Human-Centered Visual Interfaces for Image Retrieval: An Exploratory Study

Diogo Cruz, Diogo Cabral, Pedro F. Campos

Human-Centered Visual Interfaces for Image Retrieval: An Exploratory Study

Diogo Cruz[1] and Diogo Cabral[2] and Pedro F. Campos[1]

[1] ITI/LARSyS, University of Madeira, Portugal
[2] ITI/LARSyS, IST, University of Lisbon, Portugal
2030115@student.uma.pt; diogo.n.cabral@tecnico.ulisboa.pt; pedro.campos@iti.larsys.pt

Abstract. The widespread use of digital cameras in mobile phones has led to an increase in the number of pictures captured and shared. However, the development of interfaces for visualizing image collections has not matched that growth. Most search methods are based on text descriptions and retrieve a large number of results on thumbnail 2D grids, which can be hard to analyze. Therefore, it is crucial to couple image retrieval with purposely designed interfaces. This paper covers the study of four different interfaces for the visualization of collections of images, including a regular 2D grid, a variable-size 2D grid, a pile of images, and a spiral. These interfaces were evaluated in a user test involving nine participants performing search tasks. We found that both grids exhibit higher usability and lower task times than the Pile and Spiral. The Variable Size Grid had lower usability scores than the Regular Grid, but it showed higher-quality task results and was preferred by the participants in this study.

Keywords: Visual Search, Image Visualization, Retrieval Interfaces.

1 Introduction

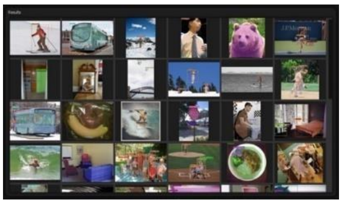
With the rise of digital cameras and the growing number of devices that use them, there has been a considerable increase in the number of images being captured. Thomee and Lew [17] claimed that finding optimal user interfaces for queries and results is one of the grand challenges for image retrieval. However, the study of novel ways to visualize collections of images has been quite limited. Most of the recent research has focused on improving search methods [1, 6, 19]. The works that proposed novel visualizations for image browsing and search [2, 8, 20] have failed to measure how the different visualizations affect human search and content analysis. While search methods can filter the images to retrieve, it is simply not enough in many cases. Therefore, it is crucial to couple image retrieval methods with purposely designed image visualization interfaces.

2D grids with image thumbnails are the predominant interface in image search engines all these years but problematic for visualizing large datasets. André et al. [2] analyzed image browsing and concluded that interfaces should: support exploration;

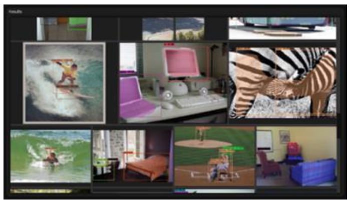
exploit thumbnail aesthetics; allow for search refinement and present its history; allow saving images for later.

Similarity functions can create a visualization of image sets [3, 10, 11]. Placing visually similar images close to each other has been proved to improve the time required to locate specific images [12]. Heesch and Rüger [8] present a visualization where the most relevant images are larger and near the center, and the least relevant are smaller and placed around. The authors also discuss another idea where the images are placed around a central point, forming a spiral. Additional research [7, 13, 18] has developed this idea further. More recent works [9, 14, 15, 21] have studied the possibility of visualizing images in virtual reality environments.

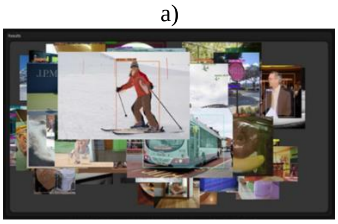
This work compares four different user interfaces for visualizing image collections: Regular Grid, Variable Size Grid, Pile and Spiral. Considering two different search tasks, we conducted a user test measuring the usability, the duration of each task, the quality of task results (Precision, Recall, and F-Measure), and users' preferences for each visualization.



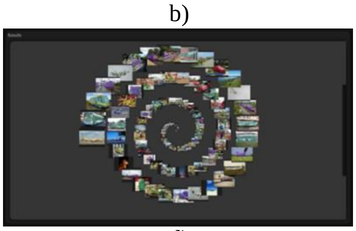
a)



b)



c)



d)

Fig. 1. Image visualizations: a) Regular Grid; b) Variable Size Grid; c) Pile; and d) Spiral

2 Visual Interfaces for Image Search

We developed four different user interfaces for visualizing collections of images as part of a visual search application: Regular Grid, Variable Size Grid, Pile and Spiral.

