Received 28 March 2023, accepted 24 July 2023, date of publication 28 July 2023, date of current version 16 August 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3299594

COMMENTS AND CORRECTIONS

Corrections to "The Structural Information Potential and Its Application to Document Triage"

VLAD ATANASIU
Department of Informatics, University of Fribourg, 1700 Fribourg, Switzerland
e-mail: atanasiu@alum.mit.edu

This work involved human subjects in its research. The author confirms that all human subject research procedures and protocols are exempt from review board approval.

In the above article [1], a few of the sections require corrections and some additional experiments to strengthen the work.

1) In the above article [1, p. 13109, col. 1, §3, Section II-A-5]," add a reference to the herein Fig. 1 in the sentence beginning with "The one domain that is almost pathologically fascinated by structural irregularity is that of the arts and architecture." The purpose of this correction is to illustrate the application of structural irregularity in the domain of architecture, with a real-world example.
2) In the above article [1, p. 13110, col. 1, §4, Section II-B]," add a reference to the herein Fig. 2 in the sentence beginning with "The problem becomes more critical as the number of documents to be processed increases." The purpose of this correction is to illustrate with a real-world example the application of triage to documents.

FIGURE 1. Example of an architecture based on the concept of redundancy minimization of edge lengths and texture—City archives, Berne, Switzerland; :mlzd architekten biel-bienne. (Credits: Vlad Atanasiu, 2022.)

FIGURE 2. Does document triage matter? An internationally famous case in point that would seem to answer in the affirmative is provided by the so-called *Pentagon Papers*, a classified history of the Vietnam War between 1945 and 1967 prepared by the U.S. Department of Defense [2]. Its publication revealed how multiple U.S. administrations misled the public on the reasons for and actions taken in the Vietnam War, and had consequential repercussions on the legal aspects of freedom of speech and journalism in the United States. The "papers" consisted of more than 7,000 pages of documents that were covertly photocopied by Daniel Ellsberg of the RAND Corporation and made available to the press in 1971. Among the major practical problems with which the journalists had to contend was that of rapidly identifying the most informative content and collating related pages so that articles could be finished before editorial deadlines [3, p. 392], [4]. Layout-based classification is a sensible strategy for both tasks under such circumstances. The difficulties posed by document triage for the evaluation of the Pentagon Papers were fictionalized in a memorable scene of the film *The Post* (2017), depicting the staff of the Washington Post at work (Credits: 20th Century Fox):
– Is that…?
– Yep.
– It's not the full report but it's over 4,000 pages of it.
– Are these in order?
– I don't think so.
– There are no page numbers.
– Yeah, that's where the "top secret" stamps were. My source had to cut 'em off.
– I was supposed to retire on Friday.
– Ben, how we supposed to comb through 4,000 pages of material?
– They're not even loosely organized?
– The Times had three months.
– There's no way we can possibly get this done.
– He's right. We've got less than eight hours.

3) In the above article [1, p. 13113, equation 9, Section III-A-5]," read "≈ 0.6942" instead of "0.6942."

4) In the above article [1, p. 13121, Section IV-A]," add after §1 of col. 2, the following section. The purpose of this correction is to strengthen the evaluation of the structural information potential (SIP) method with a series of psychophysical experiments.

3) PSYCHOPHYSICAL EXPERIMENTS

Our understanding of SIP and other pattern-ordering methods may be deepened beyond what is possible through the use of visual evidence and quantitative evaluations, on the basis of empirical insights obtained from psychophysical experiments. The primary goal of experiments such as those described below is to evaluate the agreement between computational and perceptual pattern orderings, which is especially important for SIP as a model for human behavior. The discovery of further insights into the perceptual, cognitive, and other practical aspects of SIP—for example, in the context of the triage task—represents a secondary goal of the experiments.

Experimental design—The participants in the experiments were 10 students in digital humanities from the University of Bern, Switzerland (4 females, $MD = 32$ years, $RG = 22$–43, $SD = 5.98$). The stimuli consisted of the 104 pages of the *New Yorker* magazine described above [1, Fig. 31], each reproduced on a 3-by-4.5 cm (1.18-by-1.77 in) piece of paper. This size ensured, on one hand, that all pages could be laid out within the confines of a small table and viewed at a glance; on the other hand, the small size helped to focus on the overall layout pattern, reducing the salience of semantic information such as text and picture content. To further emphasize the pattern and inhibit reading, the reproductions were placed upside-down.

The concept of the uniform–clustered–regular pattern spectrum was explained to the participants with reference to representative page samples ([1, Fig. 10] a, b, d) and through discussion. To ensure that the concept had been effectively assimilated, a practical training exercise was conducted in which the participants were required to order 16 sample documents (uniformly selected from the library cards shown in Fig. 21), with corrective inputs from the experimenter.

