Research on Recognition Method of Test Answer Sheet Based on Machine Vision

Ping Cui[1], Dan Li[1(✉)], Kailiang Zhang[1], Likai Wang[2], and Weiwei Liu[2]

[1] Xuzhou University of Technology, Xuzhou 221000, Jiangsu, China
[2] Traffic Police Detachment of Xuzhou Public Security Bureau, Xuzhou 221000, Jiangsu, China

Abstract. When using the cursor reading technology to mark the answer card, the cursor machine can only be used for special card, which is expensive and difficult to popularize. A new method of answer sheet recognition based on machine vision and image processing was proposed. Firstly, the improved curvelet algorithm was used to preprocess the image to solve the problem of low resolution and high noise caused by different acquisition methods. Secondly, Hough transform was used to detect lines and correct deformation of binary image. Finally, the answer area was segmented, and the vertical and horizontal projections were used to detect the question and option interval, generate grid lines, mark the center of rectangle and judge the option results. Experiments show that this method is accurate, efficient and robust to low resolution, tilt and noise.

Keywords: Answer sheet recognition · Automatic marking · Curvelet algorithm · Hough transform

1 Introduction

With the rapid development of computer science, cursor reading technology [1, 2] has been widely used in large-scale examinations. It is very convenient for objective questions using the cursor reader to avoid the problem of low efficiency and easy to make mistakes compared with manual marking. However, there are many problems that are difficult to be solved in the process of using them: 1) the answer card must be unified, and it is a special-purpose machine for special cards; 2) it transforms the light signal of filling position into electrical signal. When the color is light, the effect is poor; 3) the paper of the answer card has strict requirements in the printing process, the paper must be smooth, thin and thick, uniform, foldable and defaced, and 2B pencil must be used.

Aiming at the problems of low efficiency of traditional marking and special card for cursor reader, which are expensive and difficult to popularize, this paper combines machine vision [3, 4]and digital image processing technology [5–7]to study the preprocessing, feature extraction and recognition judgment technology. In order to make the automatic marking system more flexible and convenient for network transmission [8–10], it should be adaptive to the complex environment such as angle tilt, different style, low resolution, noise and so on.

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2021
Published by Springer Nature Switzerland AG 2021. All Rights Reserved
H. Song and D. Jiang (Eds.): SIMUtools 2020, LNICST 370, pp. 702–714, 2021.
https://doi.org/10.1007/978-3-030-72795-6_56

2 Image Processing Based on Curvelet

After image acquisition, it is easy to have noise, which greatly affects the subsequent image recognition [11–13]. In the preprocessing, the denoising algorithm can improve the image recognition degree and facilitate the later image recognition. E. J. Candes and D. L. Donoho first proposed the second generation Curvelet transform theory based on frequency domain [14–18], which has high approximation accuracy and sparse expression [19–21] ability. The noise and image edge information can be separated well, and the implementation is simple and the redundancy is reduced. It can not only keep the edge, but also suppress the noise well.

In this paper, the improved curvelct transform [22, 23] is used to denoise the answer card image, improve the signal-to-noise ratio and transmission rate [24–26] of the image, and obtain a clearer display effect. Applying curvelet transform to image processing requires its discrete form. Let $f[t_1, t_2]$,$0 \le t_1$,$t_2 \le n$ in Cartesian coordinate system be input, its discrete form as is shown in formula (1), and the information of scale, direction and position are respectively represented by j, k and l.

$$c^D(j,l,k) := \sum_{0 \le t_1, t_2 < n} f[t_1, t_2] \overline{\varphi^D_{j,l,k}[t_1, t_2]} \tag{1}$$

$$\Psi(\omega_1) = \sqrt{\phi(\omega_1/2)^2 - \phi(\omega_1)^2} \tag{2}$$

The above formula uses functions $\Psi(\omega_1)$ to realize multi-scale segmentation, which is defined as follows:

$$\Psi_j(\omega_1) = \Psi(2^{-j}\omega_1) \tag{3}$$

For each, $\omega = (\omega_1, \omega_2)$, $\omega_1 > 0$, the corner window is as follows:

$$V_j(S_{\theta_1}\omega) = V(2^{j/2}\frac{\omega_2}{\omega_1} - l) \tag{4}$$

$$S_{\theta l} := \begin{pmatrix} 1 & 0 \\ -\tan\theta_l & 1 \end{pmatrix} \tag{5}$$

In formula (5), where $S_{\theta l}$ is a shear matrix. Defining $\widetilde{U}_j(\omega) := \psi_j(\omega_1)V_j(\omega)$, for each $\theta_1 \in [-\pi/4, \pi/4)$, it has $\widetilde{U}_{j,1}(\omega) := \psi_j(\omega_1)V_j(S_{\theta_1}\omega) = \widetilde{U}_{j,1}(S_{\theta_1}\omega)$.