In a typical use case, the results are sorted by the confidence obtained from the visual search method. The sorting takes advantage of each interface to make sense according to their characteristics and make the most relevant results more prominent. Hovering over an image will show a tooltip with relevant information and a scaled-up version (also implemented in the other visualizations). The Regular Grid (Fig. 1a) is a standard scrollable 2D grid with image thumbnails and works as a baseline.

Then we developed a grid with thumbnails of different sizes (Fig. 1b), similar to the Regular Grid except that the rows towards the top have fewer and larger images.

The number of images per row increases gradually along the y-axis, and the images are gradually smaller. The images are sorted by relevance in descending order, so the top results are more relevant and larger.

We also developed a visualization with overlap, called Pile (Fig. 1c). The images are placed in a randomized (x, y) position, and their depth is determined by their relevance, being the most relevant images displayed on top. The user can drag the images around freely. It is also possible to bring the images to the foreground and scale with the scroll wheel, zooming in/out and clear some space. Finally, we developed a visualization where the images are placed in a spiral (Fig. 1d), exploiting the aesthetics aspect of thumbnails. The images are placed in a spiral path according to a function based on an Archimedes spiral. The most relevant images are displayed towards the outside of the spiral as it is easier to look at multiple images simultaneously. The spiral can be zoomed in/out with the scroll wheel, always over the center. It is possible to drag any image towards it before zooming in, so any image can be looked at in a large size.

3 Evaluation

We conducted a user test following a within-subject design with seven males and two females (average age: M = 27.78 years; SD = 13.28). Users were asked to perform four sets of two tasks each, one task being a broad search and the other a specific search. A broad search consists of searching for objects that belong to a group: buses, balls, dogs, and apples. A specific search consists of searching for an object with specific characteristics: red buses, footballs, beige dogs, and green apples. The order of the visualizations used for each task was counterbalanced to minimize any possible learning and dataset biases. All tasks were properly timed.

Table 1. SUS scores and time in seconds for each visualization/task

	SUS		Broad Search Task	Specific Search Task	
	M (SD)	Mdn	M (SD)	M (SD)	Mdn
Regular Grid	94.69 (9.40)	98.75	87.93 (33.86)	30.78 (8.58)	31.14
Variable Size Grid	87.50 (11.50)	91.25	92.07 (31.31)	49.41 (37.59)	36.16
Pile	48.55 (13.69)	62.50	251.17 (61.90)	134.46 (37.15)	141.87
Spiral	53.75 (26.22)	48.70	185.29 (60.09)	96.48 (46.02)	106.96

Four collections (transportation, sports, animal, and food) of 100 images each were used, one for each set of tasks. These images were obtained from the COCO dataset [5]. Each visualization and each image collection were used to perform one set of tasks. The broad search task included 20 correct images, while the specific search included five correct images that the participants should select. Participants were not told how many images there were in total nor how many were correct. However, before each task, they were told which image theme they would be browsing and what to search for. The images were presented in random order (not sorted). To accomplish each task, users could select the images that they considered as having target objects with a double click.

To compare each visualization, we measured its usability through the SUS questionnaire [4] (Table 1); the time taken by each participant to accomplish each task (Table 1); the quality of task results, through Precision, Recall and F measure [16] (Table 2); user preference, through a final questionnaire. Regarding users' preferences, the Variable Size Grid was the most preferred visualization (Mdn = 1), followed by the Regular Grid (Mdn = 2), Pile (Mdn = 3), and the Spiral (Mdn = 4).

Table 2. Effectiveness: Precision, Recall, and F Measure

	Broad Search Task						Specific Search Task					
	Precision		Recall		F Measure		Precision		Recall		F Measure	
	M (SD)	Mdn	M (SD)	Mdn	M (SD)	Mdn	M (SD)	Mdn	M (SD)	Mdn	M (SD)	Mdn
Regular Grid	0.99 (0.4)	1.00	0.88 (0.20)	1.00	0.92 (0.15)	1.00	1.00 (0.00)	1.00	0.87 (0.14)	0.80	0.93 (0.09)	0.89
Variable Size Grid	1.00 (0.00)	1.00	0.93 (0.07)	0.95	0.96 (0.04)	0.97	1.00 (0.00)	1.00	0.93 (0.10)	1.00	0.96 (0.06)	1.00
Pile	0.98 (0.03)	1.00	0.97 (0.04)	1.00	0.97 (0.02)	0.98	0.98 (0.06)	1.00	1.00 (0.00)	1.00	0.99 (0.03)	1.00
Spiral	0.98 (0.05)	1.00	0.89 (0.12)	0.95	0.93 (0.07)	0.94	1.00 (0.00)	1.00	0.89 (0.18)	1.00	0.93 (0.11)	1.00

4 Discussion and Conclusions

We presented four interfaces for visualizing image collections and compared them in a user test with nine participants. Our results show that both Grids present higher usability scores when compared with the non-traditional visualizations like the Pile and Spiral. Participants also performed both tasks, broader and specific search, fastest using both Grids. It is interesting to note that the Pile, sometimes used to show collections of images or algorithm search results, is not recommended for human analysis and search tasks. Regarding the quality of results, no significant differences were found between the different visualizations. However, it is relevant to note that the Variable Size Grid presents a higher Recall and F-Measure score when compared with the Regular Grid, particularly on specific search tasks. This result suggests that combining both visualizations in an adaptive interface could be an interesting approach for future work, which is reinforced by participants' preference, which has selected the Variable Size Grid as the preferable interface.