Ordering experiment—This experiment aimed to evaluate the suitability of SIP as a model of human pattern ordering along the uniform–clustered–regular spectrum. To this end, the participants received a stack of randomly ordered stimuli and were required to order the page reproductions along the uniform–clustered–regular spectrum.

The differences between the perceptual and SIP orderings were quantified using the topological entropy method, and the statistical summary is presented in Fig. 3b. The median value of 0.57 indicates a sizable divergence of the perceptual ordering from the computational SIP ordering ($M = 0.57$, 95% confidence interval [CI] = [0.54, 0.60]). Nevertheless, the perceptual orderings are on average closer to SIP than are the other computational methods, with the median of the latter being 0.67 ($M = 0.66$, 95% $CI = [0.58, 0.74]$; Fig. 3a).

We were able to enhance our comprehension of the results by adopting a qualitative approach, based on spontaneous remarks made by the participants when commenting on the experiment, observations of their behavior, and post-experiment debriefings. First of all, participants commented on the tediousness of the task: in terms of the need to physically arrange so many stimuli, in cognitive terms (the difficulty associated with deciding on their proper order), and in motivational terms (due to the duration of the operation, ca. 20 to 30 minutes). Second, it was difficult for the participants to order the stimuli purely in terms of their layout without interpreting ink structures in terms of their semantic classes, such as pictures and titles, despite that some of these were identified as such precisely because of the layout that made them stand out. Third, the ordering of clustered patterns was more challenging than that of uniform and regular patterns; the latter could be intuitively comprehended as "mostly empty surfaces" and "homogeneous ink distributions," while the concept of scale-space clustering was less familiar. Perhaps the results would have been different with participants drawn from other demographics, such as document designers, artists, and computer scientists.

In conclusion, the foremost finding from this experiment is the difficulty of perceptual ordering for various practical and cognitive reasons. It also provides evidence for the poor performance of human pattern triage. Since the experiment was more relevant for revealing human factors in pattern ordering than for allowing an evaluation of SIP as a perceptual model, we designed a second experimental setup to address these issues.

Improvement experiment—To diminish the influence of factors affecting free perceptual ordering, in this experiment, the participants were presented with a fixed order of stimuli (the one produced by SIP) and asked to improve this ordering by shifting the page reproductions so as to create a transition from uniform to clustered to regular layouts that they deemed to be perceptually the most "smooth."

The numerical results, with median topological entropy of 0.38, now indicate a smaller, albeit non-trivial, divergence between computational and perceptual orderings for the improvement task relative to the case of the free ordering task ($M = 0.39$, 95% $CI = [0.33, 0.45]$; Fig. 3c).

The effect size of the participants may be better appreciated through a visual examination of the resulting patterns. Unlike the numerical assessment, this approach reveals that the improvements made were comparatively limited in terms of their number and the distance of pattern shifts along the spectrum. Fig. 3g shows a representative example of the improvement produced by participant MF, which is the closest to the median divergence from SIP; clearly, this ordering is quite similar to SIP (Fig. 3h) and far from the median ordering of other methods (Fig. 3e). In terms of qualitative results, the participants stated that this task was easier than the former, and, indeed, they completed the experiment faster (in ca. 15 to 20 minutes).

