Towards the Automatic Evaluation of Visual Balance for Graphic Design Posters

Daniel Lopes, João Correia, Penousal Machado
University of Coimbra, CISUC, DEI and LASI, University of Minho
{dfl,jncor,machado}@dei.uc.pt

Abstract

Being able to evaluate aesthetics automatically is one of the fundamental needs for creating robust and autonomous computational creativity systems. In Graphic Design (GD), many aesthetic features might need to be considered simultaneously to properly evaluate GD artefacts, e.g. their visual relation to the concept of the work, legibility, innovation degree and the personal taste of the target public. Another relevant feature is the balance of the elements in the composition. This paper presents and tests an approach for evaluating the page balance of GD posters. Furthermore, it compares the evaluation computed by the developed method with the evaluation made manually by graphic designers and other creative practitioners. The results suggest the presented method can reasonably emulate the opinion of the human evaluators concerning the page balance of the presented posters. Moreover, for the presented setup, the results indicate a possible correlation between page balance and visual pleasantness, i.e. between the former and the personal taste of the human evaluators.

Introduction

More and more, Computational Creativity (CC) techniques have been explored to approach Graphic Design (GD) challenges, e.g. to speed up the GD creative process or aiding the exploration of innovative visual solutions.

To create CC systems that are capable of generating helpful GD solutions, one must first be able to create objective metrics to describe the quality of the expected designs. Nevertheless, due to the subjectivity of GD aesthetics, creating capable metrics for evaluating GD is still an open problem.

GD evaluation metrics may focus, for example, on the concept of the work, the legibility of the contents, innovation degree and other even more subjective features such as the personal taste of the target public. Another relevant feature is visual balance, which often relates to the visual weight of the items in the composition, on each side of a given axis.

Building on top of existing work (Harrington et al. 2004; Lok, Feiner, and Ngai 2004), this paper presents and tests a practical method to evaluate the page balance of GD posters. To do that, the brightness and position of each pixel in a given poster are considered to calculate a centre of mass (CM). The closer it is to a reference axis, the better the evaluation. Different axes and combinations of axes were tested. The evaluation values for each axis are weighed and summed up to calculate an overall evaluation value.

120 GD posters created by different graphic designers and gathered from a variety of sources, e.g. *typographic-posters.com*, *posters.calarts.edu* or websites from GD studios, were evaluated both automatically, using our method, and manually, by means of a user survey made with graphic designers and CC practitioners working on GD.

Subsequently, the automatically and manually obtained evaluation values were compared to assess whether or not the presented method could reasonably modulate human perception of page balance, at least, according to the opinion of the respondents of the conducted survey. Furthermore, studying the hypothesis of a correlation between page balance and visual pleasantness, the respondents were also asked how much they liked the respective posters.

The results suggest the proposed method could reasonably emulate the opinion of the respondents concerning the page balance of the presented posters. Moreover, the results indicate a possible correlation between page balance and visual pleasantness, at least, for the current experimental setup.

Related Work

As suggested before, this paper aims to contribute by introducing and testing a practical method to evaluate the visual balance of GD posters (2D), aiding the creation of CC systems for the generation of GD artefacts, such as posters. Hence, this section reviews existing work on the generation of 2D page layouts, especially focusing on page balance and CC systems for generating multipurpose GD layouts, i.e. which can be helpful in different GD briefings.

The automatic generation of multipurpose page layouts for GD has been done by numerous authors using different techniques. Constraint-based approaches are often used for displaying and aligning items on pages, e.g. using grid systems (Feiner 1988; Ferreira and others 2019; Cleveland 2010) or predefined templates (Jacobs et al. 2004). However, such systems are often unable to evaluate the generated results, so visual quality is usually controlled by humans or by restrictive hard-coded constraints, excluding such approaches from the CC domain.

Interactive Evolutionary Computation (IEC) has also been endorsed to generate GD layouts (Klein 2016; Kitamura and Kanoh 2011; Önduygu 2010). The shortcoming of IEC is the human users must still evaluate the generated results/candidates to drive the generation process.

One can also identify hybrid approaches in which the system automatically evolves layouts by fully filling in pages with a given number of text boxes, and the users are only asked to evaluate which they like the most (Rebelo et al. 2018). Nevertheless, such an approach to puzzle items into pages does not fit more generic contexts.

Geigel and Loui (2003) explored a more generic approach by constraining the layouts according to a number of hard-coded aesthetic metrics. Visual balance was automatically controlled by assessing page symmetry. However, such a metric can be reductive for defining visual balance, e.g. visual weight approaches (Harrington et al. 2004; Lok, Feiner, and Ngai 2004) can assess symmetry along with many other visual balance circumstances.

Automatic evolutionary computation (AEC) is one of the techniques that can benefit from automatic visual evaluation metrics, such as the one presented in this paper. AEC has demonstrated its potential to find solutions to complex problems (Stanley and Miikkulainen 2002), including in computational art contexts (Machado and Cardoso 2002) and some particular GD tasks, such as the generation of modular typography (Martins et al. 2016).

