Research on Synchronized Recognition System of Literary Works Based on Tesseract-OCR Model

Zilan Zhou
Wuchang Institute of Technology, Wuhan, Hubei, 430223, China

Abstract. Using mobile terminal for image and character recognition can solve the dilemma of the traditional large-scale scanning equipment, such as complicated steps and inconvenience in moving, and provide great convenience for mobile learning and office. As a direct and concrete manifestation of objective things, images are rich in information, which is incomparable in other ways. Therefore, the importance of images is self-evident. Recognition of graphics and text information of terminal equipment can solve the disadvantage of large-scale equipment, such as difficult to move and complicated use steps, and can make office and mobile learning more smoothly. Tesseract-OCR is powerful in open source OCR. Based on Tesseract-OCR model, this paper puts forward the corresponding improvement scheme for the defects such as irregular graphics and characters, and inconspicuous image and text recognition system, and carries out synchronous preview and recognition of image and text as much as possible. The graphic publication of literary works reflects readers’ impetuous reading mentality, the popular trend of literature trend and the management crisis of publishing houses.

1. Introduction
The graphic publication of literary works reflects readers’ impetuous reading mentality, the popular trend of literature trend and the management crisis of publishing houses. With the development and popularization of smart phones, mobile learning has become a fashion, and the combination of smart phones and graphic recognition software has become a new application demand [1]. Using mobile terminals for image character recognition can solve the dilemma of complicated steps and inconvenient movement of traditional large-scale scanning equipment and provide great convenience for mobile learning and office work [2]. Images are the most direct manifestation and intuitive perception of communication with the outside world in human life. As a literary form relying on words, the graphic publication of literary works has a certain negative impact on the development of literature, specifically manifested in the weakening of the written status of the works, the decline of readers’ aesthetic imagination and the differentiation of reader groups [3]. How to make use of Android system to research and develop software has gradually become a problem that workers in related fields need to study and think about. The system architecture and development mode will have a direct impact on the final application results of the system [4]. This paper implements the application of the literary artwork synchronization recognition system based on Tesser-act-OCR engine under Android platform.

Human images are in constant contact with the image. According to statistics, a person’s visually captured image information accounts for more than 70% of all information received by a person [5]. The identification of the information on the image, that is, optical recognition, that is, OCR recognition. In the open source OCR, Tesseract-ocr is more powerful. Sounds and images, especially the latter, organized aesthetics and led the audience. The direct concrete embodiment of images as objective things is rich in great information, which is incomparable to other methods, so the importance of images is
self-evident [6]. Contemporary culture is becoming a visual culture, not a printing culture. At present, some existing recognition systems for literary works are all developed for a specific literary work, which cannot be used for text positioning [7]. Identification of graphic and text information of terminal equipment can solve the disadvantages of large-scale equipment such as difficult movement and complicated use steps, and make office and mobile learning more smoothly [8]. This paper proposes an intelligent automatic recognition algorithm for literary works based on image binarization and jump detection to locate the text area, and then combines the neural network with Tesseract-ocr for text recognition.

2. Technical Essentials

The traditional reading method of pure text books has been gradually abandoned by ordinary readers, and become the insistence of academic researchers in the Ivory Tower. How to realize the transformation from pure text to image has become an important problem faced by writers and publishing houses. If we want to interpret the objective world with machine intelligence, it is necessary to acquire and process images. In the upsurge of graphic publishing, some functional and fashionable graphic documents show directly the contents closely related to real daily life in the way of mutual interpretation. For mobile applications, the first impression of user interface design is that the effect of interface design will directly affect user selection. Android tools provide a direct and simple design method and interface layout during the development period. The synchronous identification module realizes continuous synchronous identification of the preview image, and the user can see the identification effect immediately while previewing the shot image.

The networking identification module realizes the server-side identification interface and provides identification support for mobile terminals with low performance. The bottom layer uses Google's open source recognition engine Tesseract-OCR to recognize the processed images and return the recognition results to the UI interactive interface. For signals, due to band limitation, signal energy is mainly distributed in the low frequency region. Therefore, for noisy signals, the proportion of noise energy is small in the low frequency region and large in the high frequency region. Therefore, the focus of denoising should be on the high frequency region. Fig. 1 is an image analysis system structure.

![Image analysis system structure](image.png)

At the same time that literary works moved toward film and television adaptation, some paper books were published in a combination of graphics and texts, which led to a wave of graphic publishing. Considering the popularity of mobile intelligent terminals, the built-in camera function and graphic image processing technology have developed corresponding applications, which will inevitably have a better experience and a more intuitive experience. For the characteristics of images captured by mobile terminals such as mobile phones, the image is enhanced to improve the recognition rate of the recognition engine and make the image text easier to recognize [9]. After binarization of the picture, it
is necessary to locate the picture individually. Using the jumping information of pixel points, a distance threshold is taken to connect the points in the same row where the distance between two adjacent jumping points is less than the threshold to form a white line segment. Until now, the core function of the mobile terminal has not changed and is still instant messaging. This also means that although the mobile terminal integrates many functions, it is still far from professional equipment. Only by finding a balance between software and hardware and making rational use of various resources can the impact of these restrictions be reduced.

3. Requirement Analysis and Overall Design

3.1. Demand Analysis
The icon module will pre-process the image, locate and segment the area where the icon is located, and confirm the icon sequence by confirming the single icon through image analysis. The introduction of pure text production methods is difficult to form a finished image in the reader's mind, and the exquisite pictures on the one hand well explain the rationality and reliability of the production method, on the other hand stimulate the reader's desire to purchase literary works. Before the corresponding recognition and processing of the text image, it is necessary to be able to pre-process the image, and to eliminate as much as possible the influence of various adverse interference factors on the recognition result [10]. When the user does not stop the recognition mode or jump to another interface, the mobile phone will continuously collect image data for processing and recognition until the user is satisfied with the recognition result. Image enhancement is to distort the source image and transform the original clear image into an image that is easier for human or computer analysis and recognition. In this process, some local features of the image will be more distinct and prominent, and some unnecessary information will be weakened or removed at the same time.

