# Junction Based Table Detection in Mobile Captured Golf Scorecard Images

Junying Yuan(✉), Haishan Chen, Huiru Cao, and Zhonghua Guo

Department of Electronic Communication and Software Engineering,
Nanfang College of Sun Yat-Sen University, Guangzhou 510975, China
cihisa@outlook.com
http://www.scholat.com/cihisa

**Abstract.** Table detection in mobile captured images faces many challenges owning to the well-known low image quality. Recently, a few researches pioneer in detecting the tables in rich-text images, but few works for scorecard images which usually lack of texts but are rich in graphics, such as golf scorecard images. In this paper, a junction-relation based table detection method for mobile captured scorecard images is proposed. Firstly, the most distinguished junctions are determined via a simplified pattern matching method, then the fault detections are removed through filtering operations, finally the missed junctions are recovered utilizing the pair-wise relationships among neighboring junctions. The experimental results show that 98.47 % of the junctions from 90 test images are correctly detected, and thus proves the superiority of the proposed method.

**Keywords:** Mobile captured images · Junction detection · Table detection · Pair-wise relationship · Junction filtering · Junction recovery

## 1 Introduction

Table is one of the most commonly used document types due to the simple but well-defined structure for highly concise data representation [1]. Therefore, automatic table recognition has been extensively studied in the past years and are broadly applied in the business and the daily life. Traditional table image analysis methods [2–5] focus on flatbed scanned images, which are acquired with high quality and thus are easy to recognize. They usually assume that straight and parallel lines exist in the scanned images, e.g. the table ruling lines and the text lines. Based on this assumption, tables are usually recognized satisfactorily.

In recent years, smart phones with high performance digital cameras become widely used in every corner of the society. Compared to flatbed scanners, smart phones are portable, fast in response and rich in functionality [6,7]. As a result, requirements arise in demanding for realtime document image analysis using smart mobiles. However, the quality of mobile captured (MC) document images


J. Wan et al. (Eds.): Industrial IoT 2016, LNICST 173, pp. 179–188, 2016.
DOI: 10.1007/978-3-319-44350-8_18

are considerably low due to known issues, such as low effective resolution, perspective distortion, shading and noise. Therefore, traditional table analysis methods are no longer effective when applying them directly to mobile captured table images [1,8]. Seo et al. [8] proposed a junction-based table detection method for camera captured images. This method firstly locates the table region using a modified X-Y cutting technique, then detects the intersections of horizontal and vertical ruling lines, and finally label the junctions using a cost function which is constructed based on the directional information. Seo et al.'s method works well for document images with simple backgrounds, where no complex table content or text touching exists.

The contents of golf scorecards vary in table format, background and font formats. In addition, the quality of the mobile captured pictures are considerably low, such that the ruling lines are often blurred and texts sometimes touch the ruling lines. Yuan et al. [9] made a first attempt in table region detection and ruling line detection in mobile captured golf scorecard images. However, the method produces a number of miss detections and fault detections. This manuscript focuses on the problem of table detection in mobile captured golf scorecard images. Firstly, the score table region is detected by extracting the maximum connected component (CC) in the binary image [9]. Then, the candidate junctions (those pixels which have partial directional information) are captured using a simplified pattern matching technique, and are further filtered to peak out the true detections by exploiting the pair-wise junction relationships. Finally, the pair-wise relationships among neighboring junctions are further exploited to recover the missed detections. Experimental results show that the proposed method can detect ***98 %*** of the real junctions while generating few fault detections.

The rest of this manuscript is organized as follows. Section 2 briefs the related works and the observations from mobile captured golf scorecard images. The proposed work is detailed in Sect. 3. The experimental results are analyzed in Sect. 4. Section 5 concludes this manuscript.

## 2 Related Works and Observations

This section firstly briefs Seo et al.'s work in [8] and the previous work in [9], and then lists the observations from mobile captured golf scorecard images.