On the GD posters domain, there has been work using computer vision to automatically retrieve insights about whether the public is more or less interested in a given candidate poster and assigning fitness accordingly, i.e. the more a person looked at a poster, the better the fitness (Rebelo et al. 2017). While the latter work could only generate background variations on a single poster layout, there has also been work towards multipurpose systems. For example, to approximate existing layouts using different page items (Lopes, Correia, and Machado 2022).

Lastly, there has been work using Machine Learning (ML) techniques to learn features of existing layouts so one can generate new ones accordingly. For example, by learning how to drive the public's attention to given zones of the layout, detecting alignment or understanding hierarchical features (O'Donovan, Agarwala, and Hertzmann 2014), or generating layouts according to the semantic information of the page items (Zheng et al. 2019).

Specifically on the page balance domain, and besides symmetry (Geigel and Loui 2003), Harrington et al. (2004) proposed to assess visual balance by calculating the CM and measuring its distance to the centre of the page or, alternatively, assessing the difference between the visual weight on the left and on the right side of the page. The shortcoming is the authors' balance calculation is based on the average brightness of each page item, i.e. the metric is not well-fitted to be applied to raster images, especially if they're too complex. Furthermore, the average brightness of an item might be misleading if the object is visually heavier on one of its sides. Lok, Feiner, and Ngai (2004) used edge detection to assess the size, position and brightness of the page items and therefore calculated weight maps. A shortcoming of the latter approach is the page items were assumed to be uniformly weighted. As stated by the authors, a pixel-based approach might better reflect the way humans evaluate layouts.

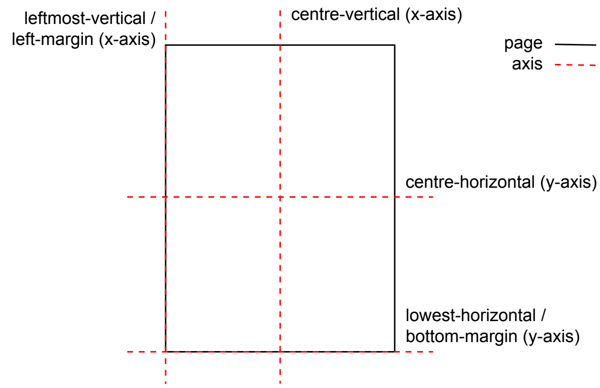

Figure 1: Page axes tested.

Approach

Inspired by the work of Harrington et al. (2004) and Lok, Feiner, and Ngai (2004), this paper presents a pixel-based method to evaluate visual balance in 2-dimensional GD artefacts, especially focusing on posters.

The proposed method, implemented in *JavaScript*, takes as input one PNG image of any size and ratio. However, to improve performance, we automatically resize the input image to 400 pixels wide. Height is set proportionally.

The CM of the given image is calculated considering the brightness and the position of each pixel. First, the default CM is assigned to a vector in the centre of the page (centred vertically and horizontally). Then, each pixel is assigned a weight equal to its inverted normalised brightness, i.e. 0 standing for brighter values and 1 standing for darker ones, so darker pixels were considered visually heavier, as often white pixels stand for emptiness/white page. This value is then squared, emphasising the differences between lighter and darker pixels. A vector referring to the position of the given pixel is then multiplied by its respective weight. Lastly, the resulting vector is added to the default CM vector. This way, darker pixels will more strongly attract the CM in their direction to the detriment of lighter ones.

After assessing the CM of the image, the distance to one or more axes can be calculated to estimate balance. In the following experiments, we evaluated posters considering the centre-vertical, centre-horizontal, leftmost-vertical (left-margin) and the lowest-horizontal (bottom-margin) axes (see Figure 1), either alone or mixing two axes together. The left-margin axis was tested in detriment to the right-margin one since the gathered posters communicate using left-to-right writing. To mix axes together, their distances to the CM were weighed and then summed up.

The full code can be downloaded from *GitHub* at *github.com/danifslopes/Visual-Balance-Evaluation*.

Experimental Setup and Analysis

As mentioned before, the system was tested by evaluating a dataset of 120 posters gathered from various sources. First, to more easily study the impact of the item's distribution on the page, we started by designing a set of 30 posters com-

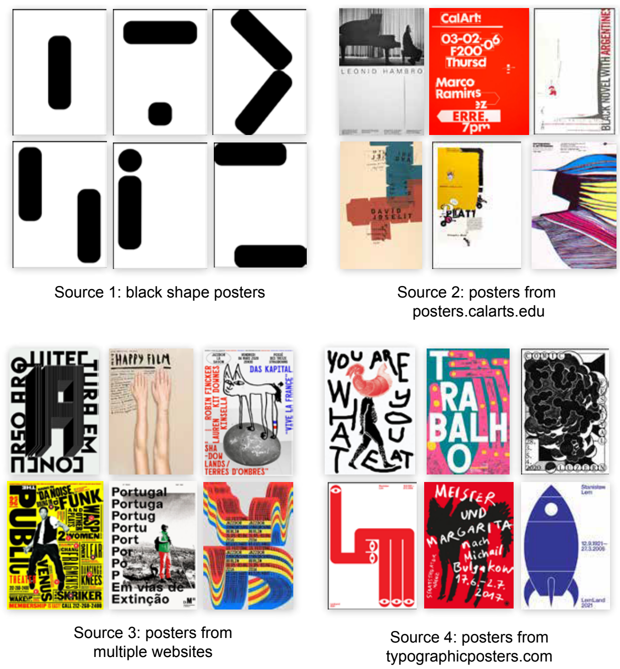

Figure 2: Examples of evaluated posters, grouped by type of source. (1) posters designed on purpose by our research team; (2) posters from the *posters.calarts.edu* archive; (3) posters from well-known GD studios, gathered from multiple websites; (4) posters from *typographicposters.com*.

