Baltic J. Modern Computing, Vol. 10 (2022), No. 4, pp. 647–663
https://doi.org/10.22364/bjmc.2022.10.4.04

On the Automatised Generation of a Japanese *Banduke* Table

Antoine BOSSARD

Graduate School of Science
Kanagawa University
2946 Tsuchiya, Hiratsuka
Kanagawa, 259-1293 Japan

abossard@kanagawa-u.ac.jp

Abstract. *Banduke* tables are a traditional Japanese mean of communicating rankings and other organisation charts. They are well-known in Japan especially thanks to their continued usage to announce details of sumo wrestling competitions. Such tables were also used in earlier centuries for instance to publish the names of the wealthiest individuals in the country. Traditional *banduke* tables feature a characteristic, visually impacting design which is difficult to reproduce on computer systems. And, the automatisation of the generation of such a chart is even more challenging. In this paper, in order to keep this cultural heritage alive, we investigate this automatisation issue and propose a system which generates *banduke* tables. This work notably starts with an analysis of the various elements and zones that are essential to the structure of a *banduke* table. After presenting technical details of the proposed generation system, like the generation flow of such an organisation chart, we address the critical character rendering issue. The details of character box calculations are formally described. The obtained results are very promising in that the charts that can be produced by our proposal look very realistic. Besides, we have also quantitatively evaluated the generated tables by comparing them to original documents, five in total, some of which were obtained with the help of museums of Japanese history. Both traditional tables and ones of our modern, printing era were used for this comparison experiment. (This paper is an extended version of Bossard (2022).)

Keywords: digitalisation, cultural heritage, Japan, character, language

1 Introduction

Traditionally, Japan has relied on organisation charts, called *banduke* (IPA: [banzɯke]), to make public, written announcements with respect to various purposes. For instance, and it is now one of the most well-known usage example of *banduke* tables, all the parties involved in sumo wrestling competitions, starting with wrestlers themselves, are detailed on such an organisation chart. Names and ranks are central to such documents.

Banduke tables are however not restricted to this characteristic sportive usage: such charts can be found in archives of libraries and museums, listing for example the wealthiest individuals in the country – such tables are typically called 長者 *chōja* "wealthy person" either 集 *shū* "list" (NMC, Tenpō period (1830–1844)), 鑑 *kagami* "example" (Katō, 1879) or 番付 *banduke* "organisation chart", literally "order assignment" (Japan Broadcasting Corporation (Nippon Hoso Kyokai), 2021; Kikuchi, 2015). This second sample usage seems also to be characteristic of *banduke* tables. A part of such an organisation chart is illustrated in Figure 1.

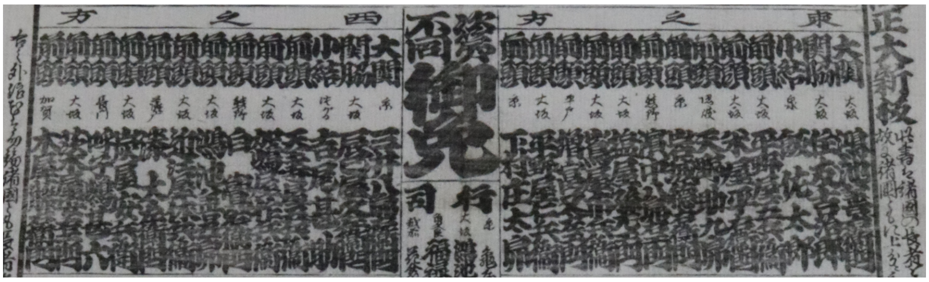


Fig. 1. A part of a *banduke* table (cropped from a photograph of NMC (Tenpō period (1830–1844)). The "crowded" aspect is characteristic of such organisation charts. (Reused from Bossard (2022))

Although characteristic of the Japanese society of the previous centuries, such listings remain part of the Japanese culture and some are thus still edited in our modern society. For example, H. Kikuchi authored a few years ago a book on post-war Japanese billionaires which is in the form of a *banduke* chart, albeit modern (Kikuchi, 2015). That is, its content conforms to that of a traditional *banduke*, but not its format: in order to benefit from facilitated editing as provided by mainstream word processing software and facilitated printing and commercial distribution, such modern editions are edited as regular books and have thus lost the typical design of *banduke* tables, and thus part of their appeal. An exception is made for sumo wrestling organisation charts (Japan Broadcasting Corporation (Nippon Hoso Kyokai), 2021), which retain the traditional, visually impacting design of ancient tables (e.g. Katō (1879); NMC (Tenpō period (1830–1844)).

The objective of this research is thus to enable automatic generation of a *banduke* table that retains their traditional, characteristic style in order to keep this cultural heritage alive for the current and next generations. In other words, we investigate and measure whether it is feasible and practical to supersede the arguably bland typesetting of modern *banduke* table editions without having to resort to manual typesetting.

The importance of ancient document preservation especially by means of information technology has been acknowledged for years. Related works that concern restoration of ancient documents are numerous; we can cite Farrahi Moghaddam and Cheriet (2009), Drira et al. (2009) and Kim et al. (2019) as examples. In this paper we take on the related albeit less researched issue of document generation preserving ancient characteristic features, such as elements of style, by means of information technology.

The rest of this paper is organised as follows: the basic structure and core properties of a *banduke* table are reviewed in Section 2. The proposed automatic generation system is described in Section 3 and experimentally evaluated in Section 4. Finally, Section 5 concludes this paper. This paper is an extended version of Bossard (2022).

2 Preliminary Analysis of *Banduke* Tables

In this section, we briefly recall essential properties of the considered writing system, Japanese, before describing in detail the structure of a typical *banduke* table.