Figure 1 below shows the spatial and frequency domain map of scale j.

In this paper, the wrapping curve transformation method [27, 28] combined with cycle spinning [29–31] is adopted. The specific steps are as follows:

Step 1: The image with noise is cyclically translated, assuming that the amount of translation is (i, j), $i \in (0, M)$, $j \in (0, N)$, the width and height of the image are M and N respectively. The row and column are translated at the same time, and the denoising results are linearly averaged to get the image $\hat{Y}(i, j)$, as shown in Eq. (6) below.

$$\hat{Y}(i,j) = \frac{1}{K_1 K_2} \sum_{i=1, j=1}^{K_1 K_2} Y_{-i,-j}(F^{-1}(T[F(Y_{i,j})])) \tag{6}$$

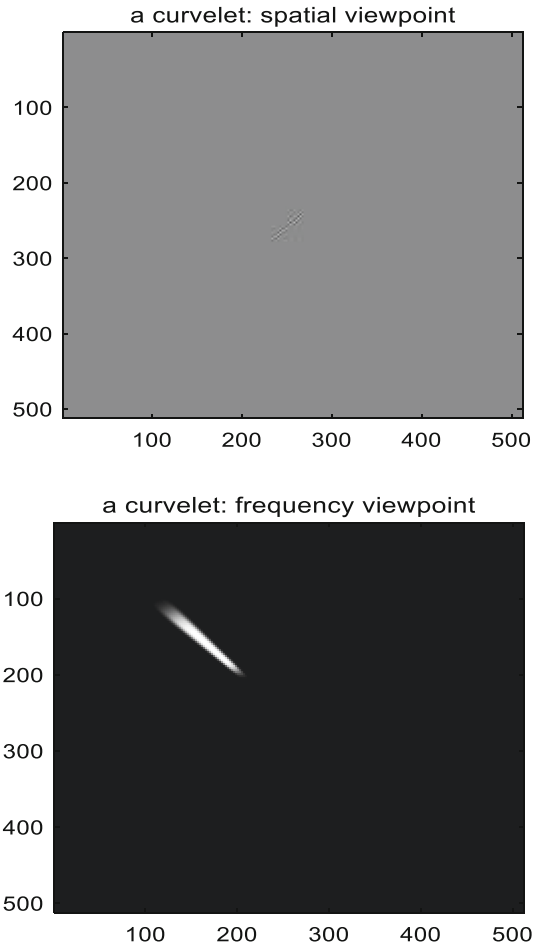


Fig. 1. Spatial and frequency domain map

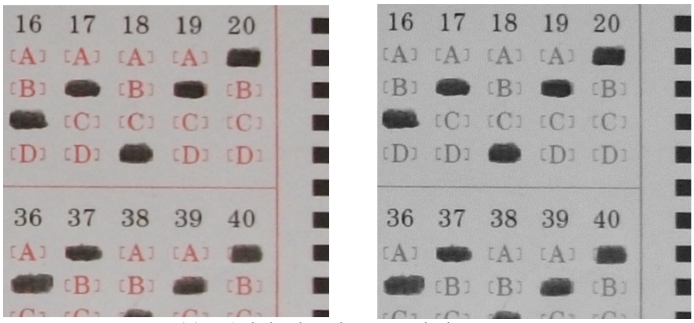
F represents wavelet transform, F^{-1} represents inverse transformation, T is the threshold, K_1, K_2 are the maximum translations of row and column respectively,$Y_{i,j}$ is the panning image, and $Y_{-i,-j}$ is the reverse panning operation after image denoising.

Step 2: The curved transform based on wrapping is used to process the translated image, and the set $C^D(j, l, k)$ of discrete curved coefficients of each scale and direction is obtained.

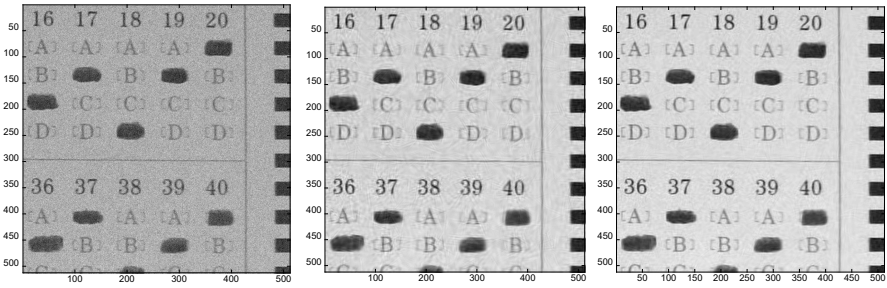
Step 3: The polynomial interpolation threshold function is used to denoise the curvelet coefficients of different scale and direction subbands.

Step 4: The curvelet algorithm performs translation operation after inverse transformation, reconstructs the image, and then obtains the denoised image.