posed of a maximum of two black geometric shapes over white background, positioned in varied dispositions. We refer to these as black shape posters. Another 30 posters were gathered online from *posters.calarts.edu*, a poster archive containing diverse posters quality-wise. The third set of 30 posters was gathered from several websites from well-known GD studios. Lastly, 30 posters were gathered from *typographicposters.com*, an online archive in which graphic designers worldwide upload their poster designs. Typically, this archive includes work by experienced designers as the users need to be approved by the administrators of the archive. Refer to Figure 2 for examples of posters from each of the aforementioned sources.

As the goal of proposed method is to aid the creation of AEC systems for GD, a group of 25 graphic designers and CC practitioners working on GD were asked to evaluate the posters concerning their visual balance and visual pleasantness. We refer to this as manual evaluation.

More specifically, all the respondents had a GD background, except for two who did not. Even so, these were working on CC for GD purposes. Also, all had Portuguese nationality except for one Brazilian living in Portugal for 2 years already. Except for 6 of them, all the respondents worked or studied at the University of Coimbra at the time the survey was conducted.

The respondents were asked, from 0 to 10, (i) *"How visually balanced do you think the posters are? Please, ignore whether you like them or not. Do not consider whether or not you like the colours, typefaces or other graphics"* and (ii) *"How aesthetically pleasing the posters seem to you, regardless of why?"*. Each respondent evaluated 24 posters, 6 of each type.

Secondly, the values resulting from manual evaluation were compared with the values resulting from automatic evaluation. To do that, the method was run over the 120 posters using different parameters. We tested evaluating the posters considering the following axes and combinations of axes: (i) Centre-Vertical alone (CV); (ii) Centre-Horizontal alone (CH); (iii) Left-Margin alone (L); (iv) Bottom-Margin alone (B); (v) Centre-Vertical & Centre-Horizontal (C+C); (vi) Left-Margin & Bottom-Margin (L+B); (vii) Centre-Vertical or Left-Margin whichever closer to CM & Centre-Horizontal (C/L+C); (viii) Centre-Vertical & Centre-Horizontal or Bottom-Margin whichever closer to CM (C+C/B); (ix) Centre-Vertical or Left-Margin whichever closer to CM & Centre-Horizontal or Bottom-Margin, also, whichever closer to CM (C/L+C/B).

For getting a unique balance value, whenever two axes were combined, their normalised distances to CM were weighted and summed up. For the vertical and horizontal axes respectively, we tested the following weights: [0.5, 0.5], [0.25, 0.75] and [0.75, 0.25].

To compare manual and automatic evaluation, we averaged the manual evaluation values and then calculated the (i) average distance (the closer to 0 the better) and (ii) the cosine similarity (the closer to 1 the better) between the manual and the automatic evaluation values. For a better comparison between metrics, the average distance was inverted, turning it into an average similarity value instead[1], i.e. the closer the value is to 1, the better.

Moreover, we tested similarities by (i) considering all 120 posters together, (ii) excluding the black shape posters from the main set of posters (i.e. using 90 posters), and (iii) using the 30 black shape posters only.

Comparing Manual and Automatic Visual Balance

The average similarity and the cosine similarity values between manual and automatic evaluation can be found in Table 1. Results for different combinations of axes and weights are presented.

By comparing the two different metrics, one can notice these return slightly different values, i.e. ranging from 0.590 to 0.807 for the average similarity and 0.900 to 0.977 for the cosine similarity metric. Even so, in one case or another, the maximum similarity values can be considered relatively high, i.e. relatively close to 1 (0.807 for average similarity and 0.977 for cosine similarity), suggesting the presented method could reasonably match the values from manual evaluation.

One can also notice that using the left-margin and bottom-margin axes to calculate the automatic balance tends to decrease similarity values when compared to the centre-vertical and centre-horizontal axes. This suggests that, in

[1]The average of the absolute differences between the manual and automatic evaluation of each poster, inverted. So the closer the value is to 1, the higher the similarity.