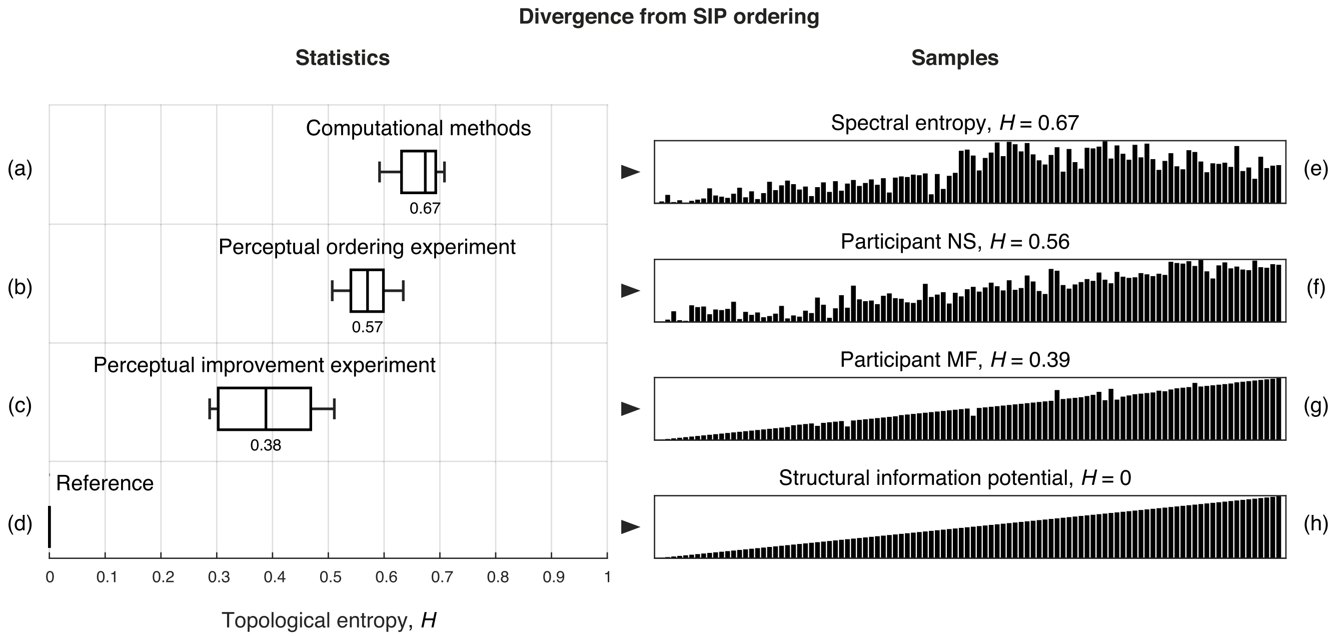

FIGURE 3. Results of the perceptual ordering and improvement experiments related to the divergence from the SIP-induced ordering. (a)–(d) Statistics on the distribution of topological entropy values, *H*, visualized as box-and-whisker plots: median, interquartile range, and range. "Computational methods" summarizes the data from [1, Fig. 19], while "Reference" refers to the value *H* of the reference SIP ordering. (e)–(h) Actual ordering samples from the corresponding statistical distributions, selected as the closest to the medians. The orderings of bars represent the permutations of the SIP ordering.

It appears from this experiment that SIP-induced ordering is very much accepted by humans as producing a smooth transition between uniform, clustered, and regular patterns. It also greatly facilitates triage.

Rating experiment—In this last experiment, we intended to compare on a perceptual basis the smoothness of the pattern spectra induced by various computational methods, which we assessed from a formal quantitative point of view in the preceding section, "Quantitative Evaluation" [1, Fig. 19]. To do this, the participants were requested to rate on a five-level scale the smoothness of the five analyzed computational orderings, along with the physical ordering of the pages, based on printed copies of Figs. 31–36 from [1].

The results are presented in Fig. 4. The rating aggregation method employed was majority judgement, chosen because it is a non-parametric ranking method that favors consensual candidates and is based on median scores rather than means [5]. Clearly, not only does SIP produce a highly consistent and smooth uniform–clustered–regular pattern spectrum, but all other methods are also given lower consistency ratings.

General assessment—In conclusion, SIP is a good model of perceptual pattern ordering along the uniform–clustered–regular spectrum, as well as aiding greatly in pattern-based triage. The psychophysical experiments described herein confirm and expand, from an empirical point of view, the conclusions derived from visual evidence and formal quantitative evaluation. In the next section, we evaluate SIP on a much larger dataset containing 1.2 million stimuli.

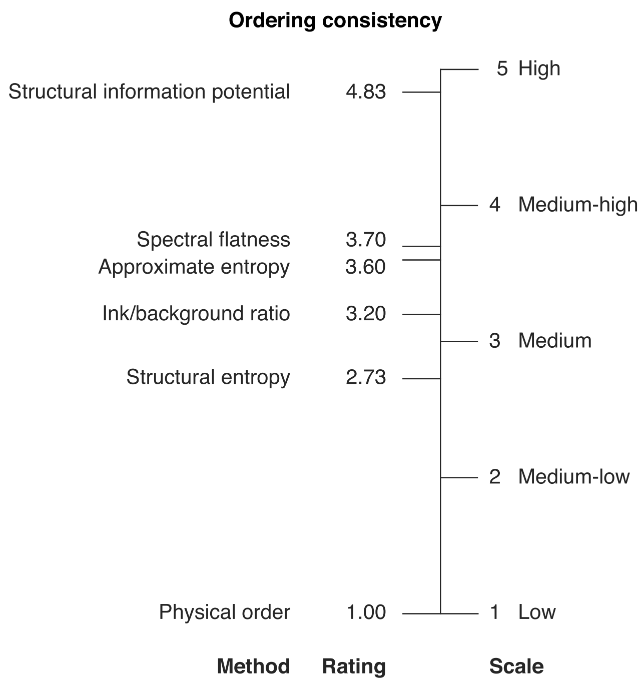

FIGURE 4. Results of the experiment on the perceptual rating of the computational ordering methods of [1, Fig. 19], as well as the physical order of pages in the document. The ratings are aggregated over 10 participants using the majority judgement method and a five-level rating scale.