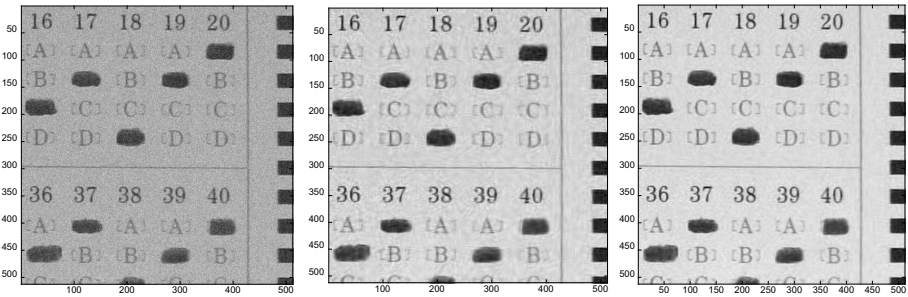
Step 5: Repeat the above steps for the image after inverse cycle translation denoising. Finally, average the results of multiple iterations to obtain the final denoising result image $I(i, j)$.



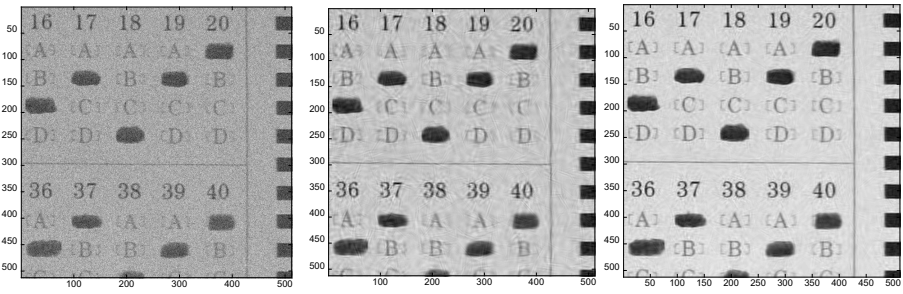
(a) Original and grayscale images



(b) Standard deviation 10, wrapcurvelet and this paper



(c) Standard deviation 20, wrapcurvelet and this paper



(d) Standard deviation 30, wrapcurvelet and this paper

Fig. 2. Comparison of image denoising algorithms of answer card

In order to test the superiority of the denoising algorithm in this paper, the gray-scale image of some areas of the answer card with the size of 512 × 512 is selected for the experiment. Fig. 2 is a comparison of wrapcurvelet algorithm and the improved algorithm when the standard deviation of noise is 10, 20 and 30.

In Fig. 2, a is the original image and the gray-scale image, b, c and d are the comparison of the denoising results of different algorithms under the condition of 10, 20 and 30 noise standard deviation respectively. From the perspective of visual effect, with the increase of noise, the wrapcurvelet algorithm appears virtual shadow and image blur after noise reduction. After noise reduction, the image is clear, which can recover the image texture well and retain the image edge information well. It still shows good advantages when the noise standard deviation increases.

3 Hough Algorithm for Deformation Correction

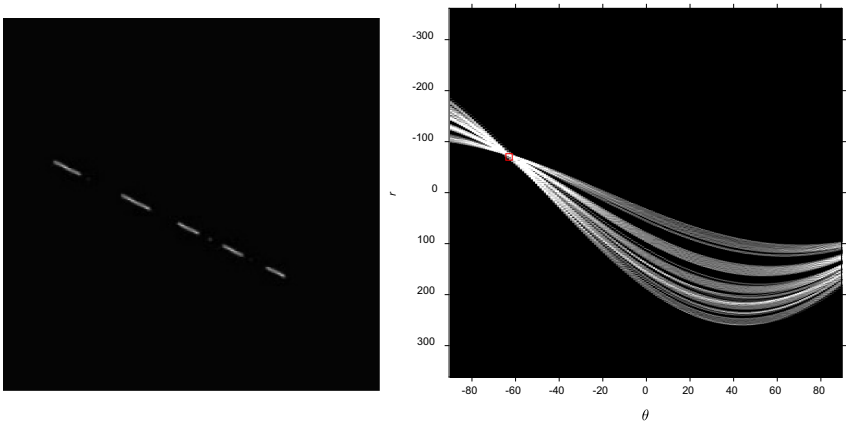
In this paper, Hough parameter space transform optimization algorithm [32–34] with good anti-interference [35, 36] is used to detect straight lines, correct image tilt, and support parallel computing.

The general expression for a line is $y = ax + b$. In Hough transformation, the line is $r = x\cos(\theta) + ysin(\theta)$, It can also be expressed as:$r = \sqrt{x^2 + y^2}sin(\theta + \varphi)$, $\tan(\varphi) = x/y$. Where r is the distance between the line and the origin of the upper left corner of the image, and the angle between the line and the perpendicular is θ. A line in x, y coordinates is mapped to a point in r, θ space; any point in x, y coordinates is mapped to a sine curve in r, θ space. Hough transform can not only detect the continuous line, but also detect the discontinuous line, so it is not affected by the linear fracture.