### 2.1 Related Works

Seo et al. consider junctions as the intersections of horizontal and vertical ruling lines, and label them into 12 patterns based on the direction information. To improve detection performance, they constructed a cost function based on the inter-junction connectivity information, and then minimized it using belief propagation algorithm. Their method is verified on a newly built image set which is composed camera captured magazines and books. The tables therein is simple in table structure, content, text font and background. In addition, there is no text touching issues in the binary table images.

The research in [9] handles with junction detection in mobile captured golf scorecard images, where the problems of diverse thickness table ruling lines, complex fonts and backgrounds, exist commonly. The table region is easily detected by finding the maximum connected components in the binary image, and the binary ruling lines are enhanced before junction detection which is performed using a simplified pattern matching technique. However, problem still remains in the commonly existed fault detections and miss detections due to text touching and the poor binarization. For further details of the above two methods, please refer to [8] and [9] respectively.

## 2.2 Observations

In this section, distinguishable observations from the mobile captured golf scorecard images are listed. Some of them may block junctions from being correctly detection, and the others could on aid the process of junction detection. As illustrated in Fig. 1, the problems blocking junction detection are listed as below:

- Table contents vary vastly in table format, fonts and background.
- Perspective distortion, shading, noise and motion blurs exist commonly in the mobile captured score table images.
- The effective resolution is low.
- Text touching exists in a part of the mobile captured images.


**Fig. 1.** A typical mobile captured golf scorecard image

In addition, some beneficial observations are also found and detailed as below:

- The score table ruling lines make up the maximum connected component in the binary image.
- The ruling lines, especially the vertical ruling lines, can be approximated as locally straight.
- Scorecards are usually taken with a very small rotation angle and thus can be neglected.

Most of the observed items have been handled or exploited in [9]. However, the problem of text touching, which is the cause of most fault detections, and the issue of miss detections, are not resolved. In this paper, the beneficial observations are further exploited to simplify the process of junction detection, the previously observed issues are carefully handled to improve the performance, and the text touching issue is resolved to minimize fault detections.

## 3 The Proposed Work

There remains two challenges, including the poorly binarized image and text touching, before moving table recognition in mobile captured scorecard images to the real world. The problem of poor binarization may reduce miss detections and fault detections, and text touching always results in numerous fault detections.

In this manuscript, candidate junctions are firstly detected via the simplified pattern matching technique [9]. Then they are filtered according the pair-wise relationships where junctions are adjacent and are horizontally or vertically connected via ruling lines. In order to avoid fault detections, a junction filtering step removes as many candidate junctions as possible, including some of the true detections. Therefore the vast majority of remaining junctions are true detections. Finally, the missed junctions are recovered iteratively by exploiting the existence probability of the connecting ruling lines. Accordingly, the junctions can be easily labeled using the pairwise relationships before extracting the table cells.

### 3.1 Junction Detection

The resolution of mobile captured images is usually high, but is often low in resolution effectiveness due to the existence of noise. Therefore, the high resolution mobile captured table images are first sampled down and then de-noised using a low-pass Gaussian filter. The table region is extracted from the maximum connected component in the binary image, and the binary ruling lines are thinned before junction detection.

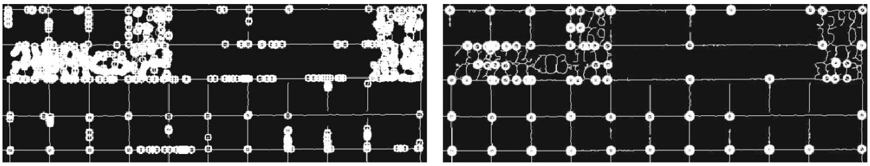
The candidate junctions are captured using the simplified pattern matching technique [9] which produces many fault detections and miss detections, as illustrated in Fig. 2(a). Note that a great number of fault detections appear around the true junctions. Therefore, the candidate junctions can be clustered to select the most probable one by exploiting the directional information in the four directions, including N(orth), (W)est, (S)outh and (E)ast. The directional information can be calculated as the existing probability of a ruling line in a given direction $D \in \{N, W, S, E\}$. Let $I$ denotes the binary table region, $L_R$ denotes the region size, $i_0$ be the binary pixel where the candidate junction locates, and $i_l^D$ be the binary pixel in the direction $D$ of $i_0$ with an offset $l \in \{1, ..., L_R\}$. The directional information $p_D$ in direction $D$ is calculated using the equation

