FISIR: A Flexible Framework for Interactive Search in Image Retrieval Systems

Sheila M. Pinto-Cáceres[1,2], Jurandy Almeida[1,3],
M. Cecília C. Baranauskas[1], and Ricardo da S. Torres[1]

[1] Institute of Computing, University of Campinas – UNICAMP, Campinas, SP, Brazil
{cecilia,rtorres}@ic.unicamp.br
[2] Institute of Science and Technology, Paulista University – UNIP, SP, Brazil
sheila.caceres@unip.br
[3] Institute of Science and Technology, Federal University of São Paulo – UNIFESP,
São José dos Campos, SP, Brazil
jurandy.almeida@unifesp.br

Abstract. This paper presents a flexible framework for interactive search in image retrieval systems. Our approach allows for the visual structure change in order to produce coherent layouts, which highlight the most relevant results according to user needs. This innovative framework is flexible in the sense it supports the dynamic creation of several hybrid visual designs, based on the combination of different visualization strategies. Results from a subjective evaluation demonstrate that the dynamic hybrid layouts created by the proposed framework provide the end-users with an effective user interface for intuitive browsing and searching experience.

Keywords: image retrieval, interactive search, result visualization, visual structure.

1 Introduction

Recent advances in technology have increased the availability of image data, creating a strong requirement for efficient systems to support image retrieval tasks. Making efficient use of image information requires that data be accessed in a user-friendly way. For this, it is important to provide users with a browsing tool to interactively search for (or query) an image in large collections, without having to look through many possible results at the same time, and therefore helping users to easily find images in which they are interested [1].

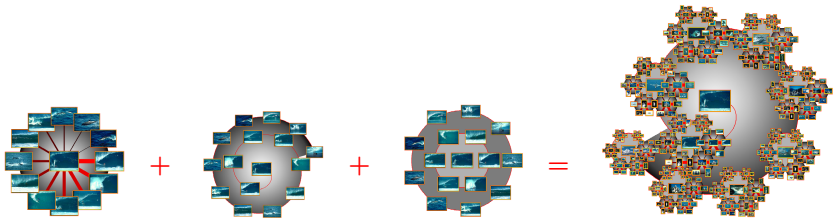
One of the main challenges of designing an interactive retrieval system is to minimize the fatigue and frustration that a user might experience when not obtaining useful information. Ideally, the layout of those systems should allow the inspection of large image collections in a coherent manner, by displaying appropriately the most relevant images according to user needs. To fulfill such requirements, a competitive layout must provide more than just images, rather, it should take advantage of different visualization strategies to improve and accelerate the comprehension of how images are correlated to each other.

A lot of research has been conducted on the proposal of browsing techniques for the interactive search on image collections [2]. In spite of all the advances,

X. He et al. (Eds.): MMM 2015, Part I, LNCS 8935, pp. 335–347, 2015.
© Springer International Publishing Switzerland 2015

many of those research approaches have displayed a set of images in a precomputed rigid structure, which cannot change (or dynamically be changed) to visually emphasize particular user interests. Usually, they support the user interaction just by exploring and showing different parts of a static structure.

In this paper, we propose a novel framework for searching and browsing in image retrieval systems. Our approach is totally flexible, supporting the dynamic creation of several hybrid and coherent layouts, based on the combination of different visualization strategies, as shown in Figure 1. In this way, this innovative framework allows the visual structure change according to user needs and, hence, it can be easily adapted to different search scenarios. Experiments were conducted both for evaluating the dynamic hybrid layouts created by the proposed framework and for comparing it with several static visualization techniques. Results from a rigorous empirical comparison with a subjective evaluation show that the hybrid visual designs produced by our framework are effective in supporting intuitive browsing and searching experience.



(a) Circle **(b)** Spiral **(c)** C. Rings **(d)** Our Framework

Fig. 1. An illustration of our framework. Users can dynamically create a hybrid layout (d) by combining different visual designs (a, b, and c). Our framework leads to a better searching experience as it is adaptive to users' needs.