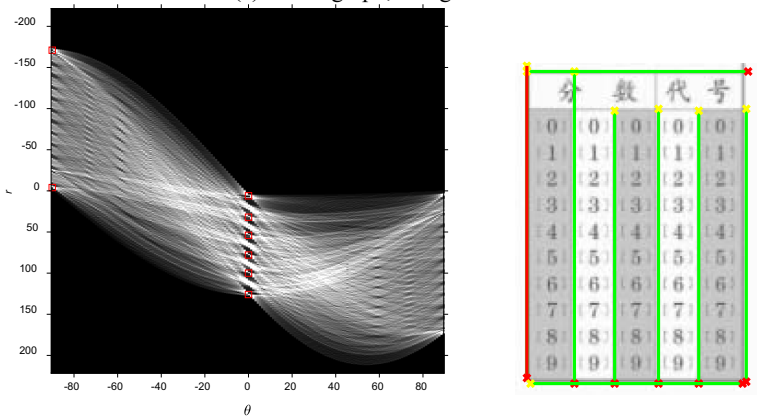
In Fig. 3, a is a number of line segments on the same line in x, y coordinates. The right figure is the Hough transform domain of these line segments in r, θ coordinate space. It can be seen that all the line segments intersect at the same extreme point in r, θ coordinate space, which means that these line segments are on the same line, and the detection of the line can be realized by accumulating and counting the points on the binary image with the accumulator. In Figure b, Hough transform is used to detect the fractional code graph. The first 8 peaks in the transform domain are selected, and then 8 extreme points are detected in the transform domain. The first 8 lines are shown in the right figure, and the longest line is the left vertical line.

Through Hough transform detection, the straight line in the answer sheet can be detected to correct the tilt angle. The center position of the image is taken as the rotation center for calculation. Suppose that the coordinate of point (x_0, y_0) after turning θ degree around point (s,t) is (x_1, y_1), and the central coordinate after rotation is (m, n), then

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & m \\ 0 & -1 & n \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -s \\ 0 & -1 & t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \tag{7}$$



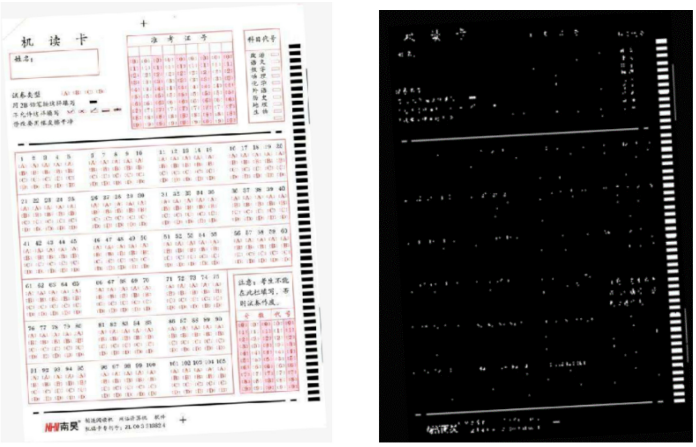
(a) Line graph, Hough transform domain



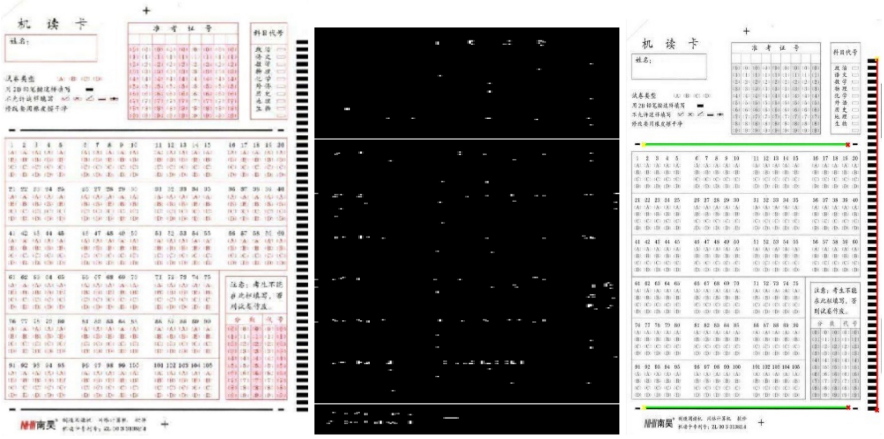
(b) Hough transform domain and detection results

Fig. 3. Hough transform domain line detection

In Fig. 4, a is the original image and the binary image before the tilt correction, b is the corrected image and the binary image. Using Hough transform, when the number of extreme points in the transform domain is 3, the straight line detection result is obtained. In the detection image, the vertical line on the right is the longest, which can be used to adjust the tilt angle. The other two lines are the horizontal lines that separate the information area and the objective question area of the examinee. You can use the binary image line length of Hough detection to arrange in descending order. The two lines with the longest horizontal direction to get the area segmentation position, so as to realize the segmentation of the examinee's information area and the objective question area.