$$p_D^i = \frac{1}{L_1} \sum_l i_l^D. \tag{1}$$

According to the directional information, the probability of pixel $i$ being a true detection can be determined using the formula

$$p^i = \frac{1}{4} \sum_D p_D^i. \tag{2}$$

As a result, the most probable true detections can be peaked out by finding the candidate junctions with the maximum $p^i$ in a local region. The clustering process is executed recursively until that the inter-junction distance between any two junctions is larger than $L_1$, as illustrated in Fig. 2(b). The clustered junction list is denoted as $J_1$ for future usage.



(a) (b)

**Fig. 2.** Junction detection and clustering

### 3.2 Junction Filtering

There may still exist fault detections in junction list $J_1$ due to the existence of text touching and blurs, as illustrated in Fig. 2(b). Fortunately, they can be filtered out by exploiting the pair-wise junction relationships.

According to the observations, table junctions are the intersections of the horizontal and vertical ruling lines. And thus they can be arranged in a two dimensional (2-D) grid. Non-null grid element indicate true detections, and a null element reveals possibly a miss detection or the existence of no junction. Therefore, the problem of junction filtering is turned to mapping the junctions in $J_1$ into a 2-D grid $G$. Since the image rotation is negligible and the ruling lines are less distorted in the vertical direction, this mapping begins by arranging junctions column-wisely from the left to the right. And then follows the determination of horizontal pair-wise relationships.

**Junction Filtering in the Vertical Direction.** Before arranging junctions column-wisely, the lateral ordinates of vertical ruling lines should be determined. Since the ruling lines are no longer straight, their lateral ordinates cannot be exactly calculated. Therefore, in this manuscript, it is proposed to estimate them by exploiting the junction distribution density from junction projection in the vertical direction. Firstly, the junction density $f_d$ is obtained by projecting the junctions in the vertical direction, then the number of junctions on a vertical ruling line is estimated, by summing up the junction density within a small range of horizontal ordinates, which is restricted by the width of the vertical ruling line. Let $x_{TL}$, $x_{BL}$, $x_{TR}$ and $x_{BR}$ denote the lateral ordinates at the top left corner, the bottom left corner, the top right corner and the bottom right corner, respectively. The maximum deviation $\dot{R}_x$ of the vertical ruling lines in the horizontal direction can be estimated using the equation

$$\dot{R}_h = \frac{1}{2} max\{abs(x_{TL} - x_{BL}), abs(x_{TL} - x_{BL})\}. \tag{3}$$

Since $\dot{R}_h$ indicates the maximum distance that a junction deviates from the corresponding vertical ruling line, the number of junctions on the vertical ruling line can be counted using the formula

$$n_J = \sum_{x-\dot{R}_h}^{x+\dot{R}_h} f_d(x). \tag{4}$$

To alleviates the impacts of text touching and noise, $n_J$ is further processed using a high-pass filter specified by the equation

$$n'_J = \begin{cases} n_J \; n_J > \tau_1, \\ 0 \quad otherwise, \end{cases} \tag{5}$$

where $\tau_1$ is threshold to remove the impact of text touching and noise in the regions with no vertical ruling lines.