2 Related Work

A comprehensive review of searching and browsing methods for image retrieval can be found in [2]. Some of the main ideas and results among the previously published methods are briefly discussed next.

Most of image retrieval systems have been designed to support a targeted search. In this kind of navigation, users provide a query pattern to the system and then receive a single ranked list containing the most similar results. The most popular approach for arranging the result set is to use list- or grid-based layouts, where images are disposed in a linear or grid manner according to their relevance to the query pattern [3]. Several techniques display the result set in a circular or elliptical form [5]. These radial methods try to centralize the user vision by setting the query at the center of the available space and then place similar images around it, allowing easy access and exploration of the result set [6]. Despite the simplicity of those methods, they fail on visually representing the intrinsic relationships among similar images.

As an alternative, extensive research has focused on the development of image retrieval systems for exploratory search. This kind of navigation is well suited

when the user is not able to formulate a specific query. In this case, the similarities between image data are exploited in order to display similar images close to each other. The most popular approach in this category is *mapping*. The main idea behind the mapping methods is to project the feature space in which the images are represented into a two-dimensional space to be displayed in a computer or other device screen. Other methods employ clustering algorithms to create groups of images, which are displayed in a grid layout or in such a way prone to overlapping. Overlapping occurs because similar clusters can be very close to each other in the feature space. For this reason, sometimes only a single image (representative) or a subset for each group is displayed in order to free up visualization space and reduce overlapping as in [7]. In [8], mapping and clustering techniques are combined to improve the provided layout.

In general, all the previous research work provide the end-users with a static structure, penalizing any customization to fit their specific needs. Therefore, they can interact just with parts of a same structure, without changing the precomputed rigid structure. This paper aims to fill such a gap. Here, we introduce a flexible framework, which is able to dynamically create several hybrid and coherent layouts by combining different visualization strategies, making our approach very adaptive to user interests.

3 Proposed Approach

We present an innovative framework for supporting the dynamic creation of several visual designs. Thus, we provide a more coherent space for searching and browsing images by organizing them in a well-defined visual structure, which can be adapted according to user interests. In this way, we introduce several original layouts, which are produced by our framework in a straightforward manner.

3.1 Features and Similarity

In our framework, we implement a traditional image retrieval system. Thus, an image descriptor is used to encode visual properties of an image into a feature vector and to measure the similarity between images by means of a distance function used for comparing their feature vectors. For that, we have adopted the Border/Interior pixel Classification (BIC) [9]. However, our framework does not impose any restriction on the descriptor used. It only uses the set of elements to be analyzed, which are not the images themselves but rather their extracted features, along with a suitable distance function. Therefore, it can be used with any searching and browsing scheme.

3.2 Intuitive Layout

An intuitive layout should ideally assist users to quickly comprehend search results preventing them from wasting time. For that, we need to coherently distribute the images over the available space. With this goal, we use clustering techniques to place search results over well-defined radial layouts. The advantage of using a hybrid scheme is to combine positive aspects of different visualization

strategies. One first step on that direction was the proposal of the Clustering Set strategy [10]. Figure 2c illustrates such strategy.

In this work, we are extending the Clustering Set principles to create novel visual structures. First, the query image is placed in the center of the visualization display. Then, the clusters are circularly distributed around the query in a clockwise order of similarity regarding the query, which is represented by the size of each cluster. This strategy is also applied inside each of the clusters. At the center, we place its most representative element. The remaining elements are sorted in a circular manner according to a clockwise order of similarity with respect to the center element, which is denoted by their size and border color.

In this scenario, we can observe the presence of two levels of radial visual structures. Notice that several radial visual structures are possible to be used for each level. It would also be possible to increment the number of levels. If we keep the same radial structure for $2 \ldots n$ levels, we can recursively create fractal-based structures, as discussed later.