(a) Original image before tilt correction, binary image



(b) After correction, binary graph and Hough transform detection

Fig. 4. Image tilt correction of answer sheet

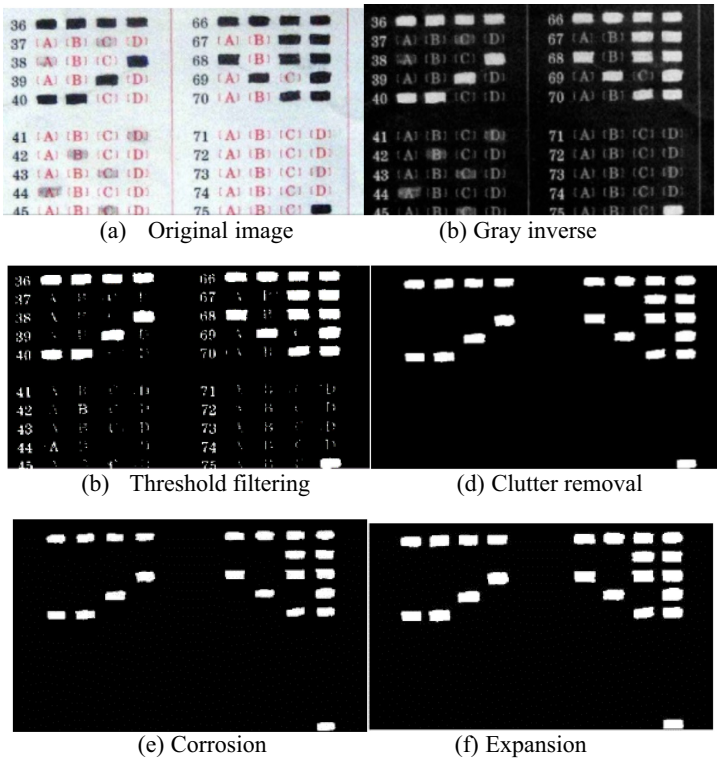
4 Option Identification and Judgment

After the correction of the answer card image, it is necessary to locate and segment different areas, complete the detection of the smear area of the answer card options and the identification and judgment of the answers.

4.1 Impurity Removal

The connected area of smear information in the binary image of the answer card needs to calculate the number of pixels in each area. The small noise points contain fewer pixels. Set the threshold T. If the connected area is smaller than T, discard it. The pixels larger than the threshold T are effective pixels. In addition, the burr on the edge of the

filling area is further removed by morphological corrosion operation [37–41], and then the edge and internal holes are filled by expansion operation to smooth the filling area.



(a) Original image (b) Gray inverse

(b) Threshold filtering (d) Clutter removal

(e) Corrosion (f) Expansion

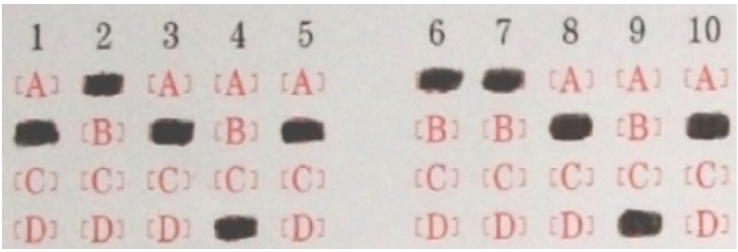
Fig. 5. Noise removal

In Fig. 5, a and b are the original image and gray-scale inverse image of part of the answer area. Through threshold filtering, the background information, separation vertical line, question number and option information with lighter gray level are removed, as shown in c. If the threshold value is increased, the information such as question number can be further removed, but the pencil smear area will reduce the area and affect the recognition. Therefore, through the above method of removing small area speckles, the connected area with small pixel area is removed while the complete smear information is retained, as shown in d. After further corrosion and expansion, as shown in e, f, the edge of the coating area is smoother.