2.1 Writing System

First and foremost, we make a recall about the elements of the Japanese writing system. It is based on three main categories of characters: Chinese characters, called *kanji*, which represent the vast majority of the characters of the Japanese writing system, and the *kana* characters of the two syllabaries *hiragana* and *katakana*. Chinese characters are logograms, each having at least one reading and one meaning (yet often several) whereas the *kana* characters are used only for their readings and thus convey no semantic information (Bossard, 2018).

In most cases, these characters are stacked vertically: in the Japanese writing system, characters traditionally flow from top to bottom and such character columns are positioned on the page from right to left (Lunde, 2009). And in some particular situations, such as titles or shop signs, characters can also be found stacked horizontally from right to left. This latter scenario is almost exclusively restricted to single line phrases though. (This is to be compared with the writing systems of Latin cultures: characters are stacked horizontally from left to right, with such character rows positioned on the page from top to bottom.) So, in the case of *banduke* tables, most text is typeset vertically and from right to left, with some exceptions, usually single lines, such as for titles which are typeset horizontally and from right to left.

2.2 Table Content and Structure

We next make several observations regarding the structure of a typical *banduke* table. These are obviously general guidelines which are more or less followed by *banduke* authors. We mostly relied on the three tables Japanese Sumo Association

(Nihon Sumo Kyokai) (2021); Katō (1879); NMC (Tenpō period (1830–1844)) which are of different origins and periods.

Such a chart can be typeset in both portrait or landscape mode, and it usually consists in five or six layers of persons' names. Layers are stacked vertically. They include lists of names written vertically from right to left, thus in accordance with traditional Japanese writings. The table is horizontally split at the centre to separate two sets of such layers: the left layers, called the "west" (西 *nishi*), and the right layers, called the "east" (東 *higashi*). East and west layers are separated by a vertical column at the centre which spans the entire table height and which includes additional information such as table details and sponsor information. A *banduke* table can have a title which spans its width or height, although some have no such title: they only rely on the middle column text to provide such information.

The typical ranks found in a *banduke* table are as follows (in descending order of importance): 横綱 *yokoduna*, 大関 *ōzeki*, 関脇 *sekiwake*, 小結 *komusubi* and 前頭 *maegashira*. It can be noticed that other, lower ranks do exist, for example 十両 *jūryō* and 幕下 *makushita* for sumo wrestling, but the corresponding titles are usually not explicitly displayed in the main, printed organisation charts (refer to Table 1(b2)).

We now detail the structure of these five to six table layers. As explained, a layer is a list of entries, such as persons' names. Each entry of a layer usually consists of three parts: the rank, at the top, the location (e.g. prefecture, district), at the middle, and the name, at the bottom. So, each layer usually consists of three rows: the rank row, the location row and the name row. There may be of course slight variations; for instance, the rank row is abbreviated for the lower layers of NMC (Tenpō period (1830–1844)).

Regarding the topmost layer, say layer 1, each entry is fully detailed: ranks are written without any abbreviation, locations are repeated even if appearing for two consecutive entries. See Figure 1.

Regarding the next layer (i.e. the layer below layer 1), say layer 2, typically the rank of the first entry (i.e. the rightmost one) is fully written but that of the subsequent entries is abbreviated to only the character 同 *dō* "same". Similarly, when the location of several consecutive entries is identical, the first location is fully written and the succeeding ones are abbreviated to only that same *kanji* character 同.

The next layer, say layer 3, typically either completely omits ranks and abbreviates repeated consecutive locations with an even simpler glyph such as 〃 or 冂 (which corresponds to 同), or slightly abbreviates the rank of the first entry, for example 前頭 *maegashira* becomes simply 前 *mae*, and succeeding ranks are denoted by 冂, just as repeated consecutive locations.

Layer 4 can be similar to the previous layer or can be even more abbreviated: for instance, repeated consecutive locations become the simple glyph 丶. The remaining layers are typeset on a par with the fourth layer.

A typical *banduke* table structure is illustrated in Figure 2. This figure shows two such organisation charts side by side (sample *banduke* tables taken from

Tōkyō Banduke Chōsakai (東京番附調査会) (1923)) so as to clearly designate the various table elements, without overlapping. The reader can also refer to the generated table given in appendix for an illustration.

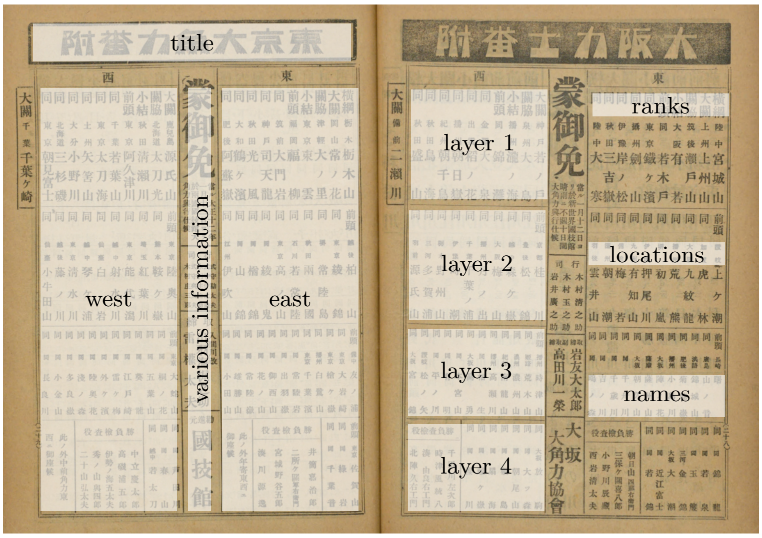


Fig. 2. Structure of a typical *banduke* table: the major elements are labelled. Two such organisation charts are here shown side by side (sample *banduke* tables taken from Tōkyō Banduke Chōsakai (東京番附調査会) (1923)) so as to clearly designate the various table elements. (Reused from Bossard (2022))

Finally, it is important to mention that some glyphs used inside *banduke* tables are characteristic of such organisation charts and thus not present in conventional character encodings, such as Unicode (The Unicode Consortium, 2021), even with support for the ideographic variation database (IVD).