3.3 Flexible Design

By using our framework, users can dynamically create new visual designs and modify the previous ones independently of the core and additional functionalities of the system. In order to provide a flexible framework, we need to support the easy creation and change of layouts. For that, we assume that visual structure developers may be interested in constructing complex diagrams from simple components. Thus, they may group components to form more complex structures.

A simple implementation could define classes for basic elements, such as a simple node containing an image, and other classes that act as containers of these basic elements, forming a structure. The problem with this approach is that it needs to handle the basic elements and the containers in a different way, which makes the application development more complex.

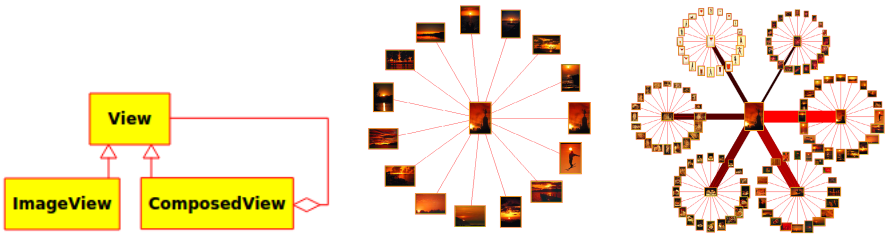
We can avoid this distinction using a recursive composition provided by the composite design pattern [11]. It establishes three basic elements: a component, a composite, and other primitive objects. The composite and the other primitive objects inherit from component. A component is a common interface that represents the object to be manipulated, it can be a primitive object or a complex set of other components (composite). It is possible to group together simple components to form bigger components that, at the same time, can be combined in order to form even more complex components.

In this way, we have separated elements and use them alone or in conjunction according to our needs. For example, when resizing the structure, we can recursively call the same function of every individual element so that all of them are resized as well, thus acting as a single unity. Thus, we can use a single interface to represent both the basic elements and the containers. This strategy enables us to treat an image or a structure of images in the same way as a simple element.

In Figure 2a, we define a primitive object that represents the basic cell (*ImageView*) containing an image, a composite object (*ComposedView*) that encompasses a list of objects (primitive of not) forming a visual structure, and a component (*View*) that can be either a primitive or a composite object. Note that a *ComposedView* is composed of a set of *View* elements. Thus, complex visual

structures can be produced by the successive composition of simpler structures. If we use a hierarchical clustering method, several levels can be included in our structure. This extension is absolutely simple, since it does not require that an element has a specific nature: a single element (image) or a complex one (set of elements).

To illustrate the potential of the composite design pattern, consider that we initially have a *ComposedView* element, which is formed by a set of basic *ImageView* elements, as shown in Figure 2b. Once they share the same interface *View*, we can equally treat a basic or a complex element. Then, suppose that we replace each basic *ImageView* element in this structure by a *ComposedView* element. In this way, we can easily create a more complex structure, as presented in Figure 2c.



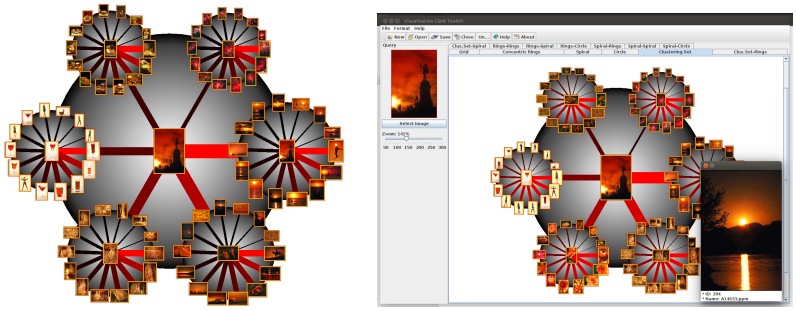
(a) Framework basic structure **(b)** Simple radial layout **(c)** Complex radial layout

Fig. 2. The potential of the proposed framework in creating complex layouts from simpler ones. A layout is composed by Views, which can be either a single ImageView or a ComposedView made of a set of Views, as shown in (a). Thus, a simple radial layout (b) can become a more complex layout (c).

