SageRef: Single Image Reflection Removal

Pulkit Madaan (pmadaan2), Parthvi Mehta (pmehta14), Rishabh Nanawati (rnanawa1), Allan Raposo(Araposo1)

1 Problem Area

Photographing a scene through transparent media, most often glasses, is frequently hampered by reflections of objects on the camera's side. The transmission image, which is meant to be reflection-free but instead merges with the reflection image, is then captured as a mixture image. Several computer vision tasks, such as segmentation, classification, recognition, etc., become extremely challenging if not impossible due to the reflections, which produce the most unpleasant image degradations. The separation and elimination of the reflection image from the acquired mixture image is a complex image restoration operation out of necessity for a variety of significant applications.

The inadequacy of typical statistical models to distinguish between reflection and transmission images has significantly hampered the work of single-image reflection removal. On the other hand, while being visually disturbed, humans are capable of mentally separating the two images. The key distinction between humans and existing models is their ability to accomplish the separation job, which heavily relies on the coherence of high-level semantics of the two images. This shows that reducing reflections based on a single combined image is still a challenge and that machine learning is a sensible and effective solution. Also, the machine learning approach can avoid the difficulty of very inadequate conditioning in directly addressing the problem using the image formation model, in which the number of unknowns substantially exceeds the number of equations, by using appropriately chosen training data.

2 Context

Over the years, researchers have proposed various learning-based methods to address the problem of reflection removal in images. Many existing reflection removal methods rely on two or more input images of the same scene with different reflections to estimate the transmission layer. These methods rely on input images that have been captured sequentially (1) or using specific photography techniques such as employing a polarized filter (2) or flash lighting (3). While these methods have proven to be satisfactory, the real world may not permit the use of such specific techniques all the time.

The problem of removing reflections using single image input has proven to be rather challenging. Many proposed methods rely on some extra information provided by the camera (4; 5) or the user (6; 7). There have also been a few attempts to solve the problem without any user assistance. Sparse non-negative matrix factorization was used by (8) to separate the reflection layer without an explicit smoothness prior. Assuming that the image is smooth, (9) used two sequential CNN networks to reconstruct a non-reflective image from the edges of the input image. This technique performed well on synthetically generated data and was improved upon by several other researchers. (10) suppressed the reflection by turning it into an optimization problem with a Laplacian data fidelity term and a total variation term. This method seemed to perform better than the rest on synthetic data but failed to be satisfactory, just like others, on real-world data.

One may check out (11) and (12) for an excellent, comprehensive review of the most prominent reflection removal techniques.

3 Data

We will be utilizing several data sources for our project. Our first dataset is the CEILNet dataset, which includes 142 image sets of reflection and transmission images. The second dataset is SIR2 Benchmark Dataset which provides us with a collection of synthetic reflection images that have the image layers separated, along with the mixed layered image, providing us with 1500 images. In total, we will have 1642 images across all datasets. These images will be accompanied by reflection-free ground truth images and their reflected images which we will split into training, validation, and testing sets.

CEILNet Dataset: A dataset of images with reflection. https://github.com/fqnchina/CEILNet

The SIR2 Benchmark Dataset: A dataset of images with reflection. https://rose1.ntu.edu.sg/dataset/sir2Benchmark/

4 Proposed Solution

We plan to use a Denoising AutoEncoder (13) for our task, i.e. reflection detection and removal. We plan on using a deep learning paradigm where an encoder-decoder approach is used to reconstruct a reflection-free image. The method involves training an end-to-end model that has 2 components. The first is designed to extract features, and the second uses these features to recover & remove glare reconstruct refined final glare-free output. To achieve this we'll use Convolutional Neural Networks and skip connections as demonstrated in (14)'s work.

Inspired from (15)'s work the model will be optimized to minimize the L2 distance between the ground-truth glare-free image and reconstructed glare-free image along with the L2 distance between the layer representations from VGG (16) network of the ground-truth and reconstructed images to avoid only focusing on low-frequency features.

We plan on evaluating our model using two comparison metrics - peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). PSNR is defined as the ratio of the maximum possible power of a signal to the power of the signal that distorts it. It takes into factor MSE and the maximum pixel value of the image. SSIM is a perceptual metric that measures the similarity between two images by comparing their structural information in terms of their luminance, contrast, and structure. We will use PSNR and SSIM to compare the ground truth images in our dataset with the outputs of our model. Our aim will be to achieve the highest value for both the metrics.

References

[1] S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski, "Image-based rendering for scenes with reflections," *ACM Trans. Graph.*, vol. 31, jul 2012.

[2] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li, "Removing photography artifacts using gradient projection and flash-exposure sampling," *ACM Trans. Graph.*, vol. 24, p. 828–835, jul 2005.

[3] Y. Y. Schechner, J. Shamir, and N. Kiryati, "Polarization and statistical analysis of scenes containing a semireflector," *J. Opt. Soc. Am. A*, vol. 17, pp. 276–284, Feb 2000.

[4] P. Chandramouli, M. Noroozi, and P. Favaro, "Convnet-based depth estimation, reflection separation and deblurring of plenoptic images," in *Asian Conference on Computer Vision*, 2016.

[5] Y. Ni, J. Chen, and L.-P. Chau, "Reflection removal based on single light field capture," *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4, 2017.

[6] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, pp. 1647–1654, 2007.

[7] S.-K. Yeung, T.-P. Wu, and C.-K. Tang, "Extracting smooth and transparent layers from a single image," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, 2008.

[8] Y. Akashi and T. Okatani, "Separation of reflection components by sparse non-negative matrix factorization," in *Computer Vision and Image Understanding*, 2014.

[9] Q. Fan, J. Yang, G. Hua, B. Chen, and D. Wipf, "A generic deep architecture for single image reflection removal and image smoothing," 2017.

[10] N. Arvanitopoulos, R. Achanta, and S. Süsstrunk, "Single image reflection suppression," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1752–1760, 2017.

[11] R. Wan, B. Shi, H. Li, Y. Hong, L. Duan, and A. C. Kot, "Benchmarking single-image reflection removal algorithms," *IEEE Transactions on Pattern Analysis amp; Machine Intelligence*, vol. 45, pp. 1424–1441, feb 2023.

[12] A. Artusi, F. Banterle, and D. Chetverikov, "A survey of specularity removal methods," *Computer Graphics Forum*, vol. 30, 2011.

[13] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, p. 3371–3408, dec 2010.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2015.

[15] Z. Chi, X. Wu, X. Shu, and J. Gu, "Single image reflection removal using deep encoder-decoder network," *CoRR*, vol. abs/1802.00094, 2018.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings* (Y. Bengio and Y. LeCun, eds.), 2015.