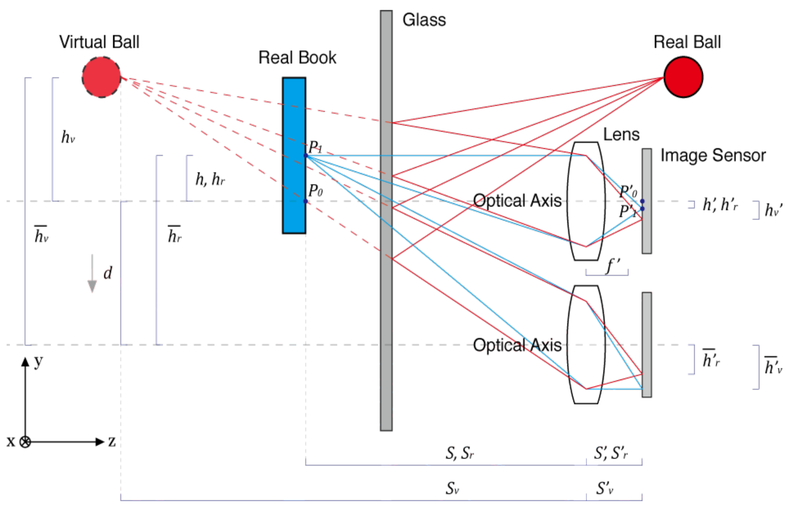
Homography-based Reflection Removal Specialized for Object Recognition

Poshen Lee
Electrical Engineering
University of Washington
185 Stevens Way Paul Allen Center
Seattle, WA
sephonlee@uw.edu

Richard E. Ladner
Computer Science & Engineering
University of Washington
185 Stevens Way Paul Allen Center
Seattle, WA
ladner@cs.washington.edu

Introduction

The camera on smartphones and tablets are frequently used to capture images for future recall or for further retrieval. For blind people the camera can be very useful in finding out about their surroundings by optical character recognition (OCR) or human interpreting services. Unfortunately blind people may not take the best pictures. In some cases, a strong reflective light may damage information on the target object. In this project, we propose a reflection removal method, demonstrated on Android-based mobile devices, specialized for object recognition aimed at digital displays, home appliance user interfaces, and documents with glossy surfaces.



$$\frac{d'_r}{d'_v} = \frac{d_r(S_v + f')}{d_v(S_r + f')} = \frac{\gamma S_r + f'}{S_r + f'} \approx \frac{\gamma S_r}{S_r} = \gamma$$

When the camera moves a distance d perpendicularly to the optical axis, the distance that an image moves depends approximately on the distance between the corresponding object and the camera. Given two images taken from different viewpoints, our method modifies one image by selecting pixels or segments with weaker reflection from the other image.