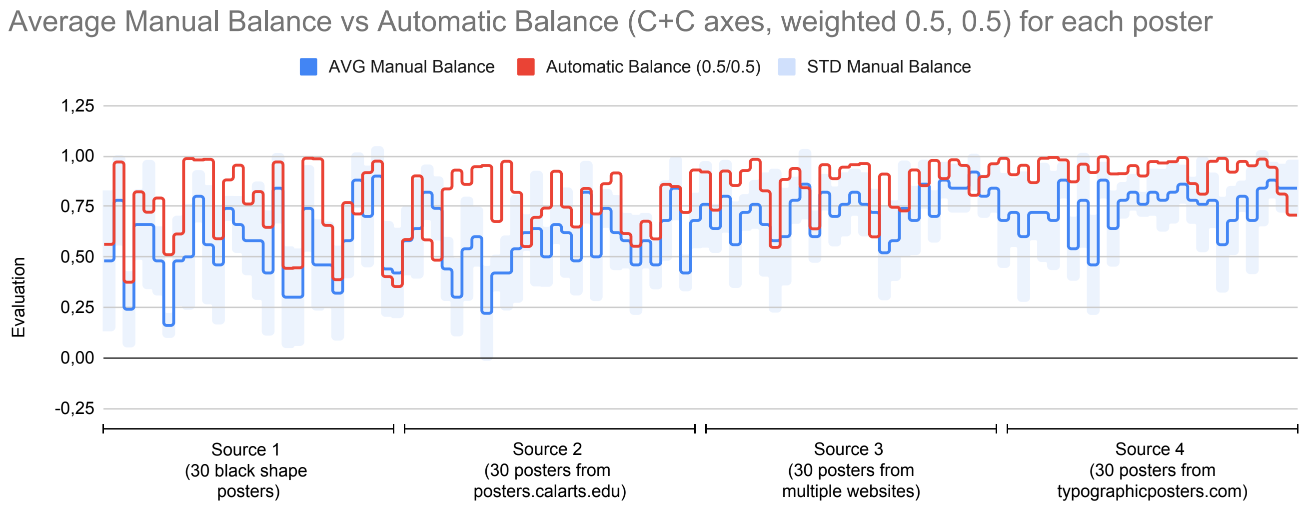

Figure 3: Automatic balance alongside Average (AVG) and Standard Deviation (STD) of manual balance, for each of the 120 posters, ordered by poster type. Automatic evaluation performed considering the C+C axes weighted equally (0.5, 0.5).

general terms, the centre-vertical and centre-horizontal axes might be better fitted to calculate visual balance. At least, when comparing to the results of the conducted survey.

Both metrics indicated the best parameterisation is using the centre-vertical and centre-horizontal axes (C+C) weighted equally (0.5, 0.5). More specifically, for such parameters, average similarity equals 0.8067 and cosine similarity equals 0.9768. Figure 3 presents a visualisation of the values obtained using the C+C axes weighted equally, alongside the average manual balance for each one of the 120 posters.

Even so, other parameterisations using one or two centre axes resulted in similar results. For instance, holding average similarities ranging from 0.788 to 0.805. This suggests that, often, the respondents evaluated better the posters in which the CM is closer to the centre of the page, e.g. either the contents align with the centre axes (one or both) or the visual weight is distributed symmetrically, relatively to the centre of the page.

Isolating the black shape posters That can also be inferred from Figure 4, which showcases all the 30 black shape posters, ordered by the respective manual evaluation values for visual balance. One can notice that most of the first 15 posters (on the top), worst evaluated, have their visual weight distributed on a single side of each centre axis (vertical and horizontal), i.e. often positioned on the corners. For instance, refer to the posters 1-8 and 10-14 of Figure 4. On the other hand, most of the posters on the bottom, better evaluated, have their visual weight better distributed on both sides of at least one centre axis. For instance, refer to posters 16 and 19-30 of Figure 4.

Best and worst-evaluated posters The assumptions above can likewise be deduced by looking at the best and worst-evaluated posters, either (a) considering all 120 posters or (b) excluding the black shape ones (see Figure 5). One can argue the visual weight of the best-evaluated posters

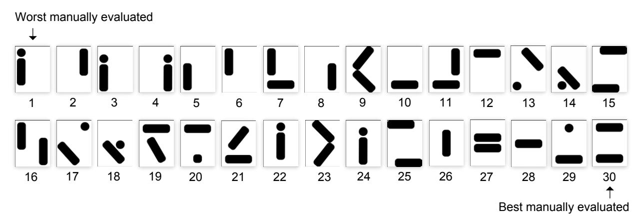

Figure 4: Black shape posters, ordered by average manual balance.

tends to be either on the centre of the page or organised in a relatively symmetrical way, relative to one or both centre axes. That can be more or less prominent in posters 2-5 of Figure 5.a and posters 2-5 of Figure 5.b. However, such an assumption can be less evident concerning poster number 1 (from either group (a) or (b) of Figure 5). Even so, one might find that the image contained in the poster reasonably balances the visual weight of the typographical elements.

On the contrary, the worst evaluated posters tend to have most of their visual weight on one side of one, or both, centre axes. For instance, all the worst posters showcased (posters 6-10 from either group (a) or (b) of Figure 5) have their contents displayed either on the left or the right of the page. Also, except for posters b.9 and b.10, all the worst-evaluated posters tend to have their contents vertically aligned to the top of the page. Poster b.9 has its contents vertically aligned to the centre of the page, and b.10 to the bottom of the page.

