Fast Frame Decomposition and Sorting by Contour Tracing for Mobile Phone Comic Images

Yusuke IN, Takashi OIE, Masakazu HIGUCHI, Shuji KAWASAKI, Atushi KOIKE and Hitomi MURAKAMI
Graduate School of Information Technology
Seikei University
3-3-1 Kitijyoujikitamati, musashino, Tokyo 180-8633 Japan
Yuhsuke.in@gmail.com, hi-murakami@st.seikei.ac.jp

Abstract: - As one of the mobile phone contents, electronic books and magazine, especially comics, are expected to have great potential. In order to view comics on mobile phones, the resolution and the size of screens must be adequate. A comic then needs to be decomposed into frames and put them in order so that they can be read properly. However, it is known that accurate decomposition is not an easy task. In this paper, we propose a fast and accurate clipping method using contour tracing and evaluate it comparing with an existing commercial system.

Key-Words: - Comic, mobile phone, image processing, layout analysis, Hough transform

1. Introduction
With the spread of the Internet in recent years, the E6book market is growing very fast in Japan as shown in Figure 1.

Figure 1: The E-book market in Japan.

The Japanese market is three times larger than that of the U.S. and more than 70% is comics for mobile phones. It would not be an exaggeration to say that the E-book market in Japan is supported by the comics for mobile phones.

It is difficult to read the comic when the entire page is displayed on the small LCD screen of the mobile phone. For this reason, a method for sequentially reading each of the whole comic pages has been proposed in [1]. In this method, for each frame, the display order and the region information are required. Thus, methods for automatic detection of this information are currently and widely under research [2-5]. One example of such systems in commercial use is the MangaGet service provided by SpicySoft [6] in Japan.

However, there are a number of problems with these conventional systems when applied to mobile phone displays.

To address the above-mentioned problems, this paper proposes a new and novel method for frame decomposition and sorting, and evaluates it against commercial methods in use.

2. Existing Frame Decomposition and Sorting Methods and their Problem

In these methods, the processing involves recursively bisecting the dividing lines of the comic image, as shown in Figure 2(a).

Figure 2: a) Existing frame decomposition and sorting method.

However, for comic images, there are cases in which there is no information outside the margin to the edge of the frame as shown in Figure 3. The above methods do not take such situations into consideration resulting in some useless margin being included in the divided region. For terminals with not enough display screen size the region of the frame around the margin becomes small and difficult to read. Thus the above approach will not always achieve the appropriate decomposition for mobile phone displays.

Presently the MangaGet system employs a decomposition and sorting method for comics that allows users to browse the comics on mobile phones.

Here, by clipping the part of the comic frame so that it can fit the mobile phone screen, performing necessary corrections, and then uploading it to the server, the user can be able to browse the comic on the mobile phone. This
concept is depicted in Figure 2(b). As for the MangaGet system, the above problem is addressed, although not publicly available. However, in some cases, the order in which the regions are displayed is not correct. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manually. To reduce this burden of manually correcting the frames, it is necessary to make the correction manually. After automatic detection, it is necessary to make the correction manual...
For the detected lines, all lines with $\theta$ less than or equal to $K^\circ$ and the difference in $\rho$ within $M$ are treated as one line. Here, we use the settings $K^\circ = 1$ and $M = 10$. In this way, extremely narrow blank margins regions between the frames are ignored. The resulting straight line image pixels are denoted by $L(\rho, \theta)$. Additionally, by using pixels $L(\rho, \theta)$ and the neighboring pixels denoted by $M(\rho, \theta)$, the intensity gradient at pixel $(x,y)$ is calculated by Equation (3) as follows:

$$g_{\rho}(x,y) = g_x(x,y) \cos \theta + g_y(x,y) \sin \theta \quad (3)$$

$g_x(x,y)$: Horizontal intensity gradient
$g_y(x,y)$: Vertical intensity gradient

There are cases in which part of the picture and the Serif extends outside the frame into the blank margin and cases where the shape of the frame is not polygonal. In these cases, frame division lines cannot be detected well. To deal with these cases, we consider $L(\rho, \theta)$’s neighboring pixels $M(\rho, \theta)$. Based on our much experience, when frame decomposition is performed, the center of the image is likely to have straight lines. For this reason, the Gauss function given by Equation (4) and shown in Figure 6 is used as weighting for images with straight lines.

$$G(\rho) = \exp\left( -\frac{\rho^2}{\sigma^2} \right) \quad (4)$$

$\sigma$: The height of the image divided by four.