Summary & References

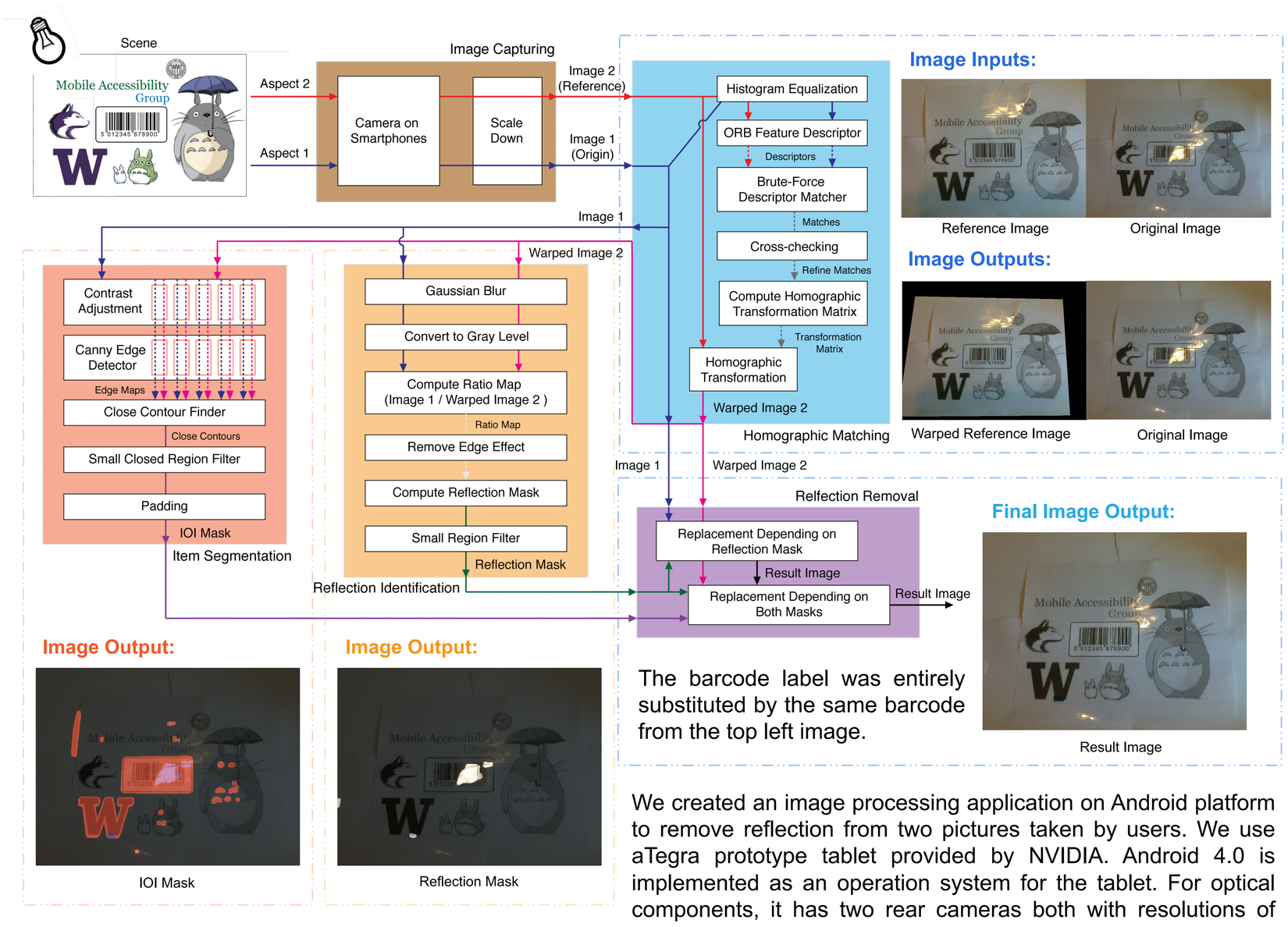
Summary

In this project, we have proposed a reflection removal method specialized for further object recognition aimed at appliance displays and documents with glossy surfaces. Our method works in a fully automated manner and without using any extra optical components on the Android platform. Given two images taken from different viewpoints, our method modifies the original image by selecting pixels or segments with weaker reflection from the reference image. From experimental results, our method effectively avoids ghost images caused by imperfect homographic transformation. We evaluated our method by using optical character recognition. Based on Levenshtein distance analysis, we obtained perfect OCR results in 59%, strong improvement in 16%, weak improvement in 20%, and deterioration in 5% of the cases.