During the automatic data acquisition by the mobile phone, the user can focus at the specified point by touching the screen of the viewing frame to acquire clearer and higher contrast processed images. With the increase of the signal amplitude, the estimated signal can reach and exceed the real value. It also has a certain enhancement effect on the radial component with larger energy, and can effectively retain the edge and contour information of the image. The comparison between this function and other contraction functions is shown in Fig. 2.

![Fig. 2 Comparison with other contraction functions](image)

3.2. Implementation of the System
In the process of image enhancement, there are often some contradictory processes, that is, on the one hand, image enhancement wants to enhance edges, on the other hand, it wants to reduce noise. For some images, some parts of the target object exceed the boundaries of the image. No matter how the initial curve is placed, it cannot meet the requirements of surrounding the target object or being located inside or outside the target object. The performance parameters of image segmentation before and after optimization are shown in Table 1. After wavelet transform and filtering optimization, the topology of
image segmentation is greatly optimized, with fewer nodes and better monitoring area. The simulation comparison of topology reliability optimization for image segmentation is shown in Fig. 1.

Table 1 Performance parameters of image segmentation image structure before and after optimization

|                        | Before optimization | After optimization |
|------------------------|---------------------|--------------------|
| The number of rows     | 112                 | 143                |
| Number of columns      | 37                  | 48                 |
| Monitoring points      | 4144                | 6864               |

The next level is the functional module layer of the system, which is responsible for the processing of business logic, including synchronization, batch and network identification. Synchronized recognition mode includes synchronous and continuous recognition of preview pictures and texts. Users can see the corresponding recognition effect in the process of preview of captured pictures and texts. When the user chooses the recognition mode of single photograph, the focus of the designated point is completed by touching the viewfinder frame, and the data of the processing image is obtained by pressing the photograph button. Since the recognition of a single shot image does not require continuous acquisition of image data, the sequence of events from image acquisition to recognition is complete, so only a single thread is required to complete the operation. Images processed in image segmentation can be gray images and color images, but color images are often converted into gray images for post-processing. After clicking on the software icon, you will be transferred to the function selection system to better complete the real-time and batch identification work and read the documents comprehensively and systematically.

4. Conclusion
Coding design corresponds to decoding and playing, and the corresponding relation between page number and video is specified. After obtaining image features, video is played according to these corresponding relations. The increase of humanized experience has significant market value. The system realizes the requirements of character recognition and processing of pictures and texts. The system basically meets the needs of library to look up data to identify and record scenes, and realizes the mobile phone character enhancement recognition algorithm based on Tesseract-OCR, which can meet the needs of mobile learning users and increase more humanized experiences. Page number recognition is fully combined with image processing technology and image characteristics, and requires image processing operations such as enhancement, binarization, correction, etc. to finally extract the digital domain. Computer software can automatically input the information of literary works into database and literary works management system, which greatly reduces the difficulty of input work and saves time and cost. The publication of pictures and texts of literary works has become an irresistible trend in today's society. Writers should continue to write more works full of artistic imagination, and readers should reserve a certain reading time for pure literary works. The image character recognition function and image processing method implemented by the system still have some defects. Continuing to improve the system function and optimize the code quality is the next research direction of the system.

References
[1] Morales Y, Satake S, Kanda T, et al. Building a Model of the Environment from a Route Perspective for Human–Robot Interaction[J]. International Journal of Social Robotics, 2015, 7(2):165-181.
[2] Spitsyn V G, Bolotova Y A, Phan N H, et al. Using a Haar wavelet transform, principal component analysis and neural networks for ocr in the presence of impulse noise[J]. Computer Optics, 2016, 40(2):249-257.
[3] Lee H J, So H, Son S, et al. Energy-efficient adaptive optical character recognition for wearable devices[J]. Electronics Letters, 2016, 52(2):113-115.
[4] Chapman M P, Moore E E, Burneikis D, et al. Thrombelastographic pattern recognition in renal
disease and trauma[J]. Journal of Surgical Research, 2015, 194(1):1-7.

[5] Yang S, Deravi F. On the Usability of Electroencephalographic Signals for Biometric Recognition: A Survey[J]. IEEE Transactions on Human-Machine Systems, 2017:1-12.

[6] Miwa K, Libben G, Dijkstra T, et al. The time-course of lexical activation in Japanese morphographic word recognition: Evidence for a character-driven processing model[J]. Quarterly Journal of Experimental Psychology, 2014, 67(1):79-113.

[7] Hausknecht R, Jacobs S, Mueller J, et al. Phylogeographic analysis and genetic cluster recognition for the conservation of Ural Owls (Strix uralensis) in Europe[J]. Journal of Ornithology, 2014, 155(1):121-134.

[8] Huang T, Chang Y C, Hsieh F F. Incomplete neutralization of sibilant consonants in Penang Mandarin: A palatographic case study[J]. The Journal of the Acoustical Society of America, 2014, 136(4):2173-2173.

[9] Wang A L, Lowen S B, Romer D, et al. Emotional reaction facilitates the brain and behavioural impact of graphic cigarette warning labels in smokers[J]. Tobacco Control, 2015, 24(3):225-232.

[10] Allen S, Rastner P, Arora M, et al. Lake outburst and debris flow disaster at Kedarnath, June 2013: hydrometeorological triggering and topographic predisposition[J]. Landslides, 2015, 13(6):1-13.