Towards the Automatic Customisation of Editable Graphics

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

Abstract

More and more, graphic designers need to deliver faster and cheaper. To help speed up the creative process, we propose a computational approach for automatically stylising graphics so that these can relate to given semantic concepts. More specifically, by automatically selecting transformation and styling properties that better relate to a list of provided keywords, according to *ConcepNet*. For testing the presented approach, 2D posters referring to different concepts were stylised using our method and evaluated through a user survey. Although the system can be further improved, the results suggest the viability of the approach to aid the automatic styling of concept-related graphics, such as posters.

Introduction

In creative fields such as Graphic Design (GD), finding disruptive visual solutions that attract people's attention is of the utmost importance, either to create communicative or artistic design artefacts, e.g. so the designs stand out over other posters on the streets or other book covers on store shelves. However, as the urgency for more effective designs grows and the GD area is increasingly democratised, graphic designers need to deliver faster and cheaper, which often leads to the adoption of trendy solutions and precludes the exploration of innovative visual solutions.

This paper proposes a computational approach for automatically styling graphics so that these can relate to given semantic concepts. More specifically, we propose using *ConceptNet* (Liu and Singh 2004) to assess the semantic similarity between given keywords (that must be set by the user to describe a given concept) and labels (not changeable by users) given beforehand to a set of mutation methods and their respective parameters. The resulting similarity values are then used to calculate the probability of each mutation method and respective parameters to be used for styling a given graphical item, e.g. given the keyword *sky* a given item may be more likely filled in *blue* in detriment to other colours. In other words, given *sky*, if the method *fillColour* is picked, the most probable parameter would be *blue*.

Although our approach might be generic enough to be used in different creative contexts, in this paper it was tested in the generation of posters, i.e. for transforming and styling a number of text boxes, images, and geometric shapes within 2D pages. Furthermore, to create a tool that could be easily integrated into a designer's workflow, the presented approach was implemented as an extension for *Adobe InDesign* — a broadly used desktop-publishing software for GD. Thus, designers can alternate between manually and automatically editing items without leaving *Adobe InDesign*.

As seeking highly subjective concepts could lead to highly subjective results, the tests focused mostly on keywords that might often be related to visual stereotypes (at least, in Western culture), e.g. in Western culture, one might find the concept *love* often related to the colour *red*. Furthermore, other parameters were explored to understand whether we could lead the system to either focus on a principal idea or exploit a more diverse range of related ideas. Lastly, the generated posters were evaluated by means of a user survey. The latter suggests that, although the system can be further improved, for the tested experimental conditions, the present approach can be viable to automatically style graphics so that these visually relate to given keywords.

Related Work

To aid the generation of concept-related graphics, there have been theoretical studies about the relationship between words and visual features, e.g. the relation between emotions and features such as colours, shapes, directions, curve sizes, or edge types (Collier 1996; Cavanaugh, MacInnis, and Weiss 2016; Rodrigues, Cardoso, and Machado 2019).

Concerning practical applications, relating words to colours seems to be often explored. One can pinpoint datasets relating words to colours (Heer and Stone 2012) and even automatic systems to do so, e.g. by extracting colour pallets (O'Donovan, Agarwala, and Hertzmann 2011) or by recolouring bitmap images automatically (Lin et al. 2013).

Besides colour-focused approaches, one can pinpoint ones, for example, for editing texture-like vector images according to one given adjective (Heath and Ventura 2016), or to generate vector sketches that illustrate keywords by using Evolutionary Computation (EC) (Cunha et al. 2020) or Machine Learning (ML) techniques (Ha and Eck 2017) to interpolate existing sketches or even by composing existing 3D models (Coyne and Sproat 2001). Zhao, Cao, and Lau (2018) presented a system to learn the most attention-catching zones in posters according to a given set of themes, such as *minimalist* or *romantic*. This system can be helpful, e.g. in the generation of layouts according to those themes.

The most popular approaches for translating semantic concepts into graphics nowadays must be those using transformers and stable diffusion to generate realistic bitmap images from given text prompts (Ramesh et al. 2022; Radford et al. 2021). A shortcoming is the enormous data and computation requirements needed to implement such systems. Furthermore, bitmap images are often not suited for creating GD artefacts, since designers often need to edit or update information or create different variations of the generated artefacts in different formats and sizes. Moreover, such systems often generate pastiche results (Toivonen and Gross 2015) (i.e. variations of existing styles). In that sense, although such systems can be helpful, e.g. to generate illustrative images, we believe these are not yet well suited for co-creatively styling GD artefacts such as posters.