References

[1] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh, “Vizwiz: nearly real-timeanswers to visual questions,” In Proc. UIST’10, 2010.
[2] Sudol, J.Dialameh, O., Blanchard, C. Dorcey, T., ”Looktel—A comprehensive platform for computer-aided visual assistance, “in Proc. IEEE Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 73 – 80, 2010.
[3] N. Ohnishi, K. Kumaki, T. Yamamura, and T. Tanaka, “Separating real and virtual objects from their overlapping images,” in Proc. European Conference on Computer Vision (ECCV), vol. 1065, pp. 636–646, 1996.
[4] Y. Y. Schechner, J. Shamir, and N. Kiryati, “Polarizationbased decorrelation of transparent layers: The inclination angle of an invisible surface,” in Proc. IEEE International Conference on Computer Vision (ICCV), pp. 814–819, 1999.
[5] H. Farid and E. Adelson, “Separating reflections from images by use of independent components analysis,” J. Opt. Soc. Am., vol. 16, pp. 2136–2145, 1999.
[6] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, “Sparse ICA for blind separation of transmitted and reflected images,” Int. J. Imaging Systems and Technology, vol. 15, no. 1, pp. 84–91, 2005.
[7] N. Kong, Y.-W. Tai, and S. Y. Shin, “A physically-based approach to reflection separation,” in Proc. IEEE Computer Vision and Pattern Recognition (CVPR), pp. 9–16, 2012.
[8] A. Levin and Y. Weiss, “User assisted separation of reflections from a single image using a sparsity prior,” In Proc. European Conference on Computer Vision (ECCV), volume 3021, pages 602–613, 2004.
[9] A. Levin and Y.Weiss, “User assisted separation of reflections from a single image using a sparsity prior,” IEEE Transactionson Pattern Analysis and Machine Intelligence (PAMI), 29:1647–1654, Sept. 2007.
[10] A. Levin, A. Zomet, and Y. Weiss, “Separating reflections from a single image using local features,” In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages I: 306–313, 2004.
[11] S. N. Sinha, J. Kopf, M. Goesele, D. Scharstein, and R. Szeliski, “Image-Based Rendering for Scenes with Reflections,” In Proc. ACM Transactions on Graphics - SIGGRAPH 2012 Conference, Volume 31 Issue 4, July 2012.
[12] V. N. Mahajan, Optical Imaging And Aberration Part I Ray Geometrical Optics, SPIE Press, 1998.
[13] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, “ORB: an efficient alternative to SIFT or SURF,” In Proc. IEEE International Conference on Computer Vision (ICCV), pp. 2564–2571, 2011.
[14] Adams, A., Talvala, E., Park, S. H., Jacobs, D. E., Ajdin, B., Gelfand, N., Dolson, J., Vaquero, D., Baek, J., Tico, M., Lensch, H. P., Matusik, W., Pulli, K., Horowitz, M., and Levoy, M., “The Frankencamera: an experimental platform for computational photography,” In Proc. ACM SIGGRAPH 2010.

