Mobile Scanner: Document Segmentation and Object Removal by Exemplar-based Image Inpainting

Jiafu Wu
Department of Electrical Engineering
Stanford University
jiafuwu@stanford.edu

Wenqi Hou
Department of Electrical Engineering
Stanford University
wenqihou@stanford.edu

***Abstract*— This proposal proposes a brief overview of the application of object segmentation and object removal techniques, who is capable to support an advantage mobile scanner functionality. It introduces various techniques in the fields of object removal techniques as well as the project overview and the tentative project schedule.**

I. Motivation

Nowadays, people need to process various images in their daily life, such as photos, receipts, written homeworks, printed documents, and name cards; nevertheless, in this digital age, an important method to process and store those paper documents is to capture them with cameras and store them via digital photos. In order to enhance readability, efficiency, and aesthetic feeling, often people needs to crop out the important document part and remove the unnecessary and unaesthetic background. Many mobile apps are made to help users automatically do that such as Evernote Scannable, iScanner, Tiny Scanner, etc. However, we can still include some features on top of them, for example, defect/undesirable object removal as well as noise reduction. As a result, in this project, we would like to achieve the scanner functionality of automatic document/photo extraction via edge detection and contour, and properly obtain a top-down, bird eye view image via perspective transform; also we would like to explore techniques for defect/object removal and de-noising. We may also a text recognition module when the target is document. A desire effect is shown on Fig 1.

(a) Original Image (b) Processed Image

Fig. 1: Before and after Our Process

II. Related Work

The first part of this project is edge detection. By definition, edge detection identifies points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities. The points at which image brightness changes sharply are typically organized into a set of curved line segments termed edges. By doing so, we can find a set of connected curves that indicate the boundaries of objects, and further locate the objects.

Most of edge detection methods can be grouped into two categories [5], search-based and zero-crossing based. The search-based methods detect edges by first computing a measure of edge strength, usually a first-order derivative expression such as the gradient magnitude, and then searching for local directional maxima of the gradient magnitude using a computed estimate of the local orientation of the edge, usually the gradient direction. The zero-crossing based methods search for zero crossings in a second-order derivative expression computed from the image in order to find edges, usually the zero-crossings of the Laplacian or the zero-crossings of a non-linear differential expression. As a preprocessing step to edge detection, a smoothing stage, typically Gaussian smoothing, is usually applied.

There has been significant development of object removal techniques in the past 20 years. Previously researchers proposed texture synthesis, which produced repetitive two-dimension textual patterns with introduced stochasticity to fill the occluded regions[1]. Out of these techniques, a particular effective one is exemplar-based techniques which introduce sampled and copied color values from the source [3]. However, while they

might work well to reproduce consistent repetitive textures, they have trouble to recover real-world scenes which often have non-repetitive textures.

In order to address this image filling issue for real-world scenes, researchers proposed inpainting techniques to fill holes in the images with isophotes, which are linear structures propagating in the target region via diffusion [4]. Such methods work finely as restoration techniques to remove small defects such as scratches and speckles, but blur would be introduced and become noticeable when filling larger regions.

Criminisi, Perez, and Toyama combined the exemplar-based and inpainting techniques into one single algorithm[2], which is a hybrid design aggregating the advantages of both methods. The linear structures using inpainting were used to determine the pixels values while the exemplar-based texture synthesis was used to determine the fill order of the color.

III. Project Overview

The goal for this project is to implement a refined scanner algorithm to recognize the area of interest in the image, transform the area to the center of image, process it by removing irrelevant objects or occlusion. First, we are going to apply an edge detection method to separate the area of interest from its background. Then we are going to perspective-transform the area to the center of the image. Based on the transformed image, we can apply various object removal methods on it to exclude irrelevant objects. We will also implement denoising methods to make it look more pleasing. At last, if the image is a document, we might consider creating a text recognition module to convert the content of the image to text. The project should be able to help users capture and process images or documents easily and nicely using a simple cell phone camera.

IV. Milestones

Time	Goal
Week of 2/12	Look for potential useful literatures, materials and documentations
Week of 2/19	Implement image segmentation, perspective transformation methods
Week of 2/26	Implement object removal methods
Week of 3/5	Text recognition and add additional features if possible
Day of 3/15	Work on poster presentation and report
Day of 3/17	Finalize project deliverables

TABLE I: Timeline

References

[1] M. Ashikhmin. Synthesizing natural textures. InProc. ACM Symposiumon Interactive 3D Graphics, pages 217–226, Research Triangle Park, NC,March 2001.

[2] A. Criminisi, P. Perez, and K. Toyama. "Object removal by exemplarbased inpainting." In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, vol. 2, pp. II-II. IEEE, 2003.

[3] A. Efros and W.T. Freeman. Image quilting for texture synthesis and trans-fer. InProc. ACM Conf. Comp. Graphics (SIGGRAPH), pages 341–346,Eugene Fiume, August 2001.

[4] C. Ballester, V. Caselles, J. Verdera, M. Bertalmio, and G. Sapiro. A vari-ational model for filling-in gray level and color images. InProc. Int. Conf.Computer Vision, pages I: 10–16, Vancouver, Canada, June 2001.

[5] D. Ziou and S. Tabbone (1998) "Edge detection techniques: An overview", International Journal of Pattern Recognition and Image Analysis, 8(4):537–559, 1998

[6] J. Canny. "A computational approach to edge detection." In Readings in Computer Vision, pp. 184-203. 1987.

[7] P. Bao, L. Zhang, and X. Wu. "Canny edge detection enhancement by scale multiplication." IEEE transactions on pattern analysis and machine intelligence 27, no. 9 (2005): 1485-1490.