Approach

For aiding the design of creative artefacts, we propose a computational approach to automatically select mutation methods and respective parameters (referred to as visual assets) that can relate to a set of keywords (i.e. given a concept) defined by users.

To demonstrate the proposed approach, we tested it for styling GD posters. We implemented it as an extension for *Adobe InDesign* so manual and automatic editions/mutations can be done interactively, fostering the collaboration between the system and human designers.

To use this system, human designers must start by inserting desired items (i.e. text boxes, geometric shapes, or images) into *InDesign* pages. Then, manual editions may or may not be done. Whenever desired, the user may define a set of keywords descriptive of the concept of the work and click a button in the extension to automatically style a number of items according to the defined keywords, e.g. by more likely selecting given colours in detriment to others or by more likely rotating items other than skewing them.

That is accomplished by picking mutation methods and respective parameters according to their semantic relation to the concepts the user wants to seek. *ConceptNet* (Liu and Singh 2004) is used to calculate the relatedness (semantic similarity) between a set of keywords (defined by the user) and a set of labels assigned in advance to the available mutation methods and respective parameters. As an example, one can retrieve such similarity values using *ConceptNet* queries such as: *https://api.conceptnet.io/relatedness?node1=/c/en/beach&node2=/c/en/yellow*. In this first iteration, the labels were defined by our research team. In future developments, we aim to gather these by means of a user survey. 357 labels were used (a complete list can be found in the supplementary materials).

The proposed procedure goes as follows: (i) Using the user interface, one must type an intended set of keywords, e.g. *sun* and *beach*; (ii) The system will call *ConceptNet* to calculate similarity values between each label and each keyword; the returned values range from -1 to 1, standing to highly dissimilar to highly similar, respectively; however, the values are truncated to range from 0 to 1, i.e. 0 is considered the minimum possible similarity; (iii) Each label will be assigned with the respective maximum similarity value, e.g. if the label *yellow* is considered 0.135 similar to *sun* and 0.056 similar to *beach*, then *yellow* will be assigned with the value 0.135; also, for each edition method, each parameter will be assigned with the maximum value of its labels, and each edition method with the maximum value of its parameters (maximum values are always picked so the system acknowledges the most important assets, whether these relate to a keyword or another); (iv) The assigned values are then used as the probability of each method/parameter being automatically picked to edit a given page item.

We implemented methods to edit the following properties: the shape of the surrounding box of the items, their size, position and order (z-position), flipping and blending modes, opacity, fill colour and tint, stroke colour and tint, stroke weight, rotation, and the item's shearing angle. Concerning text boxes, also text size, text colour, justification, vertical alignment, letter spacing, and line height were considered.

For the experiments hereby presented, the edition process went as follows: (i) For each page item, all the available methods were iterated and each one could or not be picked to edit the respective item, according to their assigned probability; (ii) Whenever a method was picked, one of the available parameters was automatically chosen using a roulette approach, according to the respective probabilities to run; (iii) Each selected method and its respective parameters were then used to mutate the given item.

Furthermore, to understand whether we could lead the system to either focus on a principal idea (i.e. to the most related assets) or explore a more diverse range of related ideas (i.e. a wider range of assets), we created two variables to remap the assigned probabilities exponentially. That is, we set up an exponent to remap method probabilities and another to remap parameter probabilities. We refer to these as probability exponents. In practical terms, by increasing these exponents, high-related assets will be even more likely to be picked than less-related ones.

Experimental Setup

Experiments were conducted to determine whether meaningful visual solutions could emerge using the proposed approach. The experimental setup was started by manually creating a base poster in *InDesign* (see Figure 1). This poster should be as visually neutral as possible while containing a reasonable amount of items, so reasonable visual changes could occur and thus stronger conclusions could be drawn. Therefore, 7 black and white items were used: 1 central circle, 4 rectangles dividing the poster into 4 parts, and 2 text boxes, one at the top and another at the bottom. There was no particular reason to choose such a layout rather than seek a medium-complexity, balanced, and neutral composition. The text within text boxes is changed along with experiments to relate to the respective keywords.

For each experiment, 30 posters were automatically generated from the base poster. Each poster was submitted to the edition process once (see Section *Approach*). Among the

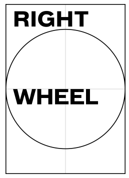

Figure 1: Base poster (created manually) from which the automatically-stylised posters were generated.

experiments, only the input concepts varied. As argued before, to avoid too much subjectivity, the tests focused mostly on concepts that we considered to have more or less obvious visual translations (at least in Western culture).