3 Methodology

In this section, general information on character rendering in our approach is presented before giving an overview of the system, which is followed by details on character processing, especially character stretching and kerning. Our implementation has been realised with the Scheme-based Racket programming language and framework (Flatt, 2012).

3.1 Character Rendering

Two aspects are essential in our approach: a traditional *banduke* table is visually appealing first and foremost because of the font it uses and tightly spaced characters (sometimes even overlapping). So, no satisfactory result could be obtained without an appropriate font. To this end, we have selected a font that is similar to the one used to typeset *banduke* tables for sumo wrestling competitions (Japan Broadcasting Corporation (Nippon Hoso Kyokai), 2021). It should be noted that the proposed system enables to seamlessly change fonts, which is obviously a significant improvement over manual production. Font selection is easily done in Racket with, for instance, the `make-font` procedure, which we use.

Second, for the realised digitalisation method to achieve optimal graphical rendering quality, notably when printing is required, the produced output needs to be based on vector graphics, and not raster graphics (i.e. bitmap). Vector graphics can be easily obtained in Racket by using, for instance, a PDF drawing context (`pdf-dc%`); we relied on this solution. Another advantage of our approach is that it induces a high accessibility since it retains character (string) information (e.g. the text is selectable, and thus searchable inside, say, the generated PDF document), which is key for instance to the text-to-speech accessibility feature (PDFlib GmbH, 2015). Furthermore, accessibility is increased thanks to vector graphics as detailed in Rotard et al. (2004). Finally, we would like to insist on this vector graphics feature since a *banduke* table could perhaps be more easily produced with bitmap graphics, that is using raster renderers, filters and so on.

3.2 System Overview

An overview of the *banduke* table generation process is shown in Figure 3: the *banduke* table generator component retrieves necessary character font, type settings and table dimension information from the respective components and subsequently generates the table either for on-screen display or as PDF, that is, vector graphics as explained previously. The generated sample tables included in this paper, except that of Figure 4, have been produced according to the latter scenario.

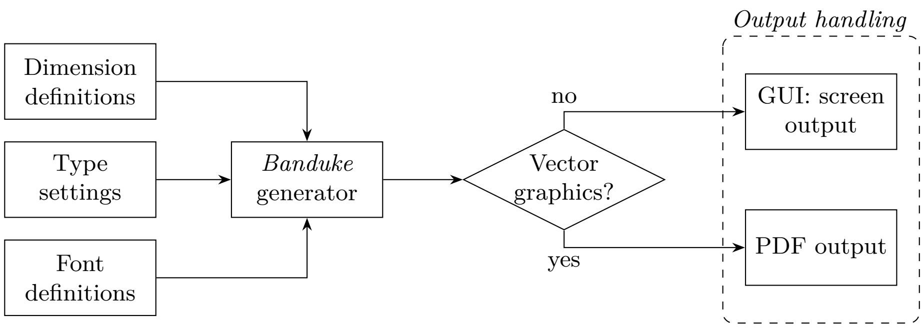


Fig. 3. An overview of the *banduke* table generation process.

Two fonts are used by the proposed system: as discussed, the characteristic *banduke* font as seen for instance on Sumo competition tables is near mandatory to achieve satisfactory results. This font is used for the larger font sizes of the generated tables, typically the names and ranks of layers 1 and 2 and the chart title, amongst a few other uses. Precisely, we have relied on the DF *sumōtai* (DF相撲体) font made by DynaComware. A second font is used to typeset smaller characters, which would not be readable enough with the main, large font. To this end, we have used the HG *gyōshotai* (HG行書体) font made by Ricoh and provided by Microsoft. Both fonts are TrueType, which is critical to be able to output vector graphics.

Next, the dimension definitions are detailed. The *banduke* table is structured according to what we call "zones", inside which text is automatically filled-in as explained in the rest of Section 3. We have declared in total 42 zones; they are materialised (highlighted) and labelled in Figure 4. Vertical typesetting applies to the vast majority of zones, with the notable exception of the title and subtitle zones. This figure additionally provides an illustration of the GUI (i.e. screen) output scenario. Note that such zones can be obviously freely adjusted upon needs (removed, added, shrunk, expanded).