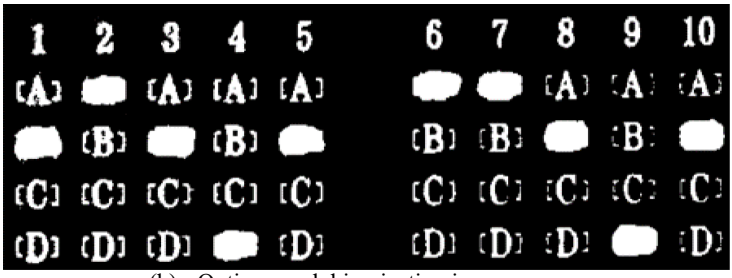
4.2 Option Positioning and Marking

Generally, the options of the answer card are regular and have the same interval, so the answer area of the answer card can use the horizontal or vertical projection curve of the gray image to determine the distance between the option areas of each row and

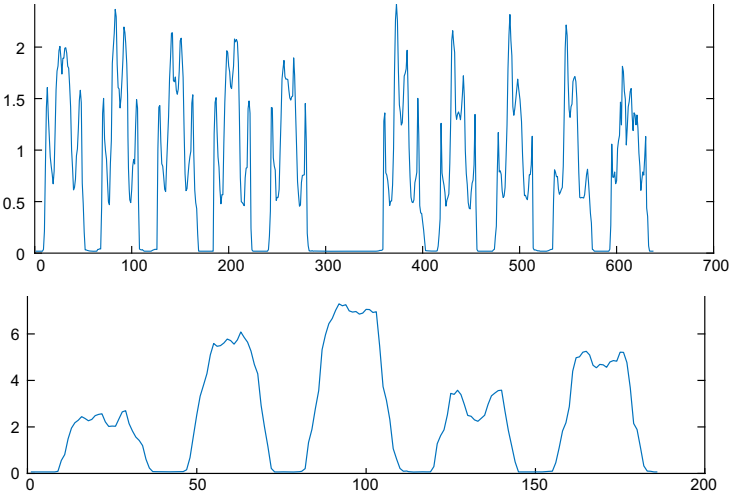
column of the image. Through the distance between each question and each option, the objective question area of the answer card is divided into grids, each grid corresponds to a question number, or an option. Store and record the question number or option information represented by grid to determine the location of correct answer.



(a) Options image



(b) Option graph binarization invers



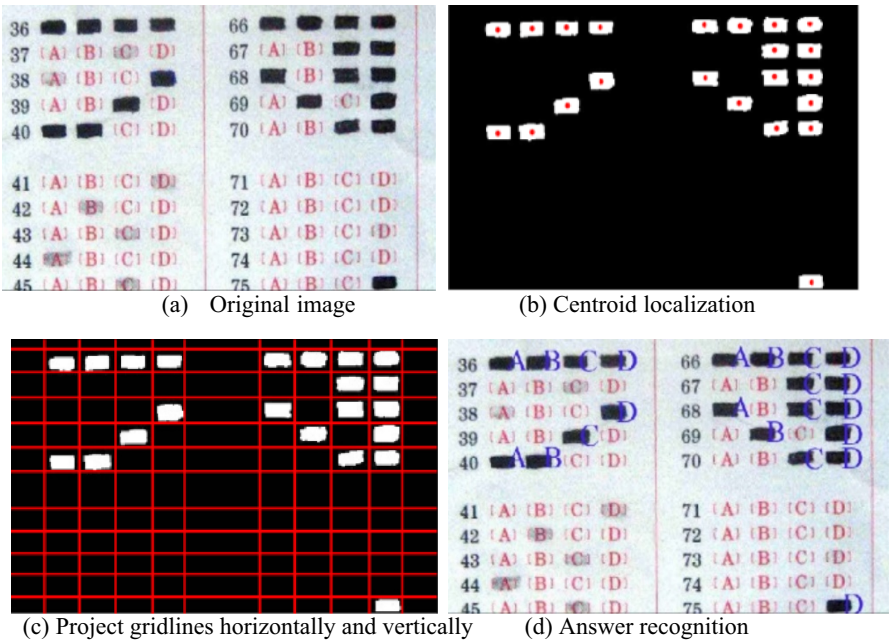
(c) Option map vertical horizontal projection

Fig. 6. Interval judgment of option area

As shown in Fig. 6, a is a partial option area image, and b is a binary inverse image of a. According to the projection of b in the vertical and horizontal directions, the horizontal

and vertical coordinates in the c represent the pixel width of the option image in the column and row directions, the change of the sum of the gray values of the questions, A, B, C and D, and the valley value represents the binary value in the image. The blank areas between the original pictures of 5 and 6 are the largest, so the valley value is the longest in the vertical projection. In the horizontal projection, the peak value of about 100 pixels in the middle is the largest, because option B is the most selected and the accumulated gray value is the largest. In this paper, the middle point whose valley value is zero in each segment is taken as the grid line to divide horizontally and vertically, so as to realize the grid line division between the questions and between the options.

Locate the connected area of the objective question after filtering out the impurity points and invalid areas, extract the central position of the objective question filling area, and mark it. Compare the location of the answer sheet options according to the central position, and determine which one of A, B, C and D options corresponds to the location, so as to realize the location and marking of the options (Fig. 7).