Acknowledgments

This research was funded by FCT/MCTES by LARSyS (UIDB/50009/2020) and by CEECINST/00122/2018 (IST-ID/300/2019).

References

1. Alexander, M., Gunasekaran, S.: A Survey on Image Retrieval Methods. (2014). 10.13140/2.1.2779.0723.
2. André, P., Cutrell, E., Tan, D., Smith, G.: Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage. In Proc. INTERACT '09. 340–353. Springer-Verlag, Berlin, Heidelberg (2009).
3. Basalaj, W.: Proximity Visualisation of Abstract Data. No. UCAM-CL-TR-509. University of Cambridge, Computer Laboratory (2001).
4. Brooke, J.: SUS: a quick and dirty usability scale. Usability evaluation in industry, 189. (1996)
5. COCO - Common Objects in Context, http://cocodataset.org/#home, last accessed 2020/08/3 (2020)
6. Dureja, A., Pahwa, P.: Image retrieval techniques: A survey. International Journal of Engineering & Technology. 7(2), pp. 215-219, (2017)
7. Goker, A., Butterworth, R., Macfarlane, A., Stumpf, S.: Presenting and visualizing results on an image retrieval user interface. In Proc. British HCI 2017 (2017).
8. Heesch D., Rüger S.: Three Interfaces for Content-Based Access to Image Collections. In Proc. CIVR 2004. LNCS, vol 3115. Springer, Berlin, Heidelberg (2004).
9. King, F., Jagadeesan, J., Bhagavatula, S., Shyn, P., Pieper, S., Kapur, T., Lasso, A., Fichtinger, G.: An Immersive Virtual Reality Environment for Diagnostic Imaging. Journal of Medical Robotics Research 1(01), p.1640003 (2016).
10. Nguyen, G., Worring, M.: Optimization of interactive visual-similarity-based search. ACM TOMM, 4(1), pp.1-23 (2008).
11. Rodden, K.: Fd, Cambridge. Evaluating Similarity-Based Visualisations as Interfaces for Image Browsing (2002)
12. Rodden, K., Basalaj, W., Sinclair, D., Wood, K.: Evaluating a visualisation of image similarity as a tool for image browsing. In Proc. InfoVis '99 (pp. 36-43). IEEE (1999).
13. Rueger, S.: Putting the User in the Loop: Visual Resource Discovery. In International Workshop on Adaptive Multimedia Retrieval (pp. 1-18). Springer, Berlin, Heidelberg. (2005).
14. Schaefer, G., Budnik, M., Krawczyk, B.: Immersive browsing in an image sphere. In Proc. IMCOM '17. ACM, New York, NY, USA, Article 26, pp. 1–4. (2017).
15. Schaefer G., Ruszala S. Hierarchical Image Database Navigation on a Hue Sphere. In Proc. ISVC 2006. LNCS, vol 4292. Springer, Berlin, Heidelberg. (2006)
16. Schütze, H., Manning, C.D. and Raghavan, P.: Introduction to information retrieval Cambridge University Press, Cambridge (2008).
17. Thomee, B., Lew, M.: Interactive search in image retrieval: a survey. International Journal of Multimedia Information Retrieval, 1(2), pp.71-86 (2012).
18. Torres, R., Silva, C., Medeiros, C., Rocha, H.: Visual structures for image browsing. In Proc. CIKM '03. ACM, New York, NY, USA, pp. 49–55 (2003).
19. Wang, H., Mohamad, D., Ismail, N.: Approaches, Challenges and Future Direction of Image Retrieval, arXiv preprint arXiv:1006.4568 (2010).
20. Zhang, L., Chen, L., Jing, F., Deng, K. and Ma, W.Y.: EnjoyPhoto: a vertical image search engine for enjoying high-quality photos. In Proc. ACM Multimedia 2006, ACM, New York, NY, USA, pp. 367-376. (2006).
21. Zheng, L., He, L., Yu, C.: Mobile Virtual Reality for Ophthalmic Image Display and Diagnosis. Journal of Mobile Technology in Medicine, 4(3), pp.35-38, (2015).