5) In the above article [1, p. 13130, Section V-B]," add after §1 of col. 2, the following paragraphs and figure. The purpose of this addition is to provide an analytical explanation

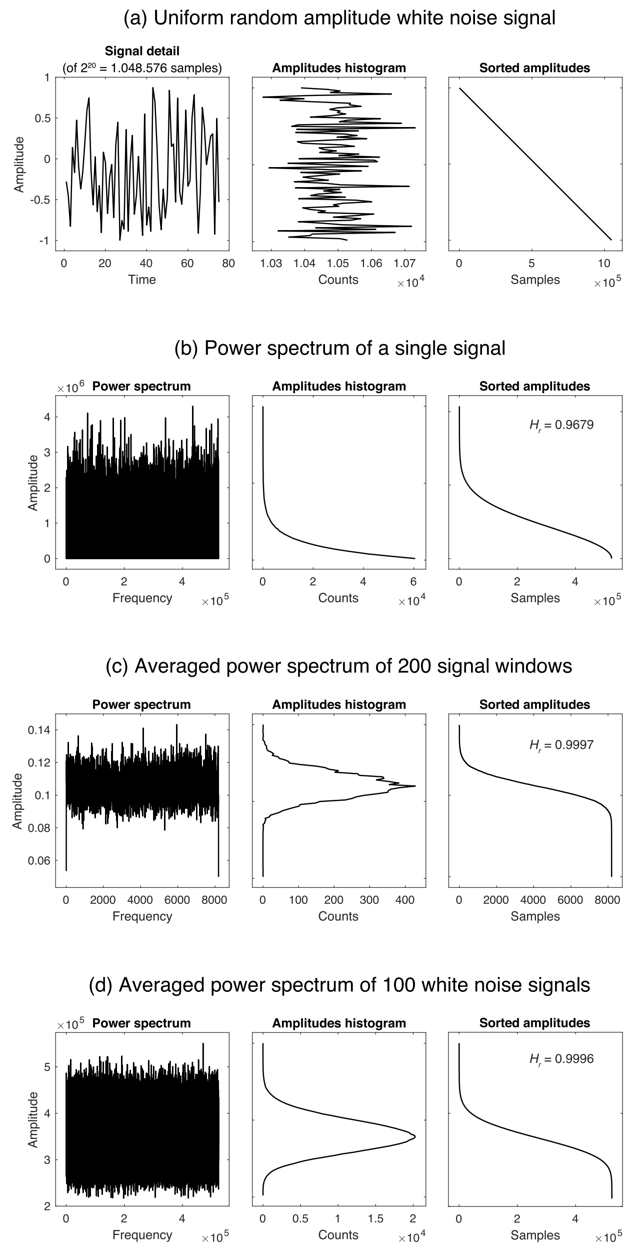

FIGURE 5. Analysis of the power spectra of uniform random amplitude white noise. (a) Left: Signal detail. Center: Histogram of the signal amplitudes. The amplitude values are in the [0, 2] range, are aggregated in 100 bins, and are displayed on the y-axis; the counts per bin are on the x-axis. Right: Amplitudes sorted in descending order. The plot is linear, as expected from a uniform random distribution. (b) The power spectrum of the signal is shown in (a). Note that the amplitude distribution is not flat, as is expected from white noise. (c) Welch's power spectral density estimate of the signal is shown in (a), in which the signal is divided into 200 overlapping windows of equal length, filtered with a Hamming window [6], [7, pp. 415–417]. Note the noisy Gaussian distribution of the amplitudes and the flattening of their sorted plot, which leads to an increase of the Shannon relative entropy, H_r, with respect to that of the power spectrum of (b). (d) Averaged power spectrum of 100 white noise signals.

of the randomness phenomena affecting the measurement of SIP.

In fact, the power spectrum of white noise with uniform random amplitude is theoretically flat, i.e. equal at all frequencies, and thus has maximal entropy [7, p. 385]. However, for a finite signal, the power spectrum distribution has a power law shape (Fig. 5 a, b). As in the case of run lengths, this is due to the unequal number of cycles per frequency available to compute the power spectrum. Various methods have been developed to overcome this limitation, such as averaging the power spectra of multiple random signals or averaging overlapping windows of the same signal (Fig. 5 c, d) [7, pp. 371–644], [8]. The distribution of the averaged amplitudes of the power spectra will now converge to a Gaussian, as per the central limit theorem. Increasing signal length also increases the distribution variance, with the power spectrum becoming uniform in the limit and its entropy maximal, as expected. The flattening of the power spectrum also characterizes the SIP method, since it comprises an averaging of the power spectra of all orientations of an image.