Moreover, to test the system in tasks of different difficulty, such concepts were selected (empirically) in a way that included keywords that were in two different categories: keywords that were equal to labels (similarity value equal to 1) and keywords that were not (similarity value smaller than 1). More specifically, we have selected 16 different keywords to build 10 different concept inputs: *"green", "big, love", "fear, dark", "sun, beach", "grape, vines", "small, shark", "big, shark", "baby, shark", "right, circle" and "colourful, balloons"* (an additional concept, removed due to space constraints, can be found in supplementary materials). 10 of those keywords were equal to labels of some method or parameter, i.e. *green, big, love, fear, dark, sun, small, right, circle,* and *colourful*.

For the aforementioned reason, one could expect the concepts *"green", "big, love", "fear, dark" and "right, circle"* to be the easiest to solve, *"baby shark" and "grape, vines"* the most difficult ones, and the remaining to be of medium difficulty. However, that might not always be true, as some mutation methods have a bigger perception impact than others, and the concepts refer to different levels of abstraction.

Although our approach can be used to pick any mutation method, to ease the interpretation of the experimental results, we have mainly focused on one of the most easily observable — colour. To do that, the methods to mutate fill colour and text colour have been set with a 100% chance of being picked regardless of the concept. The respective parameters (e.g. what colour to fill items with) and the remaining methods kept their automatically assigned probabilities. By experimentation, the method's exponent and the parameters' exponent were set to 3 and 5, respectively.

Experimental Results

Figure 2 showcases results for concepts in which keywords were used in the labelling of methods or parameters. Figure 3 showcases results for concepts in which one keyword (out of two) was included in labelling. Figure 4 showcases results for concepts in which none of the keywords was included in labelling.

The generated posters were evaluated by means of an online user survey (refer to supplementary materials). 34 people participated. To understand whether age, formation, or cultural background could influence the answers, such personal information was asked in the first instance. Also, visual disabilities were misguided. The respondents comprised an age group from 21 to 50, with 17 people being 21–22 years old. All were from Western countries (1 from Italy and 1 from the UK, both living in Portugal; the remaining were Portuguese), so one can assume all or most of them were familiar with similar cultural stereotypes. This was a relevant issue as, for different cultures, colours and symbols can have different meanings Since some questions referred to personal approaches to designing posters, the survey was specially directed to people with GD background. Even so, one of them had not. All questions were open-answer.

After gathering personal information, the respondents were asked what colours, shapes or other visual assets they would choose to design a poster for each of the keywords mentioned in Section *Approach*. The goal would be to later compare the assets chosen by the respondents with the assets automatically chosen using our approach. As previously indicated, the analysis of the answers focused mainly on colour. As the goal of our system is not to create composed or figurative images, such solutions were disregarded.

Colour-wise, the system matched the most mentioned choices of the respondents for each concept, i.e. green for "*green*", red and pink for "*big, love*", black and dark tones for "*fear, dark*", yellow and blue for "*sun, beach*", blue and grey for "*small shark*", "*big, shark*" and "*baby, shark*", colourful for "*colourful, balloons*", and purple and green for "*grape vines*". An exception is made for "*right circle*", for which the respondents more often chose the green colour and the system picked a variety of different colours. However, since the latter concept led to less consensual responses compared to the remaining, it could be predicted this concept would be more difficult to reach (same for an additional concept in supplementary materials: *right wheel*).

However, due to the set properties' exponent, the system often gave preference to one colour for each concept, whereas the respondents would sometimes use two colours in similar amounts, e.g. purple and green for *grape vines*.

Other than colour, the system sometimes matched the respondent's choices, at least for concepts included in labelling, e.g. it matched the circles in the *"right, circle"* posters, and the small items in the *"small, shark"* posters. However, further testing must be done to understand the system's effectiveness for methods that were not forced to always run, as *fillColour* and *textColour* were. Also, features that could not be achieved by the system were mentioned, e.g. repetition of items. In that regard, future work might comprise the development of additional mutation methods.

In question two, the respondents were asked whether the generated posters from Figure 2, 3 and 4 would better represent the respective concepts compared to the base poster, i.e. whether the system could improve the base poster conceptually. Also, they were asked to ignore legibility issues.