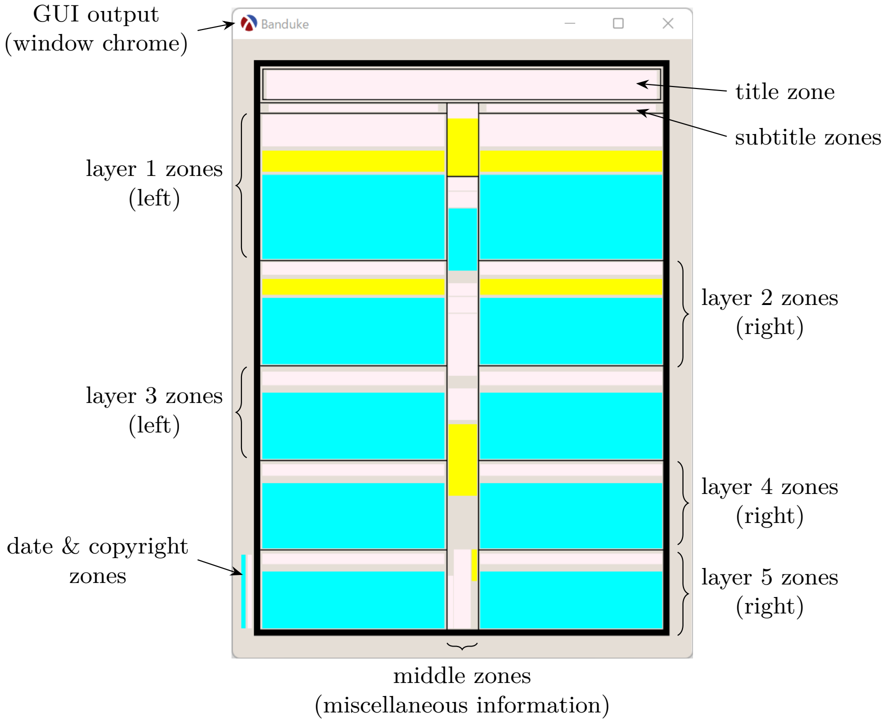


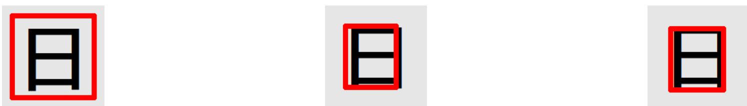
Fig. 4. Detailed view of the zones that structure the generated *banduke* tables. In addition, this figure illustrates the GUI (i.e. screen) output scenario.

3.3 Character Stretching and Kerning

The key issue of the character output routine is to stretch characters so that they fit into their destination area, but at the same time enforcing a maximum stretch ratio so that characters remain humanly readable, obviously. So, while kerning (i.e. the space between two consecutive characters) is set to the value provided as parameter, this value is in fact treated as the minimum kerning: when a stretching factor that is larger than the maximum tolerated stretch ratio would be required to make the character string fit its destination area, characters are stretched according to the maximum tolerated stretch ratio and kerning is automatically adjusted between characters so that they perfectly fit into their destination area.

The first task is to calculate the bounding box of each character, that is, to retrieve the precise width and height of each glyph. The `get-text-extent` method of the drawing context of the window canvas did not return satisfactory results. The best solution was to first retrieve the outline of the character as a path with the `text-outline` method of the `dc-path%` class and then retrieve the corresponding bounding box. The `get-path-bounding-box` method did not produce satisfactory results, although they were significantly better than those obtained with the above first method. Satisfactory results were eventually obtained by relying instead on the `get-bounding-box` method, which was applied to the result of `text-outline`.

The successive results are illustrated in Figure 5, the red frame showing the calculated bounding box for the sample character 日 “day” (here typeset with the system default Japanese font), and the acceptable method is detailed in Listing 1.1 (the source code has been simplified for the sake of clarity); the `dc` variable represents therein the drawing context that is used for graphical rendering.



(a) (b) (c)

Fig. 5. The result of each of the three bounding box (red frame) calculation solutions for the sample character 日 “day”: (a) with the `get-text-extent` method; then with the `text-outline` method combined with (b) the `get-path-bounding-box` method and with (c) the `get-bounding-box` method. (Reused from Bossard (2022))

Then, this bounding box calculation method is applied to each character of the string to print out in order to calculate the dimensions (i.e. width and height) of the entire string. To this end, kerning that we call “manual” since manually specified (constant) in the program is added between every two consecutive characters of the string. There is no need to adjust such parameter: the default value is satisfactory.

Listing 1.1. Calculation of the bounding box of each character with the `get-bounding-box` method.

```
(let ([path-dc (new dc-path%)])
  (send path-dc text-outline (send dc get-font) "日" 0 0)
  (let-values ([(left top width height)
                (send path-dc get-bounding-box)])
  (list left top width height))) ; return the bounding box
```

Then, we deduce the tentative scaling factors (i.e. horizontal and vertical) from the specified destination area and the calculated string dimensions. These obtained scaling factors remain tentative: it is indeed possible to setup a maximum stretch ratio m, which acts as a threshold to determine whether kerning should be automatic or manual. If the tentative scaling factors do not induce a stretch ratio which goes beyond this threshold, then the scaling factors are fixed. Otherwise, that is when the tentative scaling factors would overly stretch characters, the constant kerning value is discarded in favour of an automatic one which satisfies the threshold; in practice, the scaling factors are calculated so that character stretching equals the maximum stretch ratio, and space is evenly distributed between consecutive characters.

Let us consider vertical typesetting, which is the vast majority of the character strings of a *banduke* table. The horizontal scaling factor is directly fixed to $s_h = w_r/w_s$, with w_r the width of the destination rectangle and w_s the aggregated width of the character string (i.e. the width of the largest character since a vertical character string). The vertical scaling factor is tentatively set to $s_v = h_r/h_s$, with h_r the height of the destination rectangle and h_s the aggregated height of the character string. If $s_v/s_h > m$ holds, the vertical scaling factor s'_v is fixed to $m \times s_h$, and otherwise to s_v. Then, the "smart" kerning value calculation described above when $s_v/s_h > m$ holds is detailed in Listing 1.2 (the source code has been simplified for the sake of clarity). The `bounding-boxes` list contains the bounding box of each character of the string to print out.

Listing 1.2. Calculation of the automatic kerning value between consecutive characters. The value s'_v is the calculated vertical scaling factor.

```
(let ([remaining-space ; remaining space inside destination rectangle
       (- rect-height ; height of the destination rectangle
         (foldl (lambda (box current-height)
                  (let ([height (* (fourth box) s'_v)])
                    (+ height current-height)))
                0 bounding-boxes))])
  (/ remaining-space (sub1 (length characters)) s'_v))
```

A sample vertical typesetting output in the case of manual kerning only and in the case of automatic kerning are given in Figures 6a and 6b, respectively. In this example, the manual kerning method overly stretches characters, and this is corrected with the automatic kerning method: space is evenly distributed between consecutive characters so as to satisfy here a maximum stretch ratio vertical:horizontal of 2:1. (This ratio is of course a sample value: any maximum ratio can be specified.)

