Analytical Study on Typeface Visual Identification

Xavier Molinero[1,*,a], Josep Freixas[1,b] and Montserrat Tàpias[2,c]

[1]*Department of Mathematics, Universitat Politècnica de Catalunya · BarcelonaTECH, Spain*

[2]*Department of Optics and Optometry, Universitat Politècnica de Catalunya · BarcelonaTECH, Spain*

Keywords: Visual Identification, Measure of Topographies, Power Indices.

Abstract: In this study, our objective is to explore methodologies for the identification of diverse typefaces. Utilizing the gathered data, we conducted a thorough analysis of the outcomes, distinguishing between successes and failures for both uppercase and lowercase letters within each typeface. The analytical framework is anchored in three distinct recognition measures. The initial measure draws upon the relative frequency of accurate responses, providing insights into the overall performance of each typeface. The second measure is constructed around the F-score derived from confusion matrices, offering a comprehensive evaluation of recognition precision and recall. Lastly, the third measure is formulated on the well-established Shapley-Shubik index, extensively scrutinized and endorsed within the realm of classical game theory. This multifaceted approach allows us to comprehensively assess the distinct aspects of typeface recognition, contributing to a nuanced understanding of their effectiveness and characteristics.

1 INTRODUCTION

Low vision is a common visual impairment, especially among older adults, which makes it difficult for individuals to read printed materials due to reduced clarity and sharpness of vision. Typographers have developed specific typographies to enhance identification for low vision readers, by increasing contrast, reducing glare, and optimizing letter size and spacing. Serif fonts such as Times New Roman and Garamond are considered more legible as serifs guide the eye along the text. Sans-serif fonts such as Arial and Verdana provide a clearer and simpler design, improving readability. However, designers must test their typographies with low vision users, as identification or readability is subjective and can vary depending on the individual's vision impairment. By optimizing typographies for low vision readers, we can make printed materials more inclusive and improve accessibility for all.

In this research, we introduce and discuss three distinct measures for typographic identification. The initial measure is described in Subsection 2.1. Following that, the second measure is introduced and detailed in Subsection 2.2. Lastly, Subsection 2.3 presents a novel measure, which is founded on the *Shapley-Shubik index* derived from game theory. This section provides a comprehensive definition and exploration of the application of the Shapley-Shubik index in the context of typographic recognition.

[a] https://orcid.org/0000-0002-5203-4347
[b] https://orcid.org/0000-0002-9033-9432
[c] https://orcid.org/0000-0003-2421-0718
[*]Corresponding author.

2 TYPOGRAPHIC RECOGNITION

From now on, we assume that each letter of a given typography $\mathcal{T}$ is presented n times. Let be $\overline{\mu}_{\mathcal{T}} : \mathcal{T} \times \mathcal{T} \to \mathbb{N}$, where $\overline{\mu}_{\mathcal{T}}(X,Y) = k$ is the number of times that observers said Y when X is shown. For instance, $\overline{\mu}_{\mathcal{T}}(X,Y) = 5$ means that 5 times the observers have said Y when X has been shown. Thus, the most *identificable* tipography should verify

$$\overline{\mu}_{\mathcal{T}}(X,Y) = \begin{cases} n & \text{if } X = Y \\ 0 & \text{if } X \neq Y \end{cases} \tag{1}$$

for all $X,Y \in \mathcal{T}$. Next, we analyze those dataset in three different ways, i.e., applied to three visual identifications of topographies.

Molinero, X., Freixas, J. and Tàpias, M.
Analytical Study on Typeface Visual Identification.
DOI: 10.5220/0012719300003708

In *Proceedings of the 9th International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2024)*, pages 124-127
ISBN: 978-989-758-698-9; ISSN: 2184-5034

2.1 Identification from Relative Frequency

In the initial phase, we examine the foundational statistical concept of relative frequency, as outlined by Spiegelhalter in 2019 (Spiegelhalter, 2019).