For 8 out of the 10 concepts (11 if counting with *"right, wheel"*), the majority of the respondents considered the posters improved. However, as could be predicted, for *right circle*, the respondents considered the posters did not represent the concept well. Suggestions included using circles rather than ellipses and adding more green.

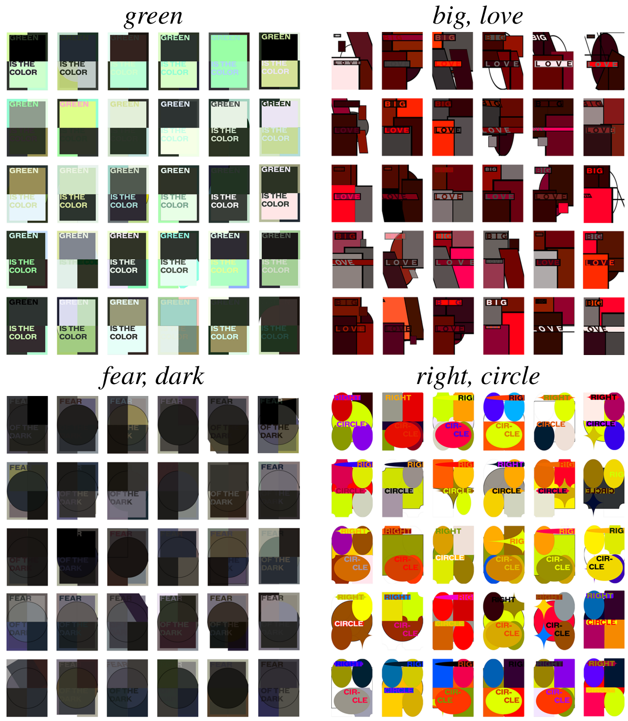

Figure 2: Posters generated from keywords used for labelling mutation methods or parameters.

Figure 3: Posters generated from one keyword used for labelling mutation methods or parameters, and one that was not.

Another non-consensual concept was *"small, shark"*, as 13 respondents considered there was no improvement, 11 that it did, and 9 that it did in part. The answers suggested that *small* was well represented. However, the rectangles did not seem to represent well the idea of *shark*, neither for *small*, *big* nor *baby, shark*. As the rectangles were already there in the base poster, it is likely that the system simply did not choose to make shape changes. Nonetheless, it would be positive if the system had transformed those rectangles into triangles, as frequently chosen by the respondents in the first question. For *"big, shark"*, adding red was also suggested. However, the big, no-fill rectangles with thick black strokes seemed to recall a more aggressive look, which pleased most respondents. Further suggestions went through decreasing dark tones for the *"green"* concept, using circles for *"grape, vines"* and *"colourful, balloons"*, and lighter colours for *"big, love"*.

In the third question, the respondents were asked whether they believed the generated aesthetics could already be used to create final GD posters, taking into consideration that the text contents could be changed. The answers suggest the respondents could see such potential in at least some of the posters. The most mentioned ones (by 20, 19, and 28 people, respectively) were the posters for *"fear, dark"*, *"colourful balloons"* and *"sun, beach"*. Furthermore, their comments suggest that illustrating given concepts can have significantly different objectives compared to designing a poster for communicating information about something related to that same concept. That can be noticed from the fact that many respondents seemed to be comfortable abdicating from the circle shapes (many times referred to in previous questions) if they were to design a poster for disseminating events related to grape vines or colourful balloons. Also, some respondents commented that the generated posters could be used as a design base, and small changes could then be manually made to create the final posters. A last noteworthy comment refers to the fact that some posters ended up being different from what the respondent chose in the first question. However, the generated solution still worked out well, suggesting the potential for the system to aid unexpected creative choices too.

The last question meant to understand whether a similar approach could aid creative (or not creative) choices in other areas rather than design. Furthermore, we asked which areas would those be. 25 respondents indicated that such approaches could be useful to hasten the beginning or during creative tasks, as long as they are used in a co-creative way, i.e. humans must be able to fine-tune the generated results. Besides some respondents still referred to GD applications, such as creating book covers or graphic identities, other suggestions were made. For instance, styling websites, picking adequate visual encoding and colour pallets for information visualisation, creating icons and glyphs, styling and creating combinations of clothing for fashion design, or picking musical features.