By far, the lateral ordinates of the vertical ruling lines can be estimated by the peaks of $n'_J$, and are denoted by $x_v(c), c \in \{1, 2, ..., N^V\}$, where $N^V$ is the number of estimated vertical ruling lines. Given the lateral ordinates, the junction list $J_1$ can be filtered to remove possible fault detections, whose lateral ordinates fall out of the scopes $[x_v(c) - \dot{R}_h, x_v(c) + \dot{R}_h]$, and the remaining junctions in the scope of $[x_v(c) - \dot{R}_h, x_v(c) + \dot{R}_h]$ is called a junction column $J(c)$, where $1 \leq c \leq N^V$.

The updated junction list still contains fault detections and can be further filtered by exploiting the pair-wise relationship in the vertical direction. Since the vertical ruling lines are less distorted, the neighboring junctions on the same vertical ruling line presents a very close slope factor. Let $\phi(j, k)$ represent the slope factor between the $j$-th and $k$-th junctions in the same column $J(c)$, the standard deviation $\sigma_\phi$ of the pair-wise junction slope factors can be calculated using the equation

$$\sigma_\phi = \frac{\sum_j \sum_{k \neq j} (\phi(j,k) - \mu_\phi)^2}{N_c(N_c - 1)}, \tag{6}$$

where $N_c$ is the number of junctions in that vertical column $J(c)$, and $\mu_\phi$ is the average slope factor of the junction column. Similarly, the standard deviation $\sigma_\phi(j)$ of the slope factors between the $j$-th junction and the remaining junctions in $J(c)$, can be calculated using the formula

$$\sigma_\phi(j) = \frac{1}{N_c - 1} \sum_{k \neq j} (\phi(j,k) - \mu_\phi(j))^2, \tag{7}$$

where

$$\mu_\phi(j) = \frac{1}{N_c - 1} \sum_{k \neq j} \phi(j,k). \tag{8}$$

If $\sigma_\phi(j) < \sigma_\phi$, the $j$-th junction is possibly a fault detection and is removed from junction column $J(c)$. After vertical junction filtering, the updated junction list is denoted as $J_2$.

**Junction Filtering in the Horizontal Direction.** Since ruling lines can be treated as locally straight, the pair-wise junction relationships in the horizontal direction, can be determined by the existence probability of a horizontal ruling line. Let $\zeta_c^r$ denotes a junction in junction column $J(c)$ with index $r$, $x_r$ and $y_r$ be the lateral and vertical ordinates of junction $\zeta_c^r$. The existence probability of a horizontal ruling line between $\zeta_c^r$ and $\zeta_{c+1}^{r'}$, can be calculated using the formula

$$\begin{aligned} y &= y_r + \left\lfloor \tfrac{(y_{r'} - y_r)}{(x_{r'} - x_r)} * (x - x_r) \right\rfloor, \\ p_h(\zeta_c^r, \zeta_{c+1}^{r'}) &= \tfrac{1}{x_{r'} - x_r} \sum_{x = x_r}^{x_{r'}} i(x, y), \end{aligned} \tag{9}$$

where $i(x, y)$ is the binary pixel value at ordinates $(x,\, y)$. If $p_h(\zeta_c^r, \zeta_{c+1}^{r'})$ exceeds a certain threshold $\tau_2$, a horizontal ruling line possibly exists between these two junctions. And if there exists a $\dot{p}_h(\zeta_c^r, \zeta_{c+1}^{r'})$ meeting the below rule, the junction $\zeta_{c+1}^{r'}$ is deemed as horizontally linked to $\zeta_c^r$:

$$\dot{p}_h(\zeta_c^r, \zeta_{c+1}^{r'}) = max\{p_h(\zeta_c^r, \zeta_{c+1}^j, j \in [1, N_{c+1}]\} > \tau_2. \tag{10}$$

When all horizontal pair-wise junction relationships are determined, the junction grid $G$ is constructed, as illustrated in Fig. 3, where it can be noted that all fault detections (in Fig. 2(b)) have been filtered.