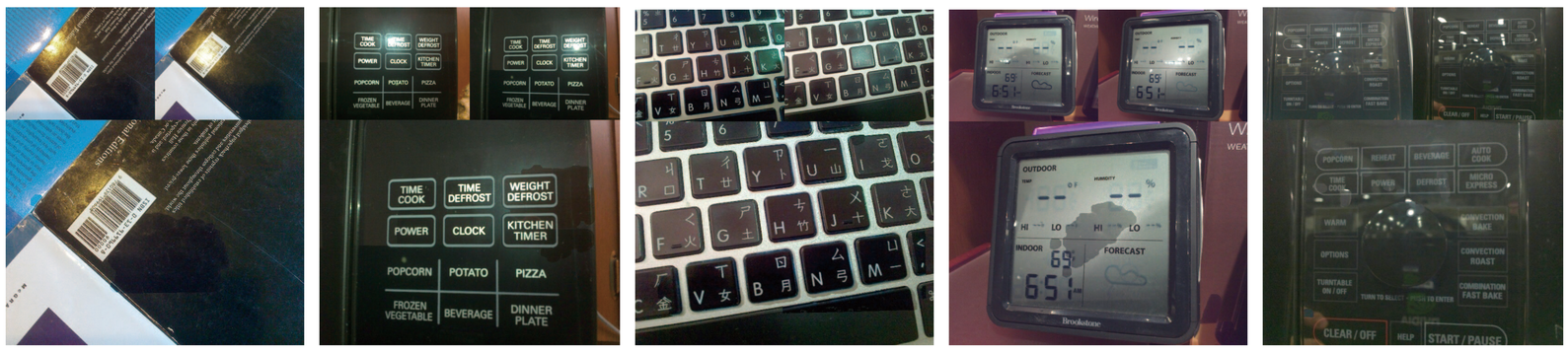
Overview



We created an image processing application on Android platform to remove reflection from two pictures taken by users. We use aTegra prototype tablet provided by NVIDIA. Android 4.0 is implemented as an operation system for the tablet. For optical components, it has two rear cameras both with resolutions of 500 million pixels and two front cameras both with resolutions of 100 million pixels.

Experimental Result

Qualitative Analysis

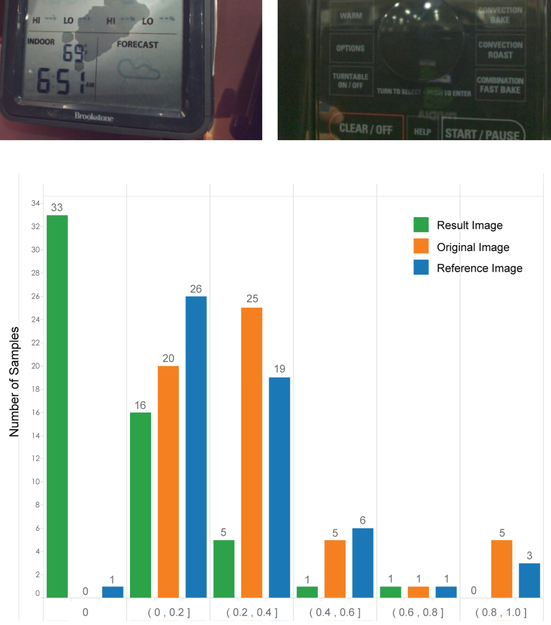


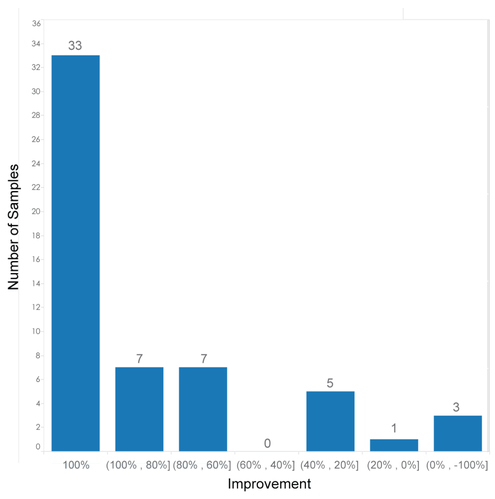
Quantitative Analysis

We focus on text documents with glossy surface and utilized the optical character recognition to show that our method improves poor object recognition caused by reflection for the reason that OCR is a mature and robust technique. We detected text targets from all the two input images and the result image by using Google tesseract-orc engine 3.02 with a GUI tool developed by VietOCR. Levenshtein distance (LD) was used to measure the minimum number of single-character edits required to modify the recognized result without regard to case. To understand the efficacy of our method, the improvement is defined by the following formula:

$$Improvement = \frac{LD(T_{original}, T_{correct}) - LD(T_{result}, T_{correct})}{LD(T_{original}, T_{correct})}$$

- **Perfectly Improved Recognition:** Improvement of 100%
- **Strongly Improved Recognition:** Improvement of 50% or more as well as LDPC of 5 % or less
- **Weakly Improved Recognition:** Improvement less than 50% and greater than 0%
- **Deterioration:** Improvement of 0% or less






Perfect Improvement


Strong Improvement. Error caused by discontinuity


Deterioration. Error caused by misalignment


Week Improvement. Error caused by misalignment

Research supported by the National Science Foundation Grant No. IIS-1116051