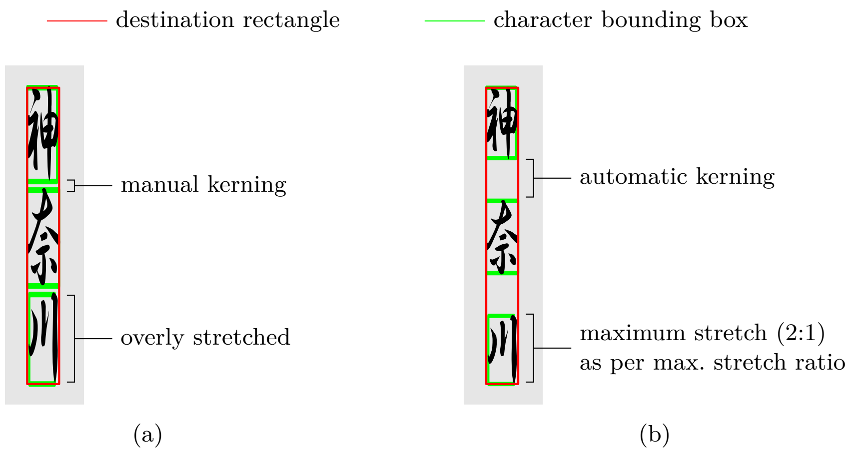


Fig. 6. (a) Manual kerning induces overly stretched characters. (b) Automatic kerning evenly distributes space between consecutive characters so as to satisfy here a maximum stretch ratio vertical:horizontal of 2:1. (Reused from Bossard (2022))

The flow of this process in the case of vertical typesetting (i.e. used for most zones) is summarised in Figure 7 and that in the case of horizontal typesetting (e.g. used for the title and subtitle zones) in Figure 8.

3.4 Character String Filling and Margins

Then, such character string fitting is automatically repeated in order to fill in large zones of the *banduke* table. This step is thus about the automatic calculation of the position and dimensions of numerous destination areas, one for each entry of the chart. It also requires to take into account the margin value to be inserted between two consecutive character strings; such margin can be set to a negative value to reproduce the "crowded" aspect of *banduke* tables. Although rather tedious, such calculations are omitted here for the sake of conciseness. The adjustment of this margin setting is illustrated in Figures 9a and 9b: a positive and a negative margin, respectively.

Finally, in order to streamline this character string filling process, we have defined several type settings which are reused as the table is being generated.

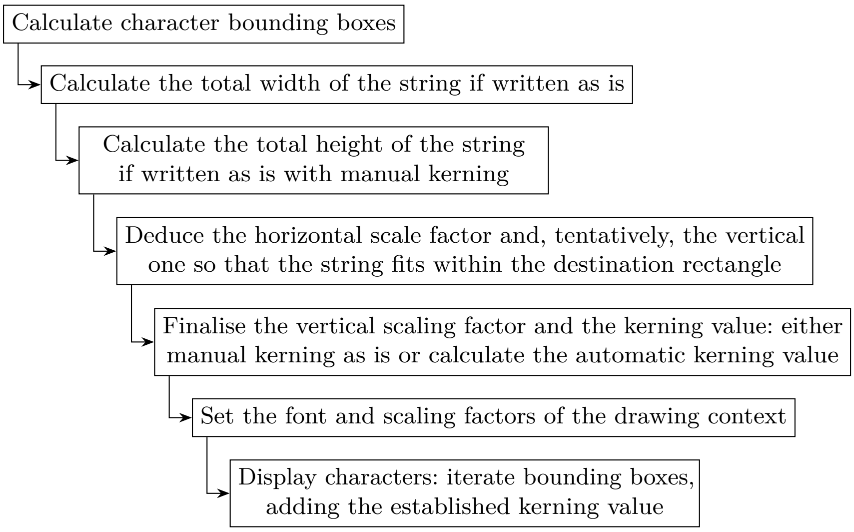


Fig. 7. A summary of the flow of the vertical typesetting process.

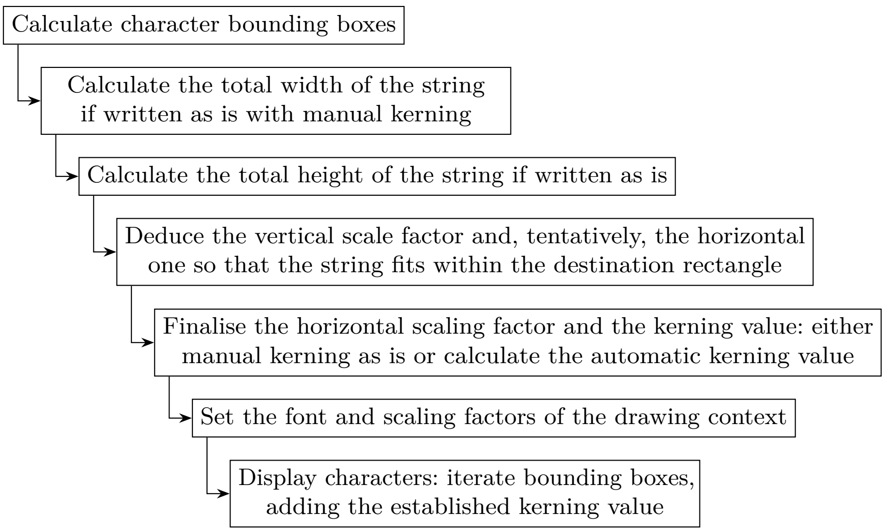


Fig. 8. A summary of the flow of the horizontal typesetting process.