Differences between automatic and manual evaluation Further insights can be drawn by visualising the differences (inverted similarity values) between automatic and manual evaluation values, for each of the 120 posters. As presented in Figure 6, the average difference was 0,193 (standing for a 0,807 similarity), indicating that automatic evaluation is, on average, around 80% aligned with the opinion of the respondents of the conducted survey.

Table 1: Average similarity between manual and automatic evaluation values. Maximum value highlighted in bold.

Average Similarity

Axis	Weights (vertical-axis, horizontal-axis) n/a	0.5, 0.5	0.25, 0.75	0.75, 0.25
CV	0,7730			
CH	0,7861			
L	0,5903			
B	0,5942			
C+C		**0,8067**	0,8028	0,7966
L+B		0,5956	0,5955	0,5939
C/L+C		0,8009	0,7993	0,7889
C+C/B		0,8050	0,8013	0,7958
C/L+C/B		0,7992	0,7979	0,7880

Cosine Similarity

Axis	Weights (vertical-axis, horizontal-axis) n/a	0.5, 0.5	0.25, 0.75	0.75, 0.25
CV	0,9708			
CH	0,9620			
L	0,8999			
B	0,9179			
C+C		**0,9768**	0,9721	0,9762
L+B		0,9319	0,9310	0,9209
C/L+C		0,9753	0,9710	0,9750
C+C/B		0,9765	0,9723	0,9759
C/L+C/B		0,9751	0,9712	0,9747

Figure 5: Best and worst posters concerning visual balance according to the average manual evaluation. On the top, considering all 120 posters. On the bottom, excluding the black shape posters. The respective manual and automatic evaluation values are indicated under the respective posters.

Figure 7 showcases the posters concerning the lowest and highest absolute difference (distance) between manual and automatic balance. Although it might be difficult to draw conclusive insights from the analysis of the showcased posters, we describe some possible yet speculative reasons for the higher distances obtained (which refer to the posters at the bottom in Figure 7), i.e. why the automatic method did not match the opinion of the respondents for these cases.

The first reason relates to a known shortcoming of the presented method. In most of the reviewed posters, page items are presented in darker tones compared to the respective backgrounds. Thus, assigning heavier visual weight to darker zones usually works reasonably to asses visual balance, as previously mentioned. However, in cases in which the background is darker than the respective contents, the calculation of the CM should be (but is not so far) inverted for the CM to still be attracted in the direction of the page items, now concerning lighter tones, and not otherwise. This shortcoming can be identified in posters 6 and 7 of Figure 7, in which the system considered the CM to be almost centred despite the contents being placed on the left of the page.

Furthermore, we highlight that such an inversion should only happen, as aforementioned, if the background is darker than its contents, not whenever there are more dark pixels than light ones (or whenever the average brightness is low). For example, if a poster is almost fully filled in with black objects over white background, most pixels will be dark. However, one may perceive the black blobs (in the majority) as the items, and the white space (in minority) as the background, so the calculus shall not be inverted in this case. Hence, a more sophisticated method must be implemented in future work to distinguish between background and foreground (whenever possible) and, therefore, decide whether the calculus shall be inverted.

Although the background-detection issue might explain the high distance between manual and automatic evaluation for posters 6 and 7 (and eventually 8) of Figure 7, such an argument cannot fit, for example, posters 9 and 10.

Therefore, we believe that the apparent visual movement of the composition might also have some degree of influence on the human perception of visual balance. As an example, poster 9 is composed of 2 shapes positioned in a way the calculated CM is close to the centre of the page, leading to a high automatic balance value (0.986). However, visually, the shapes seem to be pilled in an unstable position (if making an analogy to the physical world and considering the ground to be the bottom of the poster), which might have led the respondents to evaluate this poster with a low balance value (averaging 0.460). Even so, this is a speculative assumption.

A third reason concerns poster 10 of Figure 7 and relates back to the set of axes and respective weights used to calculate balance. Although the automatic method produced identical balance values for poster 10 and its symmetrical version (i.e. 0,987), manual evaluation resulted in considerably different values. For instance, 0,460 and 0,74 for poster 10 and