This approach can be extended to any radial-based visual design, such as illustrated in Figures 1a, 1b, 1c. In other words, any *ImageView* element can be replaced by a *ComposedView* element. Therefore, many complex radial layouts can be created from those simpler ones. Moreover, we can repeat this process many and many times so to produce fractal-based visual designs.

3.4 Visual Marks

We named as visual mark to any visual property used in a graphical display (e.g., size, color, and border). A visual mark is extremely useful tool to emphasize user needs. For instance, the background color applied in other approaches (e.g., Concentric Rings) has assisted the users to understand the distribution of the results according to their opinions. Such a characteristic is useful to indicate relevance. Clearer tone is applied to more relevant information and displayed elements get darker as they are placed far away from the query image (most relevant information tends to be nearest to the center where we locate the query image). Figure 3a shows an example of improved Clustering Set layout that considers a different background color.



(a) Improved Clustering Set **(b)** The pop-up window where a specific image is handled

Fig. 3. Some visual marks and browsing capabilities provided by our framework

3.5 Advanced Browsing Capabilities

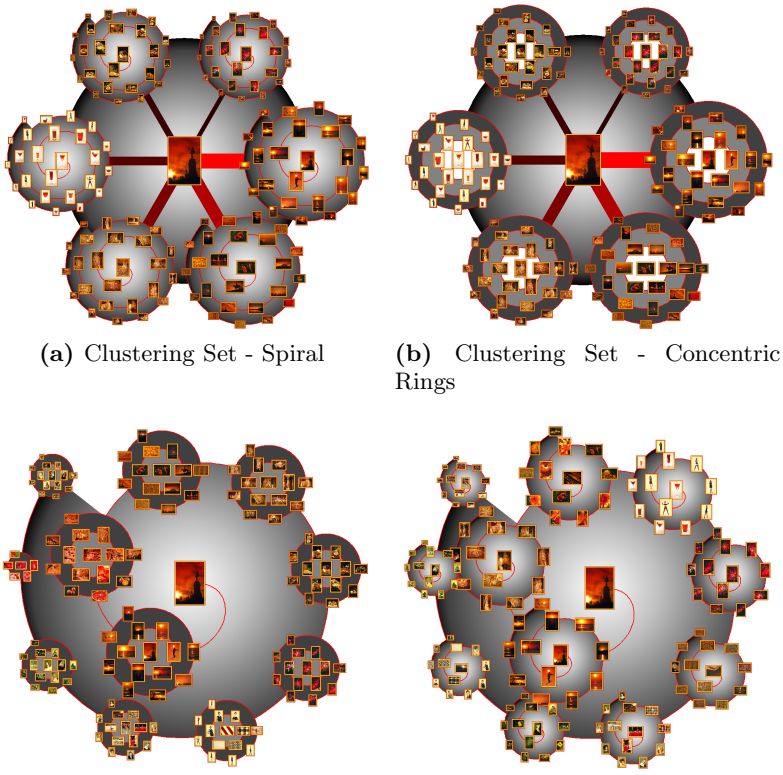
The addition of a dynamic user control through direct manipulation principles enables a user to navigate through large information spaces. In this way, given the increasing volumes of data, it is important to allow the dynamic manipulation through the image set by making available several browsing capabilities. When trying to offer a comfortable navigation, a useful interaction technique is output-as-input. It consists of providing the possibility of performing a new input into a system based on the previous output. In our proposal, anytime a user right-clicks on any already presented image (output) on the screen, it is possible to make a new query (input). The display will be automatically rearranged in order to show a new result of the most relevant clusters in relation to the last selected image. Thus, they can interleave between targeted search and exploratory search. Using different searching and browsing methods into a single environment enhance the user's inspection ability.

