**Report**

**Image pre-processing techniques**

Image preprocessing technique is a very crucial step in computer vision and image processing, and its main purpose is to eliminate irrelevant information in the image, recover useful and true information, and enhance the detectability of the relevant information, so as to improve the ability of the subsequent algorithms to identify, analyze and understand the image. (Cited)The following are some of the main techniques and methods of image preprocessing:

1. Grayscaling: converting a color image into a grayscale image to simplify the processing. A greyscale image requires only one byte per pixel to store the grey value, with a grey scale range of 0-255.
2. Geometric transformations: These include translation, rotation, scaling, etc., and are used to correct systematic errors in the image acquisition system and random errors in the position of the instrument. Geometric transformations usually require the use of greyscale interpolation algorithms such as nearest neighbor interpolation, bilinear interpolation, etc.
3. Contrast Adjustment: Improve the overall or local contrast of the image by methods such as histogram equalization to make the details of the image clearer.
4. Filtering: including spatial domain filtering and frequency domain filtering. Spatial domain filters, such as mean filter, median filter, Gaussian filter, etc., are used to smooth the image and remove noise; frequency domain filters, such as low-pass filters and high-pass filters, remove low-frequency or high-frequency noise in the frequency domain, respectively.
5. Sharpening: High-pass filters such as the Laplace operator, Sobel edge detection operator, etc. are used to highlight the edge information in the image and enhance the detail changes in the image.
6. Color space conversion: Converts an image from one color space to another, e.g. from RGB to HSV, greyscale space, etc. to facilitate subsequent feature extraction and recognition.
7. White balance correction: Ensures that the color of an image is not affected by the color temperature of the light source so that images taken under different lighting conditions have similar color performance.
8. Normalization: Adjusts the pixel intensity range to a specific interval, such as [0, 1] or [-1, 1], which helps algorithm training and comparison.

These techniques and methods can be selected and combined according to specific application scenarios and needs to achieve the best preprocessing results. In ASSIGNMENT I use greyscale and sharpening processing.

**Chosen text detection method and its working**

The working principles of OCR are (Amazon, 2023):

* Image pre-processing: Pre-process the input image, including greyscaling, binarisation, denoising, and other operations, to improve the accuracy of subsequent processing.
* Text region detection: Determine the regions in the image that may contain text through algorithms such as image analysis and edge detection.
* Text segmentation: Segment the characters in the text region for subsequent recognition of each character. Commonly used methods include segmentation based on connected regions, segmentation based on projection, segmentation based on edge detection, and so on.
* Feature extraction: feature extraction is performed on each character, and commonly used methods include projection method, template matching, neural network, and so on. The purpose is to convert the shape, texture, and other information of the character into a computer-recognizable feature vector.
* Character recognition: using trained models or algorithms, the extracted characters are recognized and converted into computer-recognizable character codes. Deep learning-based methods, such as Recurrent Neural Networks (RNN) or Long Short-Term Memory Networks (LSTM), are able to deal with text in different fonts, upper and lower case, rotation angles, and other complex situations.
* Post-processing: correct and amend the recognition results to improve the recognition accuracy. Commonly used methods include error correction, context correction, re-recognition, and so on.
* In summary, the OCR detection method converts the text in the image into editable text through a series of steps such as image pre-processing, text region detection, text segmentation, feature extraction, character recognition, and post-processing, which greatly improves the efficiency and accuracy of data processing.

**Measures and reports the speed of the text detection system**

The overall run time is less than 10 seconds, the longest of which is installing the identification tool Pytesseract and downloading the test images. This is a necessary step. Since I was using a MAC system, I used to download the test images for the entire program.

**Evaluates the accuracy of detected text**

A screenshot of a receipt

Description automatically generated

Figure 1compares the text

Text detection has low accuracy and requires more sophisticated detection and recognition.

**Discussion on the system's strengths and weaknesses**

The advantages of OCR (Optical Character Recognition) detection methods are (Tripathi, 2023):

1. Improve efficiency: OCR technology can automatically convert the text in the image into editable text, which greatly improves the speed of data entry and processing and reduces the time and error rate of manual input.
2. Widely used: OCR technology is widely used in ID cards, driver's licenses, invoices, and other document recognition, as well as in the processing of medical bills, medical records, prescriptions, and other documents, providing convenience for all walks of life.
3. Technological advancement: With the continuous development of computer vision, pattern recognition, machine learning, and other technologies, the accuracy and application scope of OCR is also improving and expanding.

The disadvantages of OCR (Optical Character Recognition) detection methods are (kalyan, 2022):

1. Recognition accuracy is limited: The recognition accuracy of OCR technology is affected by a variety of factors, such as fonts, font size, printing quality, image quality, etc.. Especially for handwritten fonts or scribbled handwriting, the recognition effect of OCR technology may be greatly reduced. In addition, OCR may also have difficulty handling documents in complex formats such as tables, graphs, and handwritten annotations.
2. Higher cost: OCR technology requires high-quality scanning equipment and software support, which is relatively costly. Also, the development and maintenance of OCR systems require specialized technical staff.
3. Risk of privacy leakage: OCR technology needs to pay more attention to security when handling sensitive information to avoid the risk of privacy leakage.
4. Operational complexity: In some cases, OCR operations may be relatively complex and require multi-device support. At the same time, the setup and adjustment of the OCR system may also require certain technical knowledge.

The drawback of the OCR detection method in this program is the lack of accuracy. In summary, OCR technology has the advantages of significant work efficiency improvement, accuracy improvement, convenient archiving, and retrieval, etc., but at the same time, there are also shortcomings such as limited recognition accuracy, high cost, risk of privacy leakage, and operational complexity. (kalyan, 2022)In practical applications, it is necessary to weigh the pros and cons according to specific scenarios and needs and choose a suitable OCR solution.

**Reference**

1. **Cited J Verne**. *Image Pre-Processing*.
2. **9 Key Advantages of OCR-Based Data Entry**. *Tripathi, Pankaj*. 2023.

<https://www.docsumo.com/blog/ocr-advantages> [Available at Sep 12, 2024]

1. **What is OCR (Optical Character Recognition)?** *Amazon Web Services*. 2023.

<https://aws.amazon.com/what-is/ocr/#:~:text=The%20OCR%20software%20uses%20pattern,is%20called%20optical%20word%20recognition> [Available at Sep 12, 2024]

1. **Advantages and Disadvantages of Optical Character Recognition (OCR)**. *Kalyan. 2022.*

<https://makemeanalyst.com/advantages-and-disadvantages-of-optical-character-recognition-ocr/> [Available at Sep 12, 2024]