### 3.3 Junction Recovery

The junction filtering process removes a vast majority of the fault detections, and also drops some true detections. Therefore, it is required to recover the missed detections in the junction grid $G$. Fortunately, this is achievable by exploiting the pair-wise relationships between neighboring junctions. It can be observed that the true detections have at least one neighboring junction in the horizontal

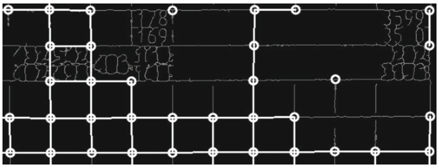


**Fig. 3.** Constructed junction grid

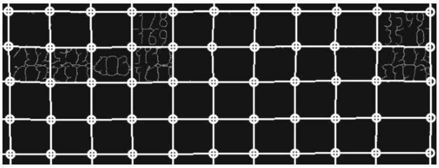


**Fig. 4.** Recovered junction grid

direction and another one in the vertical direction. Let $\zeta(r,c)$ be an element of junction grid $G$ in the $r$-th row and $c$-th column. $\zeta(r,c)$ can be deemed as a true detection if any one of the following conditions is satisfied:

$$
\begin{aligned}
(p_N(r+1,c) \geq \tau 3) \times (p_W(r,c-1) \geq \tau 3) = 1.\\
(p_N(r+1,c) \geq \tau 3) \times (p_E(r,c+1) \geq \tau 3) = 1.\\
(p_S(r-1,c) \geq \tau 3) \times (p_W(r,c-1) \geq \tau 3) = 1.\\
(p_S(r-1,c) \geq \tau 3) \times (p_E(r,c+1) \geq \tau 3) = 1.
\end{aligned}
$$

The ordinates of a recovered junction can be estimated via the neighboring junctions. The recovery process is performed repeatedly until no new junction is recovered, and the recovered and labeled junctions of Fig. 3 is shown in Fig. 4.

## 4 Experiments and Analysis

A few experiments are designed to verify the performance of the proposed work. The image database in [9] is employed in this manuscript. This database consists of 90 mobile captured golf scorecard images, which are captured from three different brands of mobile phones. There are 90 score tables, 1203 rows, 1188 columns and 15288 junctions in the captured images. Therein, the original and the full picture of Figs. 2, 3 and 4 is illustrated in Fig. 1, from which the issues in a mobile captured scorecard image can be easily observed.

The specific values of the three thresholds, which are employed in this manuscript, are fixed as $\tau_1 = 5$, $\tau_2 = 0.8$ and $\tau_3 = 0.25$ using empirical values from experiments. To evaluate the performance of junction detection, three measures are employed, including the precision $P$, the recall rate $R$, and the F-measure $F$ which is the combination of $P$ and $R$. The three measures are defined using the equation

$$
P = \frac{tp}{N_D}, R = \frac{tp}{N_R}, F = \frac{2 \times P \times R}{P + R}, \tag{11}
$$

where $tp$ is the number of true detections, $N_D$ is the number of detected junctions including true detection and fault detection, and $N_R$ is the real number of junctions. The F-measure, which is the mixture of $P$ and $R$, is susceptible to a minimum value, therefore it can better measure the overall performance.

Table 1 gives the performance when applying the three measures to the detected tables, rows and columns. Note that all the table regions are correctly detected. When excluding the short ruling lines in the split table cells, 99.00 % percent of the rows and 99.83 % percent of the columns can be captured successfully. Such a result reveals that, compared to that in [9], the proposed method may probably induce less miss detections.