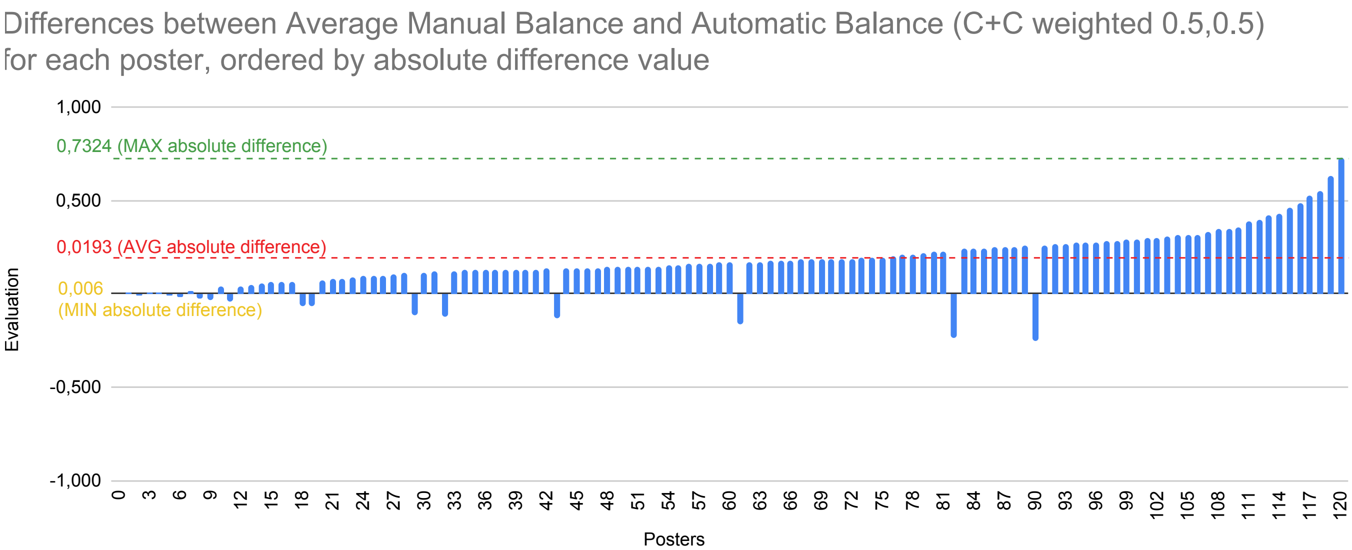

Figure 6: Differences between average manual balance and automatic balance (C+C, weighted 0.5, 0.5), for each poster, ordered by absolute difference value.

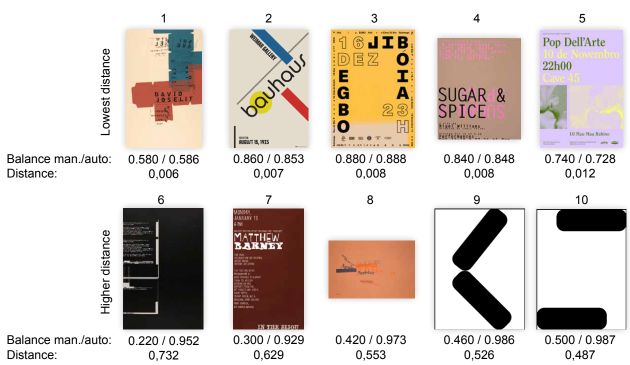

Figure 7: Posters concerning the lowest and highest absolute difference (distance) between manual and automatic balance.

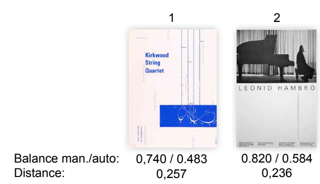

Figure 8: Two posters whose automatic evaluation is lower than manual evaluation and whose absolute difference between metrics was above average.

its symmetrical version, respectively (see posters 15 and 25 of Figure 4, respectively, for the mentioned poster and its symmetrical). A possible reason for that is the respondents considered the diagonal axis that crosses the poster from the top-left to the bottom-right corners to be more balanced than the one that goes from the top-right to the bottom-left corners. Nonetheless, all the aforementioned assumptions must require further study.

From Figure 6, one can also conclude that the automatic method often evaluates the posters optimistically compared to the average manual evaluations. For instance, 106 posters were evaluated automatically over manual evaluation, and only 14 were evaluated automatically under manual evaluation. Among the 14 under-evaluated posters, only two referred to distance values above average (see Figure 8). For instance, 0.257 and 0.236 (0.064 and 0.043 above average, respectively). Even so, it can be worth analysing such under-evaluated posters.

Although the system considered the CM was not fully centred, the respondents considered the composition of the two posters of Figure 8 to be relatively balanced, i.e. 0.740 and 0.820 manual balance, opposing to 0.483 and 0.584 automatic evaluation, respectively. A further user survey must be conducted to properly assess the reason why. Nevertheless, looking at poster 2 of Figure 8, questions concerning page division can be raised. For example, one can see that poster 2 is visually divided into two main parts — one on the top containing an image, and one on the bottom containing some typography and a wide empty zone. In future research, we shall consider whether evident divisions of the page can impact the perception of visual balance.

Conclusions In sum, considering the presented experiments and analysis, for the present experimental setup, the proposed method for automatically evaluating visual balance demonstrated to match, on around 80%, the evaluation made by the human designers and CC practitioners that participated in the conducted survey.

Therefore, we believe the present method can already be worth testing to perform fitness assignment on AEC systems. Furthermore, although the C+C axes seemed to approximate better the opinion of the respondents, we believe other pa-

Table 2: Average similarity (AVG sim.) and cosine similarity (Cosine sim.) between balance and pleasantness evaluation values, gathered through the user survey.

Poster sets	AVG sim.	Co sine sim.
All posters	0,892	0,976
Excluding black shape posters	0,923	0,991
Black shape posters only	0,800	0,934

rameterisations shall be worth trying, e.g. using C/L+C/B axes combination for allowing a wider range of possible layouts to show up.