Our approach also offers zoom and pan controls. Furthermore, our framework takes advantage of the inspection ability of the users by adding a pop-up functionality to any clicked image. Specifically, when a user clicks on an image, a pop-up window is opened on the screen providing more information about the selected element, as illustrated in Figure 3b. On the top and center of the displayed pop-up window, we make available a image visualizer. On the bottom part, we display a set of image characteristics. This technique is known as details-on-demand. The details-on-demand technique allows to interactively select parts of data to be visualized in more detail while providing, at the same time, an overview of the whole context.

3.6 Hybrid Combinations

Our framework allows the straightforward creation of new structures from the combination of basic ones. Thus, we present several hybrid visualization structures as an original contribution from this work. This set of structures is guided in the Clustering Set principles. Figure 4a presents an hybrid structure that combines the successful Clustering Set approach with the Spiral method. In this case,

we apply the clustering strategy and an approach for placing results sequentially inside each cluster. Thus, we maintain the intrinsic relationship among images arranged in a radial well-organized structure that offers a better utilization of the space. Figure 4b presents a similar approach that uses concentric rings to display an ordered list inside each cluster. As we can see in Figures 4a and 4b, we have dealt with some strategies for sequential results in order to increase the provided information inside each cluster. Then, we can follow a similar criterion to display the sequential list of clusters. Thus, we propose other methods (Figure 4c and 4d) to offer more information using the same space. In these layouts, more clusters are displayed in a coherent manner.



(a) Clustering Set - Spiral

(b) Clustering Set - Concentric Rings

(c) Spiral - Concentric Rings

(d) Spiral - Spiral

Fig. 4. Several layouts created in a straightforward manner by our framework

4 Evaluation

Unlike other research areas, evaluating a visualization strategy is not a straightforward task due to the lack of an objective ground truth. A consistent evaluation framework is seriously missing for visualization research. Presently, every work has its own evaluation methodology, often presented without any performance comparison with previously existing techniques.

4.1 Experimental Protocol

In this work, we adopted a subjective evaluation framework to assess the quality of visualization structures, known as *DECIDE* [12]. This framework incorporates the judgement of the user in evaluating the quality of a visualization structure. It guides the experimental evaluation through six well-defined steps: **D**etermine the goals, **E**xplore the questions, **C**hoose the evaluation paradigm and techniques, **I**dentify the practical issues, **D**ecide how to deal with the ethical issues, **E**valuate, interpret, and present the data. In the following subsections, each of those steps is explained in more detail.

Determine the Goals. The goal of the experiment is to evaluate a framework for visualization that provides a set of novel techniques. This framework follows the same strategy of the clustering set method, thus likewise the proposed techniques. In order to do that we need to ***(i)*** evaluate the effectiveness of the proposed techniques and ***(ii)*** compare it with current state-of-the-art methods. Therefore, we defined two stages to accomplish both aims.

Explore the Questions. In order to achieve our goals, we defined some questions for assessing the user preferences regarding the proposed techniques (first stage):

- Is it possible to identify the query pattern (e.g., query image)?
- Is it possible to determine where the most relevant results are placed?
- Is it possible to determine where the least relevant results are placed?

We have also analyzed the introduced techniques in conjunction with state-of-the-art techniques following the questions (second stage):

- Does it provide the necessary information? (not too much, not too little)
- Does it appropriately use the available space?
- Is it possible to understand the visual structure and how it distributes the result set?
- Is the user satisfied with the presented layout?
- What is the most suitable method to visualize the query results?

For obtaining unaware information, we also consider open questions for general comments about the evaluated approaches.

Choose the Evaluation Paradigm and Techniques. In this step, we adopted the Usability Test method (the standard ISO 9241). The evaluation instrument used in our experiments was a questionnaire. In this way, we are able to analyze a bigger sample in a faster manner. A bigger sample can give us a better statistical significance.

