Project ID - 27

Reflection Removal Using Ghosting Cues

[Github]
[Research Paper]

Team : Psycho07

Parth Partani 20161034
Sparsh Garg 20161025

Main Goals :

Photographs taken through glass windows often contain both the desired scene and undesired reflections. In this work, we use ghosting cues that arise from shifted double reflections of the reflected scene off the glass surface to exploit asymmetry between the transmission and reflections layers of the surface.

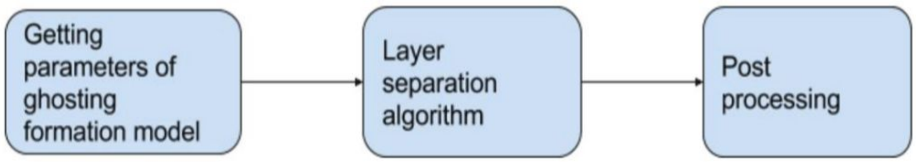
We model the ghosted reflection using a double-impulse convolution kernel, and estimate the spatial separation and relative attenuation of the ghosted reflection components. And the seperate it with Gaussian Mixture Model for regularization and removes a large fraction of reflections on both synthetic and real-world inputs.

(a) *Input* (b) *Close-up of ghosting* (c) *Recovered Transmission* (d) *Recovered Reflection*

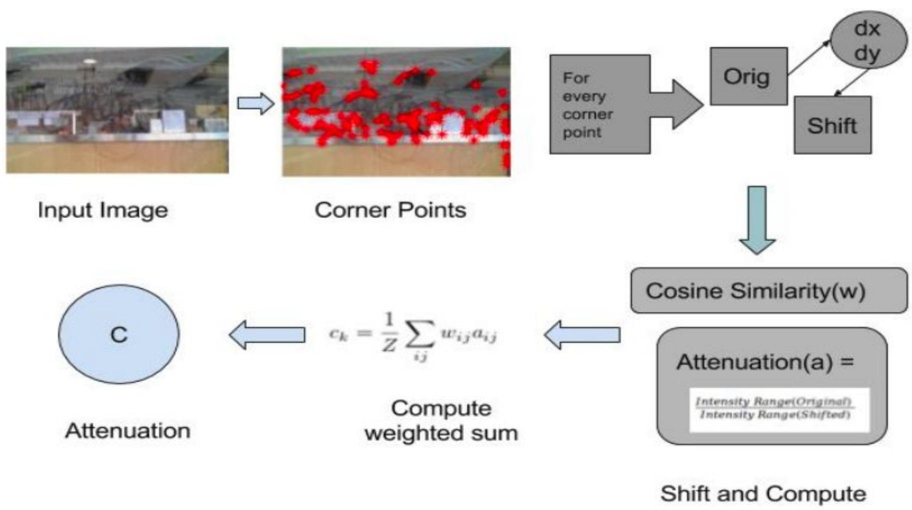
Problem Definition:

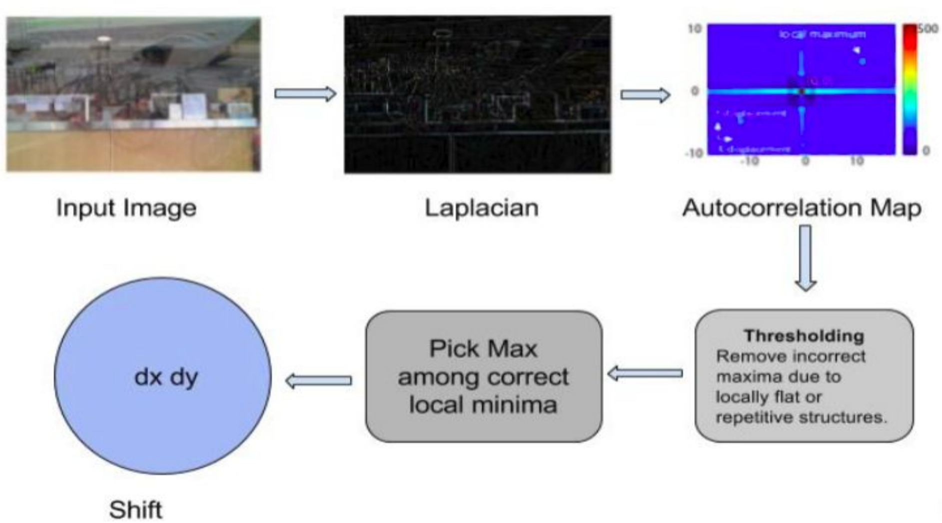
Ghosting provides a critical cue to separate the reflection and transmission layers, since it breaks the symmetry between the two layers. We model the ghosting as convolution of the reflection layer R with a kernel k. Then the observed image I can be modeled as an additive mixture of the ghosted reflection and transmission layers by R and T respectively:

$I = T + R \otimes k.$



Estimation of K :





Layer separation algorithm :

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

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

where n is additive i.i.d. Gaussian noise with variance σ^2. Given k, the above formation model leads to a data (log-likelihood) term for reconstruction of T and R:

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

However, minimizing L(T,R) for the unknowns T and R is ill-posed. Additional priors are needed to regularize the inference. The best performing priors are found using 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 GMM(P_iX) = PK j=1 π jN (P_iX; 0, Σj). The cost sums over all overlapping patches PiT in T, and PiR in R; where Pi 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 { π j}, and the covariance matrices by { Σ j}. So our final cost function with non-negativity constraint on T and R is:

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

The constraints for the above equation are per-pixel and per channel.