Also, as assessing balance alone may be reductive to evaluate GD artefacts, we suggest complementing the proposed balance metric with some other metrics, such as for assessing legibility, the innovation degree of the designs or their relation to a given concept.

The conducted analysis also suggested that, presumably, other visual features can sometimes bias people's perception of visual balance. Thus, for creating more robust visual balance methods, it might be worth studying the impact of apparent movement, or how much a visual division of the page can influence balance perception.

Comparing Visual Balance and Visual Pleasantness

As mentioned before, besides visual balance, the respondents of the conducted user survey were asked, from 0 to 10, how visually pleasing they considered the posters were. The goal was to gather some insights about whether or not visual balance relates to visual pleasantness in some way. Figure 9 presents the values for balance and pleasantness gathered through the user survey, for each of the 120 posters.

Besides analysing Figure 9, to compare balance and pleasantness evaluation, the average similarity and cosine similarity were calculated. Respectively, the similarity values consisted of 0,892 and 0,976. Such relatively high values (higher than the similarity between manual and automatic evaluation) suggest there might be a considerable degree of correlation between balance and pleasantness.

For trying to gather further insights, besides the whole 120 posters, we calculated similarity values by removing the black shape posters (using 90 posters), as well as using the latter alone (30 posters only). Such values can be consulted in Table 2.

The resulting values indicate that excluding the black shape posters leads to higher similarity values (0,923 and 0,991 average and cosine similarity values, respectively). Similarly, the black shape posters alone led to lower similarity values (0,800 and 0,934 average and cosine similarity values, respectively). This might indicate the respondents found the black shape posters less visually pleasing compared to the remaining posters, regardless of their balance. Therefore, some visual features that are not as present in the black shape posters as in the remaining ones might have influenced the perception of pleasantness.

As mentioned before, we believe it might be worthy to further study what visual features influence the most the perception of visual pleasantness. Judging from the results hereby presented, one can argue that visual balance might contribute to some extent to the perception of visual pleasantness. However, further testing must be necessary to prove such an assumption.

Conclusion

One of the requirements for developing reliable and independent computational creativity systems is the ability to autonomously evaluate aesthetics. However, finding objective metrics to do it effectively is still an open problem.

In Graphic Design (GD), to properly evaluate artefacts, it may be necessary to take into account a number of factors, e.g. the visual relationship of the given artefact to its concept, how legible it is, how innovative it is, and even whether it fits the personal taste of the target audience. Furthermore, visual balance is often a relevant feature to take into consideration.

In this paper, we have presented and tested a practical method for evaluating the page balance of GD posters. To do that, a centre of mass was calculated by taking into account the brightness and location of each pixel in a given poster. The evaluation of the balance is improved by the proximity of this centre of mass to some predefined vertical and/or horizontal axes. An overall evaluation value is then determined by weighing and adding the obtained evaluation values for each axis. Different axes and combinations of axes were tested during the experiments.

To test the presented approach, a set of 120 GD posters created by different authors and gathered from various sources were evaluated manually by graphic designers and CC practitioners, by means of a user survey. The results of the survey were then compared to the ones performed by the developed method, by crossing insights from mathematical metrics and the analysis of visual features of the posters.

In addition, the respondents were asked how visually appealing they found the posters to be, hoping to retrieve some insights into a supposed correlation between page balance and visual pleasantness.

The results suggested the proposed method could match, at around 80%, the balance evaluation made by the respondents. Moreover, the results indicated a possible correlation between page balance and visual pleasantness, at least, for the current experimental setup.

Future work must focus on testing the proposed approach as a fitness assignment method for an automatic evolutionary system. As assessing balance alone may be reductive to evaluate GD posters, we must complement it with other metrics, such as for assessing legibility or innovation degree. Lastly, we must further study how different personal backgrounds or additional visual features, such as apparent movement, hue and saturation, may impact the calculation of visual balance.

Acknowledgments

This research was funded by national funds through the FCT—Foundation for Science and Technology, I.P., within the scope of the project CISUC—UID/CEC/00326/2020 and