**Table 1.** Detected table regions, rows and columns

| | $P$ (%) | $R$ (%) | $F$ (%) |
|---|---|---|---|
| Table | 100.00 | 100.00 | 100.00 |
| Row | 99.00 | 100.00 | 99.50 |
| Column | 99.83 | 99.92 | 99.87 |

To compare with [9] (denote it as Method I), the three measures are applied to evaluate the performance when detecting all junctions $J_{all}$, the junctions on the external boundary $J_{eb}$ and junction within the table external boundary $J_{in}$. Table 2 details the produced performance metric. Note that the proposed method overpasses Method I in each and every of the measure results. The precision of $J_{all}$ of the proposed method is 98.47 %, which is 2.19 % higher than that of Method I. This means that much more junctions are correctly detected by the proposed work. The recall of $J_{all}$ is 98.91 % and is 7.43 % higher than that of Method I, meaning that only a few fault detections are produced. In addition, the F-measure of $J_{all}$ is 98.69 %, indicating the robustness of the proposed method. In a summary, the proposed method is effective in junction detection by filtering out the fault detections and exploiting the true detections.

**Table 2.** Detected junctions

| | | $P$ (%) | $R$ (%) | $F$ (%) |
|---|---|---|---|---|
| $J_{all}$ | Method I | 96.38 | 91.48 | 93.39 |
| | Proposed | 98.47 | 98.91 | 98.69 |
| $J_{eb}$ | Method I | 92.74 | 99.69 | 95.67 |
| | Proposed | 98.42 | 98.83 | 98.62 |
| $J_{in}$ | Method I | 95.26 | 92.19 | 92.91 |
| | Proposed | 98.49 | 98.95 | 98.72 |

## 5 Conclusion

This manuscript presents a junction detection scheme for mobile captured score-card images. The proposed method handles with two problems including fault detections and miss detections. By exploiting the pair-wise relationships among neighboring junctions, the vast majority of fault detections are firstly filtered out during junction grid construction, and the missed detections are then repeatedly recovered. The experimental results show that 98.47 % of the junctions can be correctly detected, and thus the effectiveness of the proposed work is proved.

**Acknowledgements.** This work is supported by the Foundation for Distinguished Young Talents in Higher Education, the Teaching Quality and Teaching Reform Project the Science and Technology Project of Guangdong, China, with Grant Nos. 2013LYM0123, ZL2013025, and 2013B090500067 respectively. Any options, findings, and conclusions or recommendations expressed in this paper are those of the authors. Also the authors would like to thank Guangzhou Gaoyou-Box Ltd. for the support on the golf scorecards.

## References

1. Kise, K.: Handbook of Document Image Processing and Recognition. Springer, London (2014)
2. Chen, J., Lopresti, D.: Ruling-based table analysis for noisy handwritten documents. In: Proceedings of the 4th International Workshop on Multilingual OCR. ACM (2013)
3. Gatos, B., Pratikakis, I., Perantonis, S.J.: An adaptive binarization technique for low quality historical documents. In: Marinai, S., Dengel, A.R. (eds.) DAS 2004. LNCS, vol. 3163, pp. 102–113. Springer, Heidelberg (2014)
4. Cesarini, F., Marinai, S., Sarti, L., Soda, G.: Trainable table location in document images. In: Proceedings of the 16th International Conference on Pattern Recognition, vol. 3, pp. 236–240. IEEE (2002)
5. Ha, J., Haralick, R.M., Phillips, I.T.: Recursive X-Y cut using bounding boxes of connected components. In: Proceedings of the Third International Conference on Document Analysis and Recognition, vol. 2, pp. 952–955. IEEE (1995)
6. Liang, J., Doermann, D., Li, H.: Camera-based analysis of text and documents: a survey. Int. J. Doc. Anal. Recogn. (IJDAR) **7**(2–3), 84–104 (2005)
7. Mirmehdi, M.: Special issue on camera-based text and document recognition. Int. J. Doc. Anal. Recogn. (IJDAR) **7**(2–3), 83 (2005)
8. Seo, W., Koo, H., Cho, N.: Junction-based table detection in camera-captured document images. Int. J. Doc. Anal. Recogn. (IJDAR) **18**(1), 47–57 (2015)
9. Yuan, J.Y., Chen, H.S., Cao, H.R.: An efficient junction detection approach for mobile-captured golf scorecard images. Procedia Comput. Sci. **55**, 792–801 (2015)