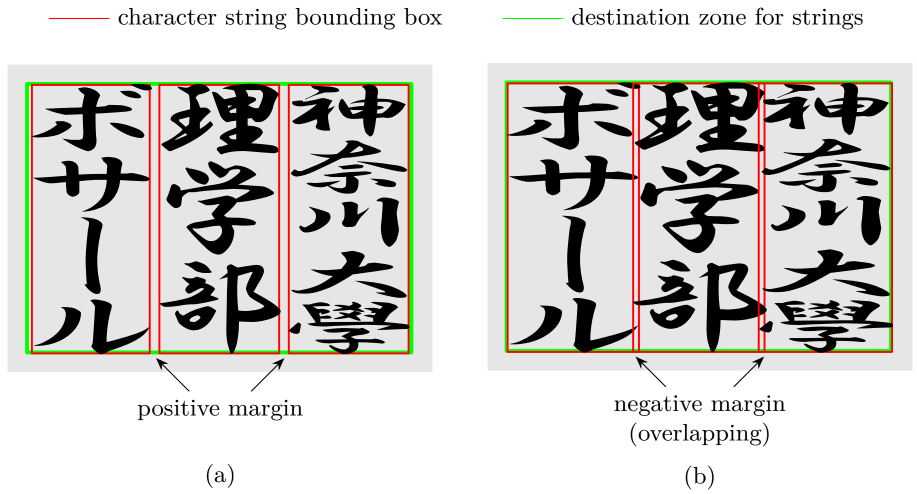


Fig. 9. An illustration of the margin setting between consecutive character strings: (a) a positive margin, (b) a negative margin, with thus character string bounding boxes overlapping. (Reused from Bossard (2022))

One single type setting is applied for the filling of each zone; it consists of the following elements:

- a maximum stretch ratio (refer to Section 3.3);
- the font to use (refer to Section 3.2);
- the manual kerning value (refer to Section 3.3).

In total, we have defined seven such type settings, five for vertical typesetting and two for horizontal typesetting. Note that although we have distinguished two sorts of type settings (vertical and horizontal), their structures are identical.

4 Experimental Evaluation

We have quantitatively evaluated the results obtained from the proposed generation system. This is detailed next.

4.1 Experimental Conditions and Results

An illustration of a generated sample *banduke* table is given in appendix; the results are very promising in that the produced chart looks very realistic. Now, in an attempt to quantitatively evaluate the proposal, we have generated a *banduke* table with the proposed automatised generation system (Table 1(b0)) and we have compared the obtained chart with three existing ones: two organisation charts for a sumo wrestling competition (Table 1(b2), (b3), respectively taken from Japanese Sumo Association (Nihon Sumo Kyokai) (2021) and Japanese

Sumo Association (Nihon Sumo Kyokai) (2014)) and that of NMC (Tenpō period (1830–1844)) (Table 1(b1)). Then, after shearing and resizing these two original tables so that their dimensions very roughly match those of the generated table, we have applied several filters to the obtained images in order to more easily visualise the difference between the generated table and original ones.

Precisely, we have calculated the difference between images, then converted the result to a monochrome image and, finally, we have pixelized the result (8×8 pixels) to better identify similarities and differences with the original *banduke* tables of Table 1(b1), (b2) and (b3). The corresponding results are shown in Table 1(d1), (d2) and (d3), respectively.

In addition, we have conducted the same experiment with the two "modern" (i.e. realised with a modern typesetting) charts of Figure 2 in an attempt to further quantify the contribution and achievements of our proposal. The obtained results are shown in Table 2(d4) and (d5).

4.2 Discussion of the Results

The result of the calculated image difference should be understood as follows: black pixels (black areas) designate similarity between the original and generated charts. And white areas accordingly emphasise dissimilarities between the two images.

The two results (d2) and (d3), that is the difference between a *banduke* table for sumo wrestling competition and the generated one, reveal larger similarities in the bottom half of the chart. It should be noted that the table generated with the proposed system has been filled with generic, mostly Japanese names, that is, not the same names that appear on the sumo wrestling table so as to avoid any possible copyright infringement.

This result (d2) induces a similarity rate of approximately 33% and the result (d3) a similarity rate of approximately 32%. One should note that this value is only given for reference: it could be easily increased by adjusting the table content so as to match the original. Rather than focusing on the exact similarity value, the purpose of this experiment is to identify areas inside the generated chart where improvements could be made in an attempt to reproduce such a *banduke* table with a higher fidelity.

On the other hand, the result (d1), that is the difference between a *banduke* table which lists wealthy individuals and the generated one, reveals larger similarities in the top half of the chart. This result (d1) induces a similarity rate of approximately 38%. These results also show how much of each table we have considered to define the proposed automatic generation system.

Regarding the second experiment which compares the generated table with modern typesetting *banduke* tables, (d4) and (d5) respectively induce a similarity rate of approximately 46% and 42%, with larger similarities in the bottom half of the two charts. This is can be explained by the fact that the bottom half of the generated chart, just as for traditional tables (e.g. NMC (Tenpō period (1830–1844)), is typeset with a lighter, less characteristic font, just as the whole tables (b4) and (b5). These results are thus positive and quantitatively show

Table 1. The difference (Δ) between original *banduke* tables and the generated one has been calculated and emphasised with pixelization. (b1) is a photograph of NMC (Tenpō period (1830–1844)), (b2) is a photograph of Japanese Sumo Association (Nihon Sumo Kyokai) (2021), (b3) is a photograph of Japanese Sumo Association (Nihon Sumo Kyokai) (2014) and (b0) is the generated table. (Partially reused from Bossard (2022))