Identify the Practical Issues.

- **Users.** In order to obtain significant results, we invited a set of undergraduate students from the Institute of Computing, University of Campinas, totalizing 28 volunteers. It is important to point out that nobody had previous knowledge of the proposed methods and, hence, they had never seen our strategy before the experiment.

- **Equipment.** Each participant had access to one computer and was free to use the operating system he/she prefers.
- **Material.** Each participant accessed a set of files. This set included:
  - User Instructions. It details what the user needs to know about the evaluation process.
  - Free and Clear Consent Terms. It clearly describes the conditions in which the user will participate in the experiment detailing risks, benefits, confidentiality, etc.
  - User Profile Form. It allows to register relevant data of an user, such as his/her familiarity with the system and frequency of using the computer.
  - Evaluation Forms. We use two evaluation forms. By using the first one, we were able to objectively capture the users' ability to find some elements in the proposed structures (first stage), and by using the second one, we were able to capture the users' judgment about the proposed methods along with state-of-the-art methods (second stage).

Decide How to Deal with the Ethical Issues. Ethical questions were clearly explained to each user verbally and in the user instructions.

Evaluate, Interpret, and Present the Data. Initially, we gave the users the "Users Instruction" and the "Free and Clear Consent Terms." Then, users read both documents and signed them. After that, they filled the user profile form online. Finally, they analyzed each visual structure prototype and evaluate the visual structures by answering the two evaluation forms.

4.2 Results

Each form allows us to obtain specific information from the users and their perceptions regarding each evaluated method.

The profile form served us to obtain personal information such as age, gender, studies, familiarity, and frequency of using computers and search systems. From the 28 participants at the moment of the experiment, nearly all of them (96.42%) were from 19 to 26 years old and only one user (representing the 3.6%) was 33 years old. The 78.57% of the users were male and the remaining 21.43% were female. All the users were undergraduate students in a computing course with a high-level familiarity and frequency use of computers. Most of them had a lot of experience with searching systems.

The first stage of the evaluation helped us to objectively analyze the real effectiveness of the used approach over the several novel methods proposed in this work and the previously proposed Clustering Set method. These methods follow the same basis: they use clustering algorithms to offer a coherent space in conjunction with well-defined radial layouts to coherently organize that information.

Figure 5 exhibits the percentage of users that correctly found the requested element or region inside each structure (% of Correct Responses). The results for the different questions are grouped by the visual designs. This graph shows that our strategies have achieved a good performance. In almost all questions involving all the proposed structures, more than 80% of the users were able to correctly identify the requested elements, demonstrating a high level of usability. No remarkable difference between the methods statistics was found, probably because

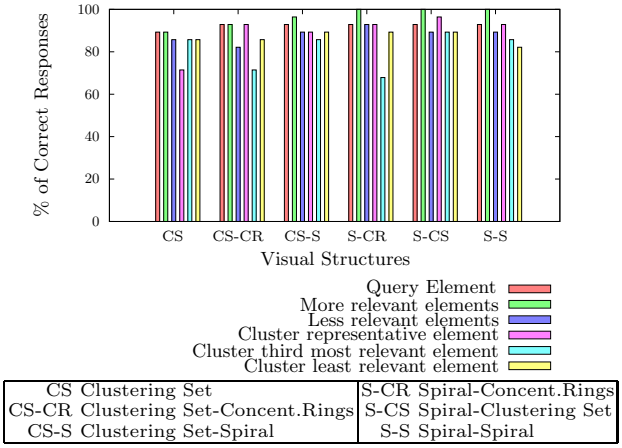


Fig. 5. Usability of each method according correct answers

they all follow the same philosophy: coherent clustered information organized in well-defined radial structures.