Given a typography $\mathcal{T}$, for each character $X \in \mathcal{T}$ that we show to observers, we determine the relative frequency of X, i.e., the number of times that we said X divided by n. This value can be computed as follows:

$$\mu_P(X) = \frac{\bar{\mu}_{\mathcal{T}}(X,X)}{\sum\limits_{Y \in \mathcal{T}} \bar{\mu}_{\mathcal{T}}(X,Y)} = \frac{\bar{\mu}_{\mathcal{T}}(X,X)}{n} .$$

In addition, we also define a general identifiable value for the given typography $\mathcal{T}$, based on the relative frequency:

$$\mu_P(\mathcal{T}) = \sum_{X \in \mathcal{T}} \mu_P(X) .$$

2.2 Identification from Confusion Matrix

In the subsequent step of our analysis, we employ *confusion matrices* (Blog at WordPress.com., 2024) as a fundamental analytical tool to assess a key metric associated with each typography $\mathcal{T}$, namely the *F-score*. Widely acknowledged in classification tasks, the F-score offers a nuanced evaluation by striking a balance between precision and recall, making it particularly pertinent to the nuances of typographic recognition. By leveraging confusion matrices, we scrutinize the performance of each typeface in terms of true positives, false positives, and false negatives, offering a comprehensive understanding of their recognition capabilities.

For a detailed exploration of the computational aspects and theoretical underpinnings of the F-score within the scope of our study, interested readers are directed to (Brabec et al., 2020; Wikipedia, 2024). These references delve into the intricacies of F-score calculations, elucidating the nuances of its application and interpretation in the context of our research. In particular, we consider $\mu_F(X) = \mu_{F-score}(X)$ defined by

$$\frac{2\texttt{TruePositive}}{2\texttt{TruePositive} + \texttt{FalsePositive} + 2\texttt{FalseNegative}},$$

where `TruePositive` is a test result which correctly indicates that a condition hold, `FalsePositive` is a result that indicates a given condition exists when it does not, and `FalseNegative` is a test result which wrongly indicates that a condition does not hold.

As we have done for the relative frequency, we define a general identifiable value for the given tipography $\mathcal{T}$ as follows

$$\mu_F(\mathcal{T}) = \sum_{X \in \mathcal{T}} \mu_F(X) .$$

Example. *Let $\mathcal{T}$ be a typography where $\{a,b,p,q\} \in \mathcal{T}$. Suppose that, after doing the experiments with 10 presentations for each letter, we obtain the following results:*

		Answered letter		
		a	b	c
Presented letter	a	7	2	1
	b	1	2	7
	c	7	1	2

From this table we are able to compute $\mu_F(a)$ as follows

$$\mu_F(a) = \frac{2 \cdot 7}{2 \cdot 7 + 2 + 2 \cdot 2} = \frac{14}{20},$$

where the red values are `TruePositive`, the blue ones are `FalsePositive`, the green ones are `FalseNegative`, and yellow numbers are the remaining values.

2.3 Identification from Shapley-Shubik Index

Thirdly, we define a specific weighted voting game (Taylor and Zwicker, 1999). In general, given n characters (or players) we assign a weight (non-negative integer) for each character and a quota q (a positive integer). We say that a coalition of characters S is winning or successful if and only if the sum of the corresponding weights is more or equal to q.

Subsequent to the consideration of a typography denoted as $\mathcal{T}$, we proceed to allocate a unique weighted voting game, denoted as $\Gamma_{SS}(X)$, for each distinct typographic element $X \in \mathcal{T}$. This allocation is carried out in accordance with a specific methodology designed to systematically capture the nuances of typographic recognition within the framework of weighted voting games. By tailoring the approach for each individual typeface element, this methodological precision allows for a detailed examination of the intricate dynamics governing typographic recognition. Consequently, this structured assignment enables a profound exploration of the interrelations between typographic elements and their corresponding recognition measures, offering insightful observations grounded in the principles of game theory.

Next, to the consideration of a typography denoted as $\mathcal{T}$, we proceed to allocate a unique weighted voting game for each distinct typographic element $X \in \mathcal{T}$, denoted as $\Gamma_{SS}(X)$. This allocation is carried out in accordance with a specific methodology designed to systematically capture the nuances of typographic identification within the framework of weighted voting games. In particular, $\Gamma_{SS}(X)$ is described as follows:

$$\Gamma_{SS}(X) := [q_{\mathcal{T}}(X); \bar{\mu}_{\mathcal{T}}(X,Y_1), \bar{\mu}_{\mathcal{T}}(X,Y_2), \ldots, \bar{\mu}_{\mathcal{T}}(X,Y_k)]$$