(a) Original image (b) Centroid localization

(c) Project gridlines horizontally and vertically (d) Answer recognition

Fig. 7. Results of filling location and option identification

The above figure shows the process of recognition of smear information in the objective question area of the answer card. After the image is preprocessed and corrected for binarization, the center of mass of the smear area can be accurately extracted, b marks the center of mass of the filled rectangle in a. C is the grid line divided by the horizontal projection and vertical projection of the answer area image. If the center of b is located in the cell corresponding to an option in the grid, then mark the corresponding option

value next to the filled area in d. According to the test result of the answer card and the standard answer, we can judge whether the filling result is correct.

5 Conclusion

In this paper, image preprocessing, segmentation, feature extraction and automatic recognition are applied in the marking of answer card. In this method, the image is preprocessed based on curvelet to improve the SNR and image quality. The Hough transform is used to detect the line effectively in polar coordinate space, and the candidates' information area and answer area are segmented effectively while the tilt angle is corrected. Using the horizontal and vertical projection segmentation option, realizes marking the filling position in the answer sheet and results recognition, improves the accuracy of marking papers.

Acknowledgements. This work is partly supported by the Ministry of Education Industry University Cooperation Project (201802312007), Science and Technology Project of Jiangsu Provincial Department of Housing and Construction (2019ZD041, 2019ZD039).

References

1. Jitngernmadan, P., Miesenberger, K.: A comparative study on Java technologies for focus and cursor handling in accessible dynamic interactions. In: 13th European Conference on the Advancement of Assistive Technology. Vol. 217, pp. 267–273 (2015)
2. Rogor, E.: Changes to the cursor stability isolation level: part 1. IBM Data Manag. Mag. **1**, 44–45 (2010)
3. Frustaci, F., Perri, S., Cocorullo, G., Corsonello, P.: An embedded machine vision system for an in-line quality check of assembly processes. Procedia Manuf. **42**, 211–218 (2020)
4. Zhang, Y., Soon, H.G., Ye, D., Fuh, J.Y.H., Zhu, K.: Powder-bed fusion process monitoring by machine vision with hybrid convolutional neural networks. IEEE Trans. Ind. Inf. **16**(9), 5769–5779 (2020)
5. Herakovic, N., Simic, M., Trdic, F., Skvarc, J.: A machine-vision system for automated quality control of welded rings. Mach. Vis. Appl. **22**(6), 967–981 (2011)
6. Jiang, D., Li, W., Lv, H.: An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. Neurocomputing **220**, 160–169 (2017)
7. Shu, Y.F., Xiong, C.W., Fan, S.L.: Interactive design of intelligent machine vision based on human–computer interaction mode. Microprocess. Microsyst. **75**, 103059 (2020)
8. Jiang, D., Wang, Y., Lv, Z., Wang, W., Wang, H.: An energy-efficient networking approach in cloud services for IIoT networks. IEEE J. Sel. Areas Commun. **38**(5), 928–941 (2020)
9. Wang, F., Jiang, D., Qi, S.: An adaptive routing algorithm for integrated information networks. China Commun. **7**(1), 196–207 (2019)
10. Huo, L., et al.: An intelligent optimization-based traffic information acquirement approach to software-defined networking. Comput. Intell. **36**(1), 151-171 (2019)
11. Jiang, D., Huo, L., Lv, Z., et al.: A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking. IEEE Trans. Intell. Transp. Syst. **19**(10), 3305–3319 (2018)
12. Huo, L., et al.: An AI-based adaptive cognitive modeling and measurement method of network traffic for EIS. Mob. Netw. Appl. (2019)

13. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **7**(1), 80–90 (2020)
14. Hou, S.J., Wu, S.L.: Image denoising by the curvelet transform for doppler frequency extraction. Chin. J. Electron. **17**(1), 178–182 (2008)
15. Raju, C., Reddy, T.S., Sivasubramanyam, M.: Denoising of remotely sensed images via Curvelet transform and its relative assessment. Procedia Comput. Sci. **89**, 771–777 (2016)
16. Altan, A., Karasu, S.: Recognition of covid-19 disease from x-ray images by hybrid model consisting of 2d curvelet transform, chaotic salp swarm algorithm and deep learning technique. Chaos Solitons Fractals **140**, 110071 (2020)
17. Jiang, D., Wang, W., Shi, L., Song, H.: A compressive sensing-based approach to end-to-end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. **7**(1), 507–519 (2020)
18. He, T., Shang, H.: Direct-wave denoising of low-frequency ground-penetrating radar in open pits based on empirical curvelet transform. Near Surf. Geophys. **18**(3), 295-305 (2020)
19. Li, J., Wang, Y., Xiao, H., Xu, C.: Gene selection of rat hepatocyte proliferation using adaptive sparse group lasso with weighted gene co-expression network analysis. Comput. Biol. Chem. **80**, 364–373 (2019)
20. Wang, Y. Jiang, D., Huo, L., Zhao, Y.: A new traffic prediction algorithm to software defined networking. Mob. Netw. Appl. 1–10 (2019)
21. Kedar, A., Ligthart, L.P.: Wide scanning characteristics of sparse phased array antennas using an analytical expression for directivity. IEEE Trans. Antennas Propag. **67**(2), 905–914 (2019)
22. Jiang, D., Wang, Y., Lv, Z., Qi, S., Singh, S.: Big data analysis based network behavior insight of cellular networks for industry 4.0 applications. IEEE Trans. Ind. Inform. **16**(2), 1310–1320 (2020)
23. Mahdinejad, N., Mota, H.O., Silva, E.J., Adriano, R.: Improvement of system quality in a generalized finite-element method using the discrete Curvelet transform. IEEE Trans. Magn. **53**(6), 1–4 (2017)
24. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. PLoS ONE **13**(5), 1–23 (2018)
25. Qi, S., Jiang, D., Huo, L.: A prediction approach to end-to-end traffic in space information networks. Mob. Netw. Appl. 1–10 (2019)
26. Jiang, D., Zhang, P., Lv, Z., et al.: Energy-efficient multi-constraint ligent optimization-brouting algorithm with load balancing for smart city applications. IEEE Internet Things J. **3**(6), 1437–1447 (2016)
27. Luo, J., Chai, S., Zhang, B., Xia, Y., Gao, J., Zeng, G.: A novel intrusion detection method based on threshold modification using receiver operating characteristic curve. Practice and Experience, Concurrency and Computation (2020)
28. Vuppala, A., Krämer, A., Braun, A., Lohmar, J., Hirt, G.: A new inverse explicit flow curve determination method for compression tests. Procedia Manufacturing **47**, 824–830 (2020). https://doi.org/10.1016/j.promfg.2020.04.257
29. Caraka, R.E., Chen, R.C., Toharudin, T., Pardamean, B., Yasin, H.: Ramadhan short-term electric load: a hybrid model of cycle spinning wavelet and group method data handling (CSWGMDH). IAENG Int. J. Comput. Sci. **46**, 670–676 (2020)
30. Luo, P., Yao, W., Susmel, L.: An improved critical-plane method and cycle counting method to assess damage under variable amplitude multiaxial fatigue loading. Fatigue Fract. Eng. Mater. Struct. **43**(9), 2024–2039 (2020)
31. Kozhemyakin, G.N., Kovalev, S.Y., Soklakova, O.N.: Fabrication of bismuth films by a melt spinning method and the influence of annealing on their microstructure. Inorg. Mater. Appl. Res. **11**(3), 727–730 (2020)
32. Fernández, A., Alonso, J.R., Ayubi, G.A., Osorio, M., Ferrari, J.A.: Optical implementation of the generalized Hough transform with totally incoherent light. Opt. Lett. **40**(16), 3901 (2015). https://doi.org/10.1364/OL.40.003901

33. Chandrasekar, L., Durga, G.: Implementation of hough transform for image processing applications. In: International Conference on Communication and Signal Processing, pp. 843–847 (2014)
34. Wei, W., Dong, X.Q., Shen, Y.Y.: Research on a two value generalized Hough transform method of identification. In: Proceedings of 2011 International Conference on Computer Science and Network Technology, ICCSNT 2011, Vol. 1, pp. 278–281 (2011)
35. Zhang, W., Zhu, P., Cheng, L., Zhu, H.: Improved centripetal force type-magnetic bearing with superior stiffness and anti-interference characteristics for flywheel battery system. Int. J. Precis. Eng. Manuf. Green Technol. **7**(3), 713–726 (2020)
36. Ma, J., Yang, Y., Li, H., Li, J.: FH-BOC: generalized low-ambiguity anti-interference spread spectrum modulation based on frequency-hopping binary offset carrier. GPS Solutions **24**(3), 1–16 (2020).
37. Mandhala, V.E.N., Bhattacharyya, D., Kim, T.H.: Face detection using image morphology - a review. Int. J. Secur. Appl. **10**(4), 89–94 (2016)
38. Nazre, B., Rama, C.: Fast detection of facial wrinkles based on Gabor features using image morphology and geometric constraints. Pattern Recogn. **48**(3), 642–658 (2015)
39. Sangpongsanont, Y., Chenvidhya, D., Chuangchote, S., Kirtikara, K.: Corrosion growth of solar cells in modules after 15 years of operation. Sol. Energy **205**, 409–431 (2020)
40. Ryu, H.S., Lee, H.S., Jalalzai, P., Kwon, S.J., Aslam, F.: Sodium phosphate post-treatment on al coating: morphological and corrosion study. J. Therm. Spray Technol. **7**(2), 1–21 (2019)
41. Jafari, A., et al.: Statistical, morphological, and corrosion behavior of pecvd derived cobalt oxide thin films. J. Mater. Sci. Mater. Electr. **30**(24), 21185–21198 (2019). https://doi.org/10.1007/s10854-019-02492-6