A summarized plot of the answers of this stage is displayed in Figure 6a which represents a Tukey-style boxplot [13] of the number of correct answers by structure for all users. The x axis varies from 1 to 6 as the evaluation form for this stage requested six questions for each method. The y axis represents each evaluated method. We can observe that the hybrid methods involving a combination of Clustering Set with Spiral (S-CS and CS-S) obtained very good results with a low dispersion of data and a highly compact boxplot, representing that most users was able to correctly identify all the elements in such structures. Thus, spiral layouts seem to be more intuitive for users.

The second stage of the evaluation consisted of a rigorous subjective evaluation to capture the users judgments of several methods from literature and from the proposed framework. After collecting the data from this second evaluation form, we obtained the results shown in Figures 6b, 6c, 6d, 6e. These graphs present an overall analysis of each criterion also using the Tukey-style boxplots. Such figures indicate that there is no statistical difference between the proposed method and the already familiar Grid method. Therefore, the strategies created by our framework can be seen as a truly alternative to the popular Grid method. This is a significant achievement since it is difficult to compete with a widely known method (Grid) considering that users are already familiarized with it.

Users were also asked to rank the structures according to their suitability for visualizing query results. Figure 6f exhibits the distribution of ranks attributed by the participants for different approaches. The comparison of the average rank of those methods is showed in Table 1.

The non-parametric Friedman test [14] was performed to verify the statistical significance of those results. It checks whether the measured average ranks are significantly different from the mean rank. The analysis of the experiment confirms that there is no statistical difference among those methods (p-value

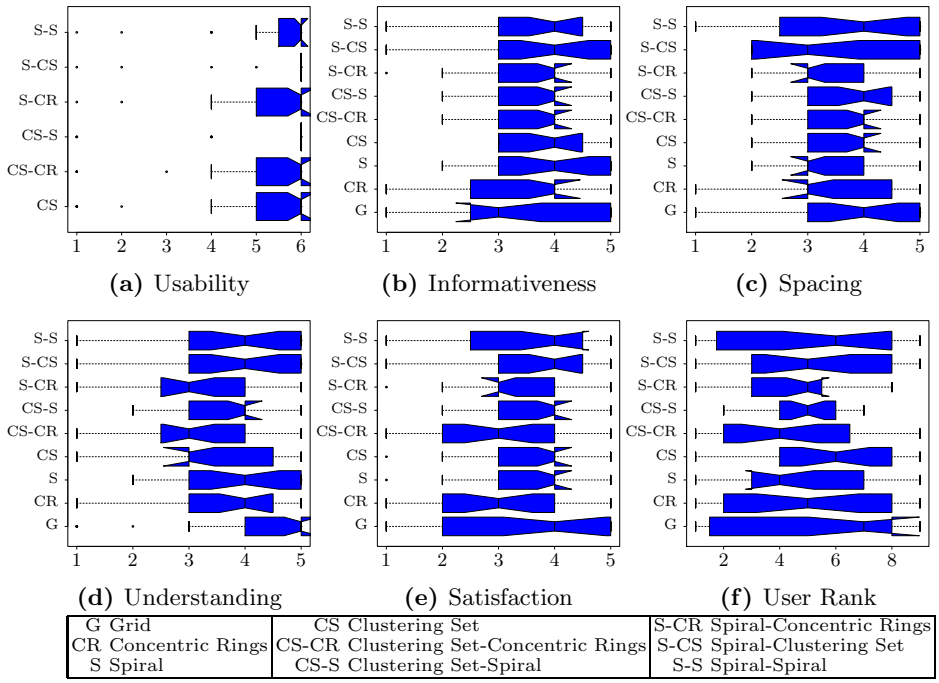


Fig. 6. Results from our experiments

Table 1. Average rank achieved by several approaches

Method	Average Rank
Clustering Set	5.65
Spiral - Clustering Set	5.48
Grid	5.19
Spiral - Spiral	5.11
Concentric Rings	4.98
Spiral	4.80
Clustering Set - Spiral	4.80
Clustering Set - Concentric Rings	4.65
Spiral - Concentric Rings	4.35

= 0.788). According to a Nemenyi test, the critical distance (CD) for pairwise comparisons between different strategies at a confidence of 95% is 2.31.