where Y_i, for $i \in \{1,\ldots,k\}$, are all characters of $\mathcal{T}$. Considering the 80% success rate, for each weighted voting game $\Gamma_{SS}(X)$, we stated the quota $q_{\mathcal{T}}(X)$ as the 80% of the total n presentations considered, i.e., $q_{\mathcal{T}}(X) = 0.8 \cdot n$. From these weighted voting game $\Gamma_{\mathcal{T}}(X)$, we are able to compute the corresponding Shapley-Shubik value of X with respect to $\Gamma_{\mathcal{T}}(X)$, denoted by $\mu_{ss}(X)$ (Shapley and Shubik, 1954). In essence, $\mu_{ss}(X)$ is the number of times that X is *pivot* (it makes that a coalition of characters will be successfully) divided by all possible permutations. Finally, we consider a general identifiable value for the given tipography $\mathcal{T}$ based in those values as follows

$$\mu_{SS}(\mathcal{T}) = \sum_{X \in \mathcal{T}} \mu_{SS}(X) .$$

Example. *Let $\mathcal{T}$ be a typography where $\{a,b,p,q\} \in \mathcal{T}$. Suppose that, after doing the experiments with 10 presentations for each carhacter, we obtain the following results:*

$$\begin{array}{l} \Gamma_{\mathcal{T}}(a) = [8;8,1,0,1] \to \mu_{SS}(a) = 1 \\ \Gamma_{\mathcal{T}}(b) = [8;2,8,0,0] \to \mu_{SS}(b) = 1 \\ \Gamma_{\mathcal{T}}(p) = [8;0,1,7,2] \to \mu_{SS}(p) = 2/3 \\ \Gamma_{\mathcal{T}}(q) = [8;0,0,4,6] \to \mu_{SS}(q) = 1/5 \end{array}$$

Thus,

$$\mu_{SS}(\mathcal{T}) = 1 + 1 + \frac{2}{3} + \frac{1}{5} = \frac{43}{15}.$$

Note that this last method is the main novelty of this work. It gives us a new point of view to determine whether a typography is indistinguishable or not. In fact, it let us to see how relevant is a character to make a coalition of characters a successfully coalition.

3 CONCLUSION AND FUTURE WORK

In this work, we define three different values or measure to identify *how good* is a given typeface or typography. Subsequently, it is necessary to apply these measures to specific typefaces and systematically assess their effectiveness. This comprehensive evaluation will not only provide insights into the performance of individual typefaces but will also contribute to a nuanced understanding of the applicability and robustness of each measurement method in the context of our study.

A prospective avenue for further research involves the exploration of alternative power indices, such as Banzhaf, Deegan-Packel, Holler, among others. This expansion should allow a more comprehensive understanding of the typographic recognition. Additionally, our future effors will also focus on an examination of typography recognition within the realm of social networks, where we intend to investigate the interplay and relationships among characters as a social network.

The final results help us to choose the best typography and improve the design of specific characters in the considered typography, in terms of visual identification tasks and readability (Braida et al., 2018). It let us to rank optotypes in order to classify them according to the visual identification.

ACKNOWLEDGEMENTS

J. Freixas and X. Molinero has been partially supported by funds from the Ministry of Science and Innovation grant PID2019-104987GB-I00 (JUVOCO) and the Catalan government [2021 SGR 01419 ALBCOM].

REFERENCES

Blog at WordPress.com. (18th March, 2024). https://classeval.wordpress.com/introduction/basic-evaluation-measures/. Basic evaluation measures from the confusion matrix.

Brabec, J., Komárek, T., Franc, V., and Machlica, L. (2020). On model evaluation under non-constant class imbalance. In Krzhizhanovskaya, V. V., Závodszky, G., Lees, M. H., Dongarra, J. J., Sloot, P. M. A., Brissos, S., and Teixeira, J., editors, *Computational Science – ICCS 2020*, pages 74–87, Cham. Springer International Publishing.

Braida, D., Ponzoni, L., Verpelli, C., and Sala, M. (2018). Chapter 8 - visual object recognition task: A translational paradigm to evaluate sustained attention across species. In Ennaceur, A. and de Souza Silva, M. A., editors, *Handbook of Object Novelty Recognition*, volume 27 of *Handbook of Behavioral Neuroscience*, pages 139–150. Elsevier.

Shapley, L. S. and Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, 48(3):787–792.

Spiegelhalter, D. (2019). *Simple games: Desirability relations, trading, pseudoweightings*. Pelican, Penguin Books.

Taylor, A. and Zwicker, W. (1999). *Simple games: Desirability relations, trading, pseudoweightings*. Princeton University Press, Princeton, NJ.

Wikipedia (18th March, 2024). https://en.wikipedia.org/wiki/F-score. F-score.