By this weighting function, lines that are close to the center of the image are assigned big weights, while those away for the center of the image are assigned small weights.

Using Equations (1), (2), (3) and (4), the values $A(\rho, \theta)$ determined by the line division process for straight lines $L(\rho, \theta)$ are given by Equation (5).

$$A(\rho, \theta) = G(\rho) \sum_{(x,y) \in L(\rho, \theta)} \left[ g_{\rho}(x,y) C(x,y) \right] \quad (5)$$

In the above equation, if $\theta$ is between 45$^\circ$ and 135$^\circ$ candidate horizontal lines assumed, and if the condition is not satisfied, vertical candidate lines are assumed. Also, since a large weight $C(x,y)$ is added to straight line along the frame, straight lines near the center of the image are detected as candidate frame division lines. In this way, when performing frame decomposition, large blank margins are detected as frames but they are deleted during sorting, resulting in high precision decomposition and sorting of frames.

The frame decomposition and sorting process is done as follows. First, using the comic as input image, the detected division lines with the largest values of $A(\rho, \theta)$ are separated. In this case, whether the division lines that will become the target surround the whole rectangular area of the entire image or not, is used as a criteria deciding division lines and non-division lines. In the case of horizontal dividing line, top and bottom edges of the rectangular region are decided, in the case of vertical dividing lines, the left and right edges are decided, and by the focusing on these division lines, the image is partitioned into two. If the above is not the case, then the process is terminated without division. For the divided region, a square surrounding region is again formed, and using the above process recursively to determine the dividing lines, we split the comic image into two. When performing the division, regions that are split into the top and button by a horizontal dividing line are processed from the top to the bottom region. Similarly regions that are split by a vertical line are processed from the left region to the right region. In this way, the frame sorting is performed by a tree structure. At the same time, regions with no information are deleted. The processing flow chart is shown in Figure 7.
The tree structure example is shown in Figure 8. The numbers are shown by the order of the frame.

Figure 8: The tree structure that is used for frame sorting (Hor. Div. Line: Horizontal Division Line, Ver. Div. Line: Vertical Division Line).

The results of comic image processing by the conventional method and the Comic-CFD are shown in Figure 9. In addition, Figure 10 shows a typical example that easily shows the difference between the conventional method and proposed method.

Figure 9: Comparison of the conventional and proposed methods.

Figure 10 shows a character with a piece protruding the drawing in which the extra information is excluded and the frame is naturally divided.

Figure 10: Comparison of the conventional and proposed methods [2]. (Source: Title SLAM DUNK, Publisher: Shueisha, Vol. 1, p.23)

4. Experiments

For the proposed Comic-CFD method, the performance of the method with respect to the required region and the processing order for a mobile phone display was evaluated for each using comic database images.

A. Experiment Setting

In total, 2037 pages of 32 different comics were used in the experiment. The images of cartoon genre boy, girl/woman and youth were selected from published copies [9] in representative genre for young comics. With the assistance of Seikei University’s Cartoon Research, Persian Cat members, a database of correct image regions and processing order was built.

Table 1 below shows the quantity and the resolution of the images at the time of scanning. The computing environment is shown in Table 2.

Under these conditions, the Comic-CFD method was run. Table 3 shows the average processing time for each genre of
the comics. In this case, multithread processing [10] was
carried out, resulting in high processing speed.

Table 1: Number of images and resolution.

<table>
<thead>
<tr>
<th>Cartoon Genre</th>
<th>Number(page)</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy</td>
<td>961</td>
<td></td>
</tr>
<tr>
<td>Girl/woman</td>
<td>512</td>
<td>144dpi</td>
</tr>
<tr>
<td>Youth</td>
<td>564</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Computing Environment.

<table>
<thead>
<tr>
<th>CPU</th>
<th>Intel Core2Quad Q6600 2.4GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Memory</td>
<td>4GB</td>
</tr>
<tr>
<td>OS</td>
<td>Windows Vista</td>
</tr>
</tbody>
</table>

Table 3: Average processing time for each genre.