5 Conclusion

This paper presents a novel framework for creating visual structures to support search and browse services in image retrieval systems. The proposed framework allows the construction of coherent layouts that emphasize the more relevant results and, at the same time, the comprehension of the relations among them.

User studies have demonstrated that several hybrid visual structures created by the proposed framework yield, according to different criteria, similar results than those observed for traditional methods (e.g., Grid method). This is a relevant result considering the fact that the users have never saw our proposed visual structures before, and they were very familiar with the Grid-based structures.

Future work includes the extension of our framework to consider other descriptors (e.g., local features [15]), learning-to-rank methods (e.g., genetic programming [16]) for combining descriptors and hierarchical clustering methods [17].

Acknowledgments. We would like to thank CAPES, CNPq, FAPESP for the financial support. Thanks to the participants of the experiments.

References

1. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Computer Survey 40(2), 1–60 (2008)
2. Heesch, D.: A survey of browsing models for content based image retrieval. Multimedia Tools Appl. 40(2), 261–284 (2008)
3. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by image and video content: The QBIC system. IEEE Computer 28(9), 23–32 (1995)
4. Mukherjea, S., Hirata, K., Hara, Y.: Amore: A world wide web image retrieval engine. World Wide Web 2(3), 115–132 (1999)
5. Draper, G.M., Livnat, Y., Riesenfeld, R.F.: A survey of radial methods for information visualization. IEEE Trans. Visualization and Computer Graphics 15, 759–776 (2009)
6. Torres, R.d.S., Silva, C.G., Medeiros, C.B., Rocha, H.V.: Visual structures for image browsing. In: ACM CIKM, pp. 49–55 (2003)
7. Schaefer, G., Ruszala, S.: Hierarchical image database navigation on a hue sphere. In: Bebis, G., et al. (eds.) ISVC 2006. LNCS, vol. 4292, pp. 814–823. Springer, Heidelberg (2006)
8. Nguyen, G.P., Worring, M.: Interactive access to large image collections using similarity-based visualization. J. Visual Languages and Computing 19(2), 203–224 (2008)
9. Stehling, R.O., Nascimento, M.A., Falcão, A.X.: A compact and efficient image retrieval approach based on border/interior pixel classification. In: ACM Intl. Conf. Information and Knowledge Management, pp. 102–109 (2002)
10. Pinto-Caceres, S.M., Almeida, J., de Almeida Néris, V.P., Baranauskas, M.C.C., Leite, N.J., da Silva Torres, R.: Navigating through video stories using clustering sets. Intl. J. Multimedia Data Engineering and Management 2(3), 1–20 (2011)
11. Erich Gamma, R.J., Helm, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley (1994)
12. Preece, J., Rogers, Y., Sharp, H.: Interaction Design. John Wiley & Sons, Inc., Chichester (2002)
13. Tukey, J.W.: Exploratory Data Analysis. Addison Wesley (1977)
14. Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. American Statistical Association 32(200), 675–701 (1937)
15. Almeida, J., Rocha, A., da, R., Torres, S., Goldenstein, S.: Making colors worth more than a thousand words. In: ACM Intl. Symp. Applied Computing (ACM-SAC 2008), pp. 1180–1186 (2008)

16. Andrade, F.S.P., Almeida, J., Pedrini, H., da, R., Torres, S.: Fusion of local and global descriptors for content-based image and video retrieval. In: Iberoamerican Congress on Pattern Recognition (CIARP 2012), pp. 845–853 (2012)
17. Rocha, A., Almeida, J., Nascimento, M.A., Torres, R., Goldenstein, S.K.: Efficient and flexible cluster-and-search for CBIR. In: Blanc-Talon, J., Bourennane, S., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2008. LNCS, vol. 5259, pp. 77–88. Springer, Heidelberg (2008)