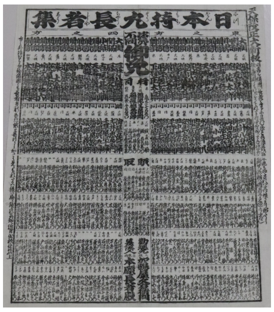
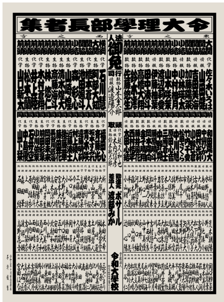
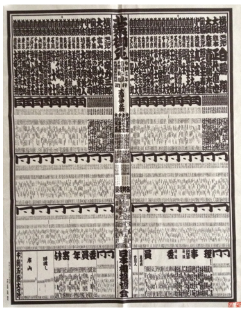
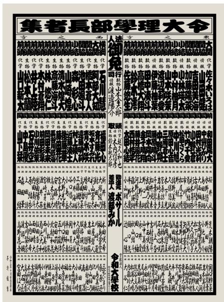
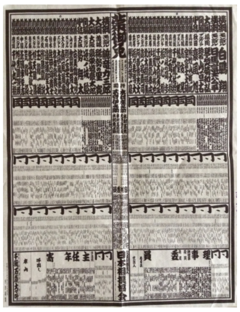
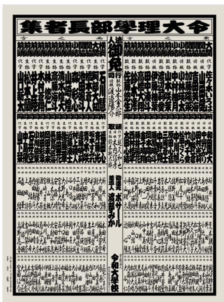
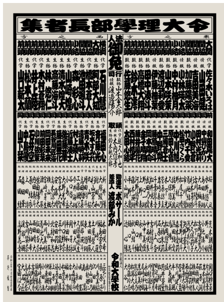
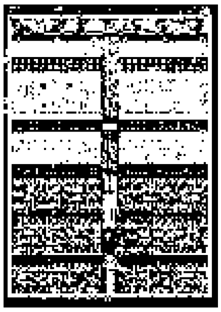
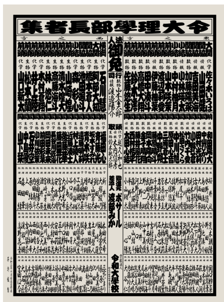
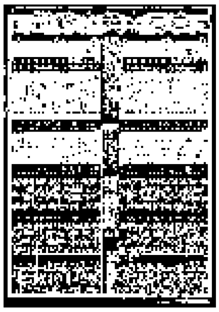
Image 1		Image 2		Difference
	Δ		→	
(b1)		(b0)		(d1)
	Δ		→	
(b2)		(b0)		(d2)
	Δ		→	
(b3)		(b0)		(d3)

Table 2. The difference (Δ) between original *banduke* tables and the generated one has been calculated and emphasised with pixelization. (b4) and (b5) are a digitalisation of two pages of Tōkyō Banduke Chōsakai (東京番附調査会) (1923) and (b0) is the generated table.

Image 1		Image 2		Difference
 (b4)	Δ	 (b0)	→	 (d4)
 (b5)	Δ	 (b0)	→	 (d5)

that our proposal succeeded in enhancing the reproduction fidelity of traditional *banduke* tables.

5 Conclusions

Banduke tables are traditional Japanese organisation charts which are used for various purposes. Well-known usage examples include sumo wrestling competition tables and national wealth rankings. Due to their complex layout, so as to benefit from computer technologies the produced modern charts are mostly uncharacteristic, mere lists such as those published in books. In this paper, we have investigated the feasibility of automatising the creation of *banduke* tables while, importantly, retaining their characteristic design. The obtained results are very promising: the generated organisation chart looks very realistic. In addition, we have quantitatively measured the results in an attempt to identify high and low similarity areas in the output.

Regarding future works, some properties of vertical typography, which is an inherent part of the Japanese writing system, such as the rotation of several glyphs, remain to be implemented. In addition, the declaration of the zones that structure generated *banduke* charts is a rather tedious process: simplification or even automatisation with supporting tools to define them would be another meaningful future work. And more generally, an interface could be realised to facilitate the usage of this *banduke* table generator.

Acknowledgements

The author is grateful towards the Miyazaki Prefectural Museum of Nature and History (Japan) for providing and granting permission to use the *banduke* table of NMC (Tenpō period (1830–1844)) for this study. The *banduke* tables included in this paper are copyright of their respective authors and owners.

References

Bossard, A. (2018). *Chinese characters, deciphered*, Kanagawa University Press, Yokohama, Kanagawa, Japan.

Bossard, A. (2022). On the automatisation of the realisation of a *banduke* table, *Proceedings of the 15th International Baltic Conference on Digital Business and Intelligent Systems (DB&IS; Riga, Latvia, 4–6 July)*, Vol. 1598 of *Communications in Computer and Information Science*, Springer, pp. 80–92. DOI: 10.1007/978-3-031-09850-5_6.

Drira, F., LeBourgeois, F., Emptoz, H. (2009). Document images restoration by a new tensor based diffusion process: Application to the recognition of old printed documents, *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR; Barcelona, Spain, 26–29 July)*, pp. 321–325. DOI: 10.1109/ICDAR.2009.109.

Farrahi Moghaddam, R., Cheriet, M. (2009). RSLDI: Restoration of single-sided low-quality document images, *Pattern Recognition* **42**(12), 3355–3364. New Frontiers in Handwriting Recognition. DOI: `10.1016/j.patcog.2008.10.021`.

Flatt, M. (2012). Creating languages in Racket, *Communications of the ACM* **55**(1), 48–56. DOI: `10.1145/2063176.2063195`.

Japan Broadcasting Corporation (Nippon Hoso Kyokai) (2021). Ōzumō banduke tte, ittai nannano? (大相撲 番付って、いったい何なの？), `https://www3.nhk.or.jp/sports/story/13400/index.html`. In Japanese. Last accessed in April 2022.

Japanese Sumo Association (Nihon Sumo Kyokai) (2014). *Banduke* table (番付表) of October 27th, 2014. In Japanese.

Japanese Sumo Association (Nihon Sumo Kyokai) (2021). *Banduke* table (番付表) of March 1st, 2021. In Japanese.

Katō, T. (1879). Dainippon marumochi chōja kagami (大日本持丸長者鑑). In Japanese.