In conclusion, uniform random-amplitude white noise patterns of finite extent naturally exhibit substantial homogeneity, while their measured entropy is further increased through spectral averaging as part of the SIP method.

6) In the above article [1, p. 13132, col. 1, Section V-G]," replace §1 with the following paragraph. The purpose of this correction is to extend the discussion of SIP in relation to psychology.

The field of perceptual organization [9] studies from a psychological perspective many aspects related to SIP, among them structure [10], informativeness [11], entropy [12], uncertainty [13], redundancy [14], complexity [15], randomness [16], Gestalt Prägnanz [17], scene statistics [18], illusions [19], attractivity [20], and aesthetics [21]. Configurations are, however, difficult psychophysical research stimuli; this is more so the case for real data, such as documents, where semantics, aesthetics, and user experience (among many other factors) play an important role in their evaluation. As the pattern ordering resulting from the SIP is primarily intended to have humans as end users, it is nevertheless desirable to investigate in future work the relevance of modifying the SIP method to account for human pattern perception.

REFERENCES

[1] V. Atanasiu, "The structural information potential and its application to document triage," *IEEE Access*, vol. 10, pp. 13103–13138, 2022.

[2] Wikipedia. (Feb. 15, 2022). *Pentagon Papers*. [Online]. Available: https://en.wikipedia.org/wiki/Pentagon_Papers

[3] D. Ellsberg, *Secrets: A Memoir of Vietnam and the Pentagon Papers*. New York, NY, USA: Penguin Books, 2003.

[4] B. Gallagher, J. Harlan, and J. Scott. (2021). *'We're Going to Publish' An Oral History of the Pentagon Papers*. New York, NY, USA: The New York Times. [Online]. Available: https://www.nytimes.com/interactive/2021/06/09/us/pentagon-papers-oral-history.html

[5] M. Balinski and R. Laraki, *Majority Judgment: Measuring, Ranking, and Electing*. Cambridge, MA, USA: MIT Press, 2010.

[6] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, no. 2, pp. 70–76, Jun. 1967.

[7] S. M. Alessio, *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*. Cham, Switzerland: Springer, 2016.

[8] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.
[9] J. Wagemans, *The Oxford Handbook of Perceptual Organization*. New York, NY, USA: Oxford Univ. Press, 2015.
[10] E. Leeuwenberg and P. A. van der Helm, *Strutural Information Theory: The Simplicity of Visual Form*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
[11] J. Lauwereyns, *Brain and the Gaze: On the Active Boundaries of Vision*. Cambridge, MA, USA: MIT Press, 2004.
[12] R. D. Luce, ''Whatever happened to information theory in psychology?'' *Rev. Gen. Psychol.*, vol. 7, no. 2, pp. 183–188, Jun. 2003.
[13] W. R. Garner, *Uncertainty and Structure as Psychological Concepts*. New York, NY, USA: Wiley, 1962.
[14] F. Attneave, ''Some informational aspects of visual perception,'' *Psychol. Rev.*, vol. 61, no. 3, pp. 183–193, May 1954.
[15] F. Attneave and M. D. Arnoult, ''The quantitative study of shape and pattern perception,'' *Psychol. Bull.*, vol. 53, no. 6, pp. 452–471, Nov. 1956.
[16] R. Falk and C. Konold, ''Making sense of randomness: Implicit encoding as a basis for judgment,'' *Psychol. Rev.*, vol. 104, no. 2, pp. 301–318, Apr. 1997.
[17] M. Wertheimer, ''Untersuchungen zur Lehre von der Gestalt. II,'' *Psychologische Forschung*, vol. 4, no. 1, pp. 301–350, Jan. 1923.
[18] A. Desolneux, L. Moisan, and J.-M. Morel, *From Gestalt Theory to Image Analysis: A Probabilistic Approach*. New York, NY, USA: Springer, 2008.
[19] A. G. Shapiro and D. Todorovic, *The Oxford Compendium of Visual Illusions*. New York, NY, USA: Oxford Univ. Press, 2017.
[20] W. Prinz, ''Quantitative versuche über die Prägnanz von punktmustern,'' *Psychologische Forschung*, vol. 29, no. 4, pp. 297–359, Dec. 1966.
[21] R. Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley, CA, USA: Univ. California Press, 1974.