Separation of T and R and optimisation:

An optimisation scheme based on half quadratic regularisation method is used here. A set of patches z_iN 1, one for each overlapping patch P_ix in the image, yielding the following cost function:

$$c_{p,\beta}(\mathbf{x}, \{\mathbf{z}^{\mathbf{i}}\} \,|\mathbf{y}) = \frac{\lambda}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \sum_i \frac{\beta}{2}\left(\|\mathbf{P}_i\mathbf{x} - \mathbf{z}^i\|^2\right) - \log p(\mathbf{z}^i)$$

As β ➔ ∞ the patches P_ix are restricted to be equal to the auxiliary variables z_i and the solution of the following equation will converge.

$$f_p(\mathbf{x}|\mathbf{y}) = \frac{\lambda}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 - EPLL_p(\mathbf{x})$$

For a fixed value of β , optimizing 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.The following function will be the cost function for this problem.

$$\begin{aligned}
\min_{T,R,z_T,z_R} \; & \frac{1}{\sigma^2}\|I - T - R \otimes k\|_2^2 \\
& + \frac{\beta}{2}\sum_i \left(\|P_iT - z_T^i\|^2 + \|P_iR - z_R^i\|^2\right) \\
& - \sum_i \log(\mathbf{GMM}(z_T^i)) - \sum_i \log(\mathbf{GMM}(z_R^i)) \\
& \text{s.t. } 0 \le T, R \le 1
\end{aligned}$$

Starting with β = 200, and increase the value after each iteration. As β is increased, the values of P_iT and z_iT are forced to agree, similarly for the values of P_iR and z_iR. Value of σ = 5×10−3. We solve for T and R simultaneously by transforming the term || I−T− R⊗k ||^2 to || I−AX ||^2. Here X vertically concatenates vectors T and R, i.e., X = [T; R], and A horizontally concatenates the identity matrix I and convolution matrix k, i.e. A = [I_{Ik}]. An extended **L-BFGS** is used to handle box constraints. Since P_i contains only diagonal elements, and k contains only two non-zero entries for each pixel, the pixel domain L-BFGS solver is very efficient. Find component with the largest likelihood in the GMM model, and then perform Wiener filtering using only that component.

$$\hat{\mathbf{x}} = \left(\boldsymbol{\Sigma}_{k_{max}} + \sigma^2\mathbf{I}\right)^{-1}\left(\boldsymbol{\Sigma}_{k_{max}}\mathbf{y} + \sigma^2\mathbf{I}\mu_{k_{max}}\right)$$

A good initialization is crucial in achieving better local minima. The initialisation for the GMM-based model is done with a sparsity inducing based model, with a convex L1 prior penalty:

$$\min_{T,R} \quad \frac{1}{\sigma^2}\|I - T - R \otimes k\|_2^2 + \sum_j \|f_j \otimes T\|_1 + \sum_j \|f_j \otimes R\|_1$$

The L1 optimization can be efficiently performed using ADMM and the sparsity inducing filters $\{f_i\}$ are set that include gradients and Laplacians.

Results :

We evaluate our algorithm on both synthetic and real-world inputs, demonstrating how ghosting cues help in the recovery of transmission and reflection layers resulting image classification on the recovered transmissions and automated de-ghosting and test its performance from dataset.

We synthesize inputs from randomly sampled images for T and R, attenuation cK between 0.5 And 1.0, and ghosting distance dK between 4 to 40 pixels and try to achieve SSIM and PSNR for the transmission layer.

All the images are taken between 0.3 to 1 meter away from the window, and the angles between camera and glass range from 10 to 35 degrees. The glass thicknesses are between 4 and 12mm.

We then create images from each pair of transmission and reflection image. The ghosting kernel k is generated with a random attenuation ck between 0.5 and 1, and random shift dK generated by sampling a shift between -20 to 20 pixels in both the x and y directions. Then we compare recovered transmission layers to the input images.

Proposed Timeline :

10 OCT - 20 OCT :

Calculating dk with laplacian filter and working on the estimation of the kernel k using the parameters dk and attenuation factor ck.

21 OCT - 31 OCT :

Working on Layer separation algorithm using gosting kernel k and regularising interface with additional priors to minimize errors.

1 NOV - 14 NOV :

Working on optimisation of the algorithm using quadratic regularization and finding the function parameters.