<table>
<thead>
<tr>
<th>Cartoon Genre</th>
<th>Average processing time [ms/page]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy</td>
<td>44.45</td>
</tr>
<tr>
<td>Girl/woman</td>
<td>61.01</td>
</tr>
<tr>
<td>Youth</td>
<td>45.05</td>
</tr>
</tbody>
</table>

B. Success Rate of Frame Decomposition

In this section, we verify the accuracy of the proposed frame decomposition method. We compared the frame region information generated by the Comic-CFD with true information generated manually. As shown in Figure 11, if the difference in coordinates of all the four corner points is within a distance of K pixels, then correct region recognition is deemed to have been achieved. Based on our experience, we set the value of K to 30 pixels.

Figure 11: Comparison of frame region distance.

Figure 12 shows the success rate of frame decomposition. For evaluation, the Precision given by Equation (6), the Recall ratio given by Equation (7) and the Degree of Accuracy given by Equation (8) were used.

\[
P = \frac{S}{N} \quad (6), \quad R = \frac{S}{C} \quad (7), \quad F = \frac{2PR}{P + R} \quad (8)
\]

S: Number of correct frames detected
N: Number of frames detected
C: Number of correct frames to be detected.

\[
P = \frac{S}{N} \quad (6), \quad R = \frac{S}{C} \quad (7), \quad F = \frac{2PR}{P + R} \quad (8)
\]

S: Number of correct frames detected
N: Number of frames detected
C: Number of correct frames to be detected.

B. Success Rate of Frame Decomposition

In this section, we verify the accuracy of the proposed frame decomposition method. We compared the frame region information generated by the Comic-CFD with true information generated manually. As shown in Figure 11, if the difference in coordinates of all the four corner points is within a distance of K pixels, then correct region recognition is deemed to have been achieved. Based on our experience, we set the value of K to 30 pixels.

Figure 11: Comparison of frame region distance.

Figure 12 shows the success rate of frame decomposition. For evaluation, the Precision given by Equation (6), the Recall ratio given by Equation (7) and the Degree of Accuracy given by Equation (8) were used.

\[
P = \frac{S}{N} \quad (6), \quad R = \frac{S}{C} \quad (7), \quad F = \frac{2PR}{P + R} \quad (8)
\]

S: Number of correct frames detected
N: Number of frames detected
C: Number of correct frames to be detected.

B. Success Rate of Frame Decomposition

In this section, we verify the accuracy of the proposed frame decomposition method. We compared the frame region information generated by the Comic-CFD with true information generated manually. As shown in Figure 11, if the difference in coordinates of all the four corner points is within a distance of K pixels, then correct region recognition is deemed to have been achieved. Based on our experience, we set the value of K to 30 pixels.

Figure 11: Comparison of frame region distance.

Figure 12 shows the success rate of frame decomposition. For evaluation, the Precision given by Equation (6), the Recall ratio given by Equation (7) and the Degree of Accuracy given by Equation (8) were used.

\[
P = \frac{S}{N} \quad (6), \quad R = \frac{S}{C} \quad (7), \quad F = \frac{2PR}{P + R} \quad (8)
\]

S: Number of correct frames detected
N: Number of frames detected
C: Number of correct frames to be detected.
C. Frame Sorting

Based on successful frame division, frame sorting accuracy was evaluated. We use MangaGet method for comparison. In order to evaluate sorting, the order of frames belonging to the same set is changed, and then the number of steps required to restore correct order is used as a performance measure. For example, a comic image having frames with order A, B, C is changed to order A, C, B. By interchanging B and C, the correct order can be restored in one step.

Using the evaluation method defined above, the results are shown in Figure 15.

The verification result shows that the MangaGet method gives a frame sorting success rate of 42% (corresponding to 0 steps) while the proposed Comic-CFD method gives a very high accuracy of 99%. The MangaGet method gives fewer mistakes for sorting from to bottom, but the number of mistakes increase when sorting from left to right. In addition, the Comic-CFD approach only gives sorting errors for the case shown in Figure 16. The error arises in this case since the final result depends on whether the division is done in the vertical direction or horizontal.

5. Conclusion

In this paper, we proposed and evaluated a method of frame decomposition and sorting using information bearing regions of the comic image for the purpose of displaying comics on a mobile phone screen. Frame decomposition results depend on the comic genre, with the girl/woman genre showing low precision while the boy and youth genre gave a high precision of about 80%. In addition, frame sorting gave a very high accuracy of 99% by the proposed method for all of genre, The boy, Girl/Women and Youth.

Acknowledgment

A part of this work was supported by MEXT Grant-in-Aid for Building Strategic Research Infrastructures.

References: