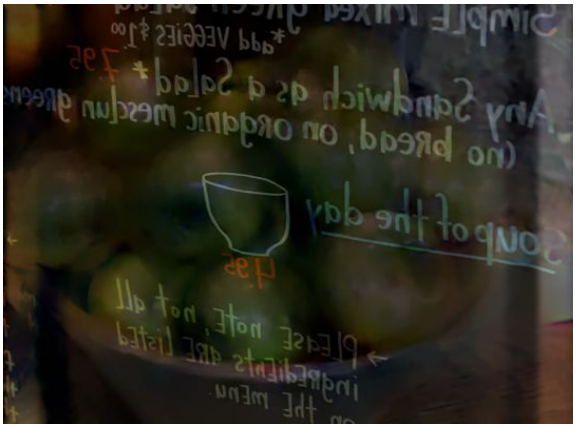
Transmitted

Reflected

Input Image 2

Transmitted

Reflected



Failure Cases

Input Image

Transmitted

Reflected