Document Scanner and Perspective Correction
ECE501 - Digital Image Processing
Group-1

Members:

Name	Enrolment No.
Dhruvik Shah	AU2340113
Ayush Makwana	AU2340138
Divyam Thakkar	AU2340227
Jainik Patel	AU2340

1. Introduction

With the increasing use of mobile phones and digital cameras, capturing documents using handheld devices has become common. However, unlike flatbed scanners, photographs of documents often suffer from perspective distortion, uneven lighting, and background noise. Perspective distortion occurs because the document surface is not parallel to the camera sensor, causing the rectangular page to appear as a skewed quadrilateral in the image.

To address this, computer vision techniques are used to transform a photographed document into a clean, "scanned-like" image. The process typically involves detecting the document's boundaries, applying perspective correction, and enhancing the image for readability.

Input: Photograph of a document

Output: Flattened, cropped, and enhanced document image

2. Core Process

2.1 Contour Detection

- The first step is locating the document edges within the captured image.

- Methods like **Canny edge detection** followed by **contour approximation** are often used.

- The largest quadrilateral contour is usually assumed to be the document boundary.

2.2 Perspective Transformation

- Once corners of the quadrilateral are identified, the document is mapped to a rectangle using **homography transformation**.

- A homography is a projective mapping that can warp the image so that the distorted quadrilateral becomes a flat rectangle.

- This corrects perspective distortion, simulating the top-down view of a scanner

.

2.3 Cropping and Enhancement

- The transformed image is cropped to remove the background.

- Enhancements like **grayscale conversion, adaptive thresholding, and contrast adjustment** are applied to improve text clarity.

- The final result is a clean, flattened document image ready for further use (e.g., OCR).

3. Related Work

- **Boundary-based methods:** Use the rectangular boundary of the document to compute the homography

.

- **Text/layout-based methods:** Exploit text alignment, paragraph structure, or repeated patterns (e.g., tables) for perspective correction

.

- **Geometric methods:** Some approaches rely on orthogonality of document edges to restore the correct aspect ratio

.

Both categories demonstrate that effective perspective correction can be achieved either from the page geometry or from its content.

4. Applications

- Mobile document scanning apps (e.g., CamScanner, Adobe Scan).

- Digitization of books, receipts, and IDs.

- Preprocessing step for **OCR (Optical Character Recognition)**.

- Automated reading of signboards, license plates, and forms.

5. Conclusion

Document scanning with perspective correction bridges the gap between traditional scanners and portable devices. By combining contour detection with homography-based transformation, a simple photograph can be turned into a high-quality, flattened, and enhanced document image. This improves readability and enables downstream tasks like OCR.

6. Referenced Documents

https://www.researchgate.net/publication/221253834_A_new_method_for_perspective_correction_of_document_images

http://cvit.iiit.ac.in/images/ConferencePapers/2005/jagannathan05Perspective.pdf