Kikuchi, H. (2015). *Nihon no chōja banduke: sengo okuman chōja no seisui (*日本の長者番付 : 戦後億万長者の盛衰*)*, Heibonsha, Tokyo, Japan. In Japanese.

Kim, Y. J., Hazra, D., Byun, Y., Ahn, K.-J. (2019). Old document restoration using super resolution GAN and semantic image inpainting, *Proceedings of the International Workshop on Artificial Intelligence and Education (WAIE; Singapore, 25–27 November)*, pp. 34–38. DOI: `10.1145/3397453.3397459`.

Lunde, K. (2009). *CJKV information processing*, second edn, O'Reilly Media, Sebastopol, CA, USA.

NMC (Tenpō period (1830–1844)). Nihon marumochi chōja shū (日本持丸長者集). In Japanese.

PDFlib GmbH (2015). A technical introduction to PDF/UA. PDFlib Whitepaper.

Rotard, M., Otte, K., Ertl, T. (2004). Exploring scalable vector graphics for visually impaired users, *in* Miesenberger, K., Klaus, J., Zagler, W. L., Burger, D. (eds), *Computers Helping People with Special Needs*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 725–730. DOI: `10.1007/978-3-540-27817-7_108`.

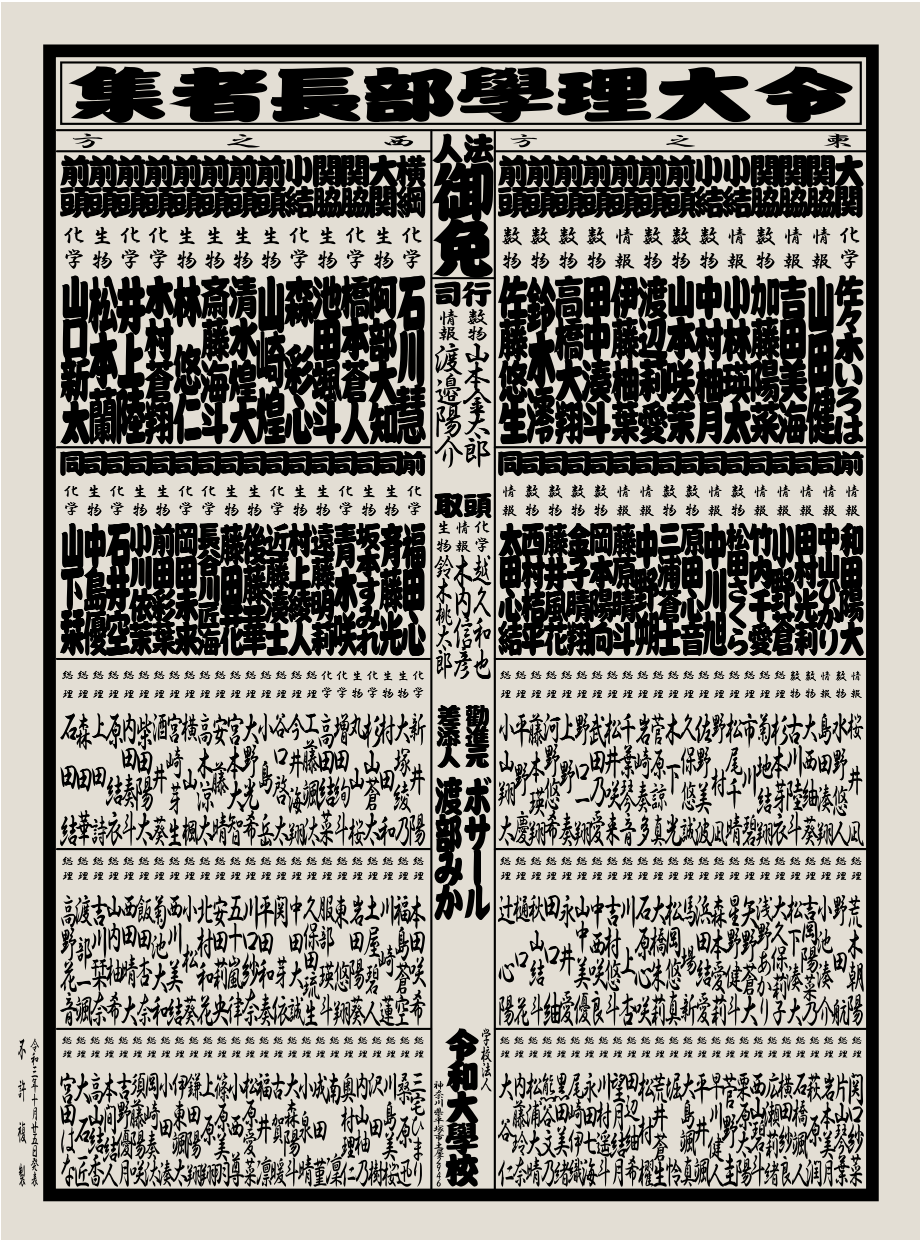
The Unicode Consortium (2021). *The Unicode Standard, Version 14.0*, Print on demand.

Tōkyō Banduke Chōsakai (東京番附調査会) (1923). *Kinko ō-banduke: nanajū yorui (*今古大番附 : 七十余類*)*, Bunsankan Shoten, Tokyo, Japan. Retrieved from the National Diet Library of Japan (identifier no. info:ndljp/pid/1185061). DOI: `10.11501/1185061`.

A Complete Sample Output

We reproduce the generated sample *banduke* table of Table 1(b0) in a full-page sized version below to allow for a more precise inspection. It be noted that this table has been realised for research purposes only: names are entirely fictitious; any similarity to actual persons, organisations or other entities is purely coincidental.

Received October 7, 2022 , accepted November 29, 2022


令大理學部長者集
東之方
西之方
法人
御免
勧進元 ボサール
差添人 渡部みか
学校法人
令和大學校