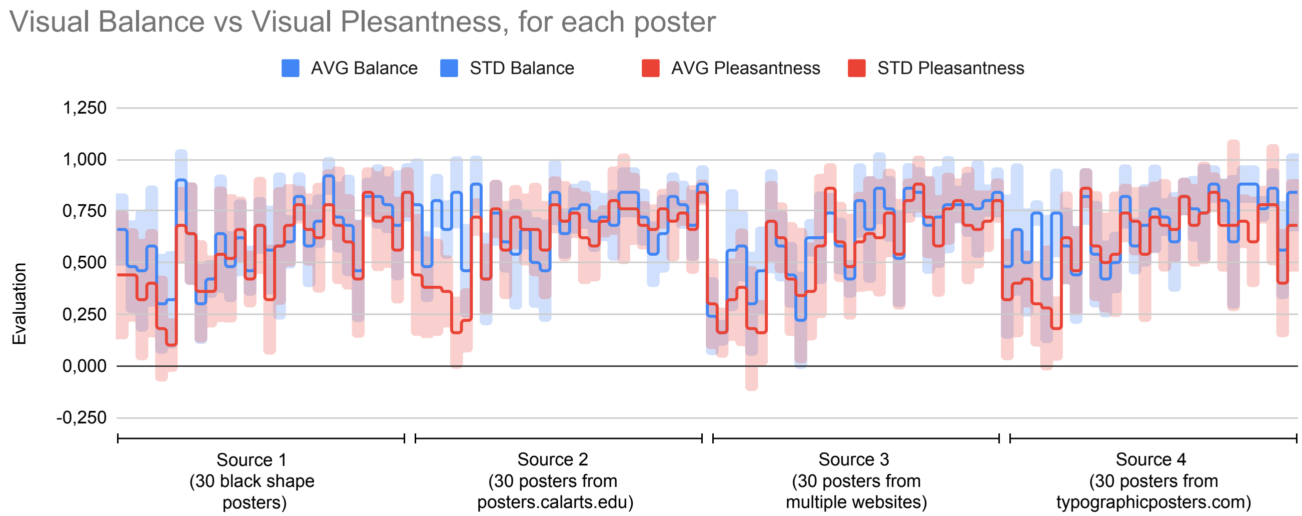

Figure 9: Average (AVG) and Standard Deviation (STD) for visual balance and visual pleasantness, for each of the 120 posters, ordered by poster type. The values were gathered through a user survey.

by the European Social Fund, through the Regional Operational Program Centro 2020 and is partially supported by *Fundação para a Ciência e Tecnologia*, under the grant SFRH/BD/143553/2019.

References

Cleveland, P. 2010. Style based automated graphic layouts. *Design Studies* 31(1):3 – 25.

Feiner, S. 1988. A grid-based approach to automating display layout. In *Proceedings Graphics Interface*, volume 88, 192–197.

Ferreira, D., et al. 2019. Design editorial algorítmico. Master's thesis, Universidade de Coimbra.

Geigel, J., and Loui, A. 2003. Using genetic algorithms for album page layouts. *IEEE MultiMedia* 10(4):16–27.

Harrington, S. J.; Naveda, J. F.; Jones, R. P.; Roetling, P.; and Thakkar, N. 2004. Aesthetic measures for automated document layout. In *Proceedings of the 2004 ACM Symposium on Document Engineering*, DocEng '04, 109–111. NY, USA: Association for Computing Machinery.

Jacobs, C.; Li, W.; Schrier, E.; Bargeron, D.; and Salesin, D. 2004. Adaptive document layout. *Communications of the ACM* 47(8):60–66.

Kitamura, S., and Kanoh, H. 2011. Developing support system for making posters with interactive evolutionary computation. In *2011 Fourth International Symposium on Computational Intelligence and Design*, volume 1, 48–51. Piscataway, USA: IEEE.

Klein, D. 2016. Evolving layout: next generation layout tool. Last accessed 17 Dec 2018.

Lok, S.; Feiner, S.; and Ngai, G. 2004. Evaluation of visual balance for automated layout. In *Proceedings of the 9th International Conference on Intelligent User Interfaces*, IUI '04, 101–108. New York, NY, USA: Association for Computing Machinery.

Lopes, D.; Correia, J.; and Machado, P. 2022. Evodesigner: Towards aiding creativity in graphic design. In Martins, T.; Rodríguez-Fernández, N.; and Rebelo, S. M., eds., *Artificial Intelligence in Music, Sound, Art and Design*, 162–178. Cham: Springer International Publishing.

Machado, P., and Cardoso, A. 2002. All the truth about NEvAr. *Applied Intelligence, Special Issue on Creative Systems* 16(2):101–119.

Martins, T.; Correia, J.; Costa, E.; and Machado, P. 2016. Evotype: from shapes to glyphs. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, 261–268. ACM.

Önduygu, D. C. 2010. *Graphagos: evolutionary algorithm as a model for the creative process and as a tool to create graphic design products*. Ph.D. Dissertation, Sabanci University.

O'Donovan, P.; Agarwala, A.; and Hertzmann, A. 2014. Learning layouts for single-pagegraphic designs. *IEEE Transactions on Visualization and Computer Graphics* 20(8):1200–1213.

Rebelo, S.; Martins, P.; Bicker, J.; and Machado, P. 2017. Using computer vision techniques for moving poster design. In *6.º Conferência Internacional Ergotrip Design*. Aveiro, Portugal: UA Editora.

Rebelo, S.; Fonseca, C. M.; Bicker, J.; and Machado, P. 2018. Experiments in the development of typographical posters. In *6th Conference on Computation, Communication, Aesthetics and X*. Porto, Portugal: Uni. of Porto.

Stanley, K. O., and Miikkulainen, R. 2002. Evolving neural networks through augmenting topologies. *Evolutionary Computation* 10(2):99–127.

Zheng, X.; Qiao, X.; Cao, Y.; and Lau, R. W. H. 2019. Content-aware generative modeling of graphic design layouts. *ACM Trans. Graph.* 38(4).