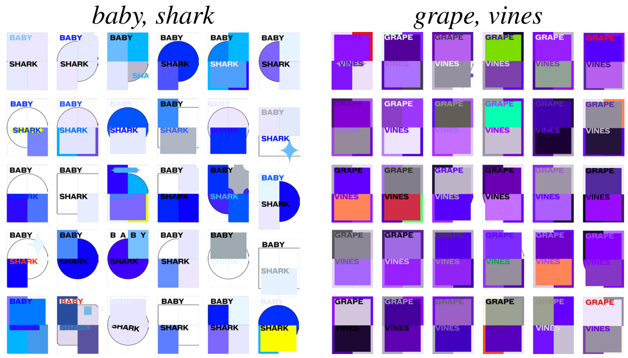

Figure 4: Posters generated from keywords that were not used for labelling mutation methods or parameters

Conclusion

More and more, designers need to be more competitive, delivering faster and cheaper, which often culminates in the adoption of trendy solutions. Thus, to speed up the exploration of innovative solutions, we presented an approach for automatically stylising graphics according to given semantic concepts, i.e. a set of keywords. To do that, *ConceptNet* (Liu and Singh 2004) was used to assess the relatedness value between given conceptual keywords and the labels of a set of mutation methods and respective parameters. The bigger the relatedness of a method, the more likely it is to be used.

We tested our approach by styling posters, i.e. transforming and styling a number of text boxes, images and geometric shapes within 2D pages. The presented approach was implemented as an extension for *Adobe InDesign* — a broadly used desktop-publishing software for GD —, so designers could alternate between manually and automatically editing/mutating the posters within the same software.

The results drawn from the user survey suggest the presented approach can be viable to aid the exploration of concept-related solutions, at least in poster design and, especially, taking into consideration that GD concepts can be transmitted in abstract ways. However, human collaboration might still be essential to curate and fine-tune the results and transform the generated ideas into final GD applications.

Future work can comprise (i) assigning weights (importance levels) to the keywords; (ii) resetting the labels according to the insights gathered through another user survey; and (iii) associating assets, e.g. if both the shape *wave* and the colour *blue* are picked because of the same keyword *sea*, then *wave* would more likely run over blue items (or neutral ones) and vice-versa.

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 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

Cavanaugh, L. A.; MacInnis, D. J.; and Weiss, A. M. 2016. Perceptual dimensions differentiate emotions. *Cognition and Emotion* 30(8):1430–1445. PMID: 26308182.

Collier, G. L. 1996. Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion* 20(1):1–32.

Coyne, B., and Sproat, R. 2001. Wordseye: An automatic text-to-scene conversion system. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, 487–496. New York, NY, USA: Association for Computing Machinery.

Cunha, J. M.; Lourenço, N.; Martins, P.; and Machado, P. 2020. Visual blending for concept representation: A case study on emoji generation. *New Generation Computing* 38(4):739–771.

Ha, D., and Eck, D. 2017. A neural representation of sketch drawings.

Heath, D., and Ventura, D. 2016. Creating images by learning image semantics using vector space models. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Heer, J., and Stone, M. 2012. Color naming models for color selection, image editing and palette design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, 1007–1016. New York, NY, USA: Association for Computing Machinery.

Lin, S.; Ritchie, D.; Fisher, M.; and Hanrahan, P. 2013. Probabilistic color-by-numbers: Suggesting pattern colorizations using factor graphs. *ACM Trans. Graph.* 32(4).

Liu, H., and Singh, P. 2004. Conceptnet — a practical commonsense reasoning tool-kit. *BT Technology Journal* 22(4):211–226.

O'Donovan, P.; Agarwala, A.; and Hertzmann, A. 2011. Color compatibility from large datasets. In *ACM SIGGRAPH 2011 Papers*, SIGGRAPH '11. New York, NY, USA: Association for Computing Machinery.

Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; Krueger, G.; and Sutskever, I. 2021. Learning transferable visual models from natural language supervision.

Ramesh, A.; Dhariwal, P.; Nichol, A.; Chu, C.; and Chen, M. 2022. Hierarchical text-conditional image generation with clip latents.

Rodrigues, A.; Cardoso, F.; and Machado, P. 2019. A dynamic approach for the generation of perceptual associations. In *International Conference on Computational Creativity (ICCC)*. address: organizations.

Toivonen, H., and Gross, O. 2015. Data mining and machine learning in computational creativity. *Wiley Int. Rev. Data Min. and Knowl. Disc.* 5(6):265–275.

Zhao, N.; Cao, Y.; and Lau, R. W. H. 2018. What characterizes personalities of graphic designs? *ACM Trans. Graph.* 37(4).