Large Image Collections – Comprehension and Familiarization by Interactive Visual Analysis

Krešimir Matković[1], Denis Gračanin[2], Wolfgang Freiler[1], Jana Banova[3], and Helwig Hauser[4]

[1] VRVis Research Center in Vienna, Austria
{Matkovic,Freiler}@VRVis.at
[2] Virginia Tech
gracanin@vt.edu
[3] PRIP, Vienna University of Technology, Austria
banovaja@prip.tuwien.ac.at
[4] University of Bergen, Norway
helwig.hauser@uib.no

Abstract. Large size and complex multi-dimensional characteristics of image collections demand a multifaceted approach to exploration and analysis providing better comprehension and appreciation. We explore large and complex data-sets composed of images and parameters describing the images. We describe a novel approach providing new and exciting opportunities for the exploration and understanding of such data-sets. We utilize coordinated, multiple views for interactive visual analysis of all parameters. Besides iterative refinement and drill-down in the image parameters space, exploring such data-sets requires a different approach since visual content cannot be completely parameterized. We simultaneously brush the visual content and the image parameter values. The user provides a visual hint (using an image) for brushing in addition to providing a complete image parameters specification. We illustrate our approach on a data-set of more than 26,000 images from *Flickr*. The developed approach can be used in many application areas, including sociology, marketing, or everyday use.

Keywords: Interactive Visual Analysis, Coordinated Multiple Views. Large Image Collections.

1 Introduction

The popularity of digital photography and the exponential growth of Internet resulted in the emergence of a new phenomenon — very large Internet image collections. They offer new and interesting opportunities to share images with friends, relatives, or public. Image collections can be used in a number of ways.

Sociologists can investigate the available data about the users uploading images, the number of comments an image initiated, details about the people who commented, etc. Marketing professionals can analyze preferences and characteristics of a specific target group. For example, they could determine that men

A. Butz et al. (Eds.): SG 2009, LNCS 5531, pp. 15–26, 2009.

between 25 and 30 prefer blue, highly saturated images with flowers (this is just a fictitious example). After identifying a preference, they can fine-tune marketing campaigns, e.g., for personal marketing. Individual users can also profit from exploring such image collections. We are fascinated by "new and unknown worlds" captured in images and related information (e.g. Google Earth). We can establish social contacts with other users through such online sites. There is a lot of, still mostly unused, information in Internet image collections.

Due to sheer size and complexity of image collections it is impossible to understand them by only viewing images or by only investigating abstract image parameters. We use multiple, coordinated views to simultaneously explore images and related parameters and introduce innovative ways of parameter selection. The user can also use images in the selection process. This is not an image query system, or a tool for browsing of image collections. The main idea is in understanding hidden characteristics of images and collections. The user can intuitively and in an almost effortless way interact and alternate between images and their parameters to perform various types of exploration (e.g. multivariate parameter analysis). We propose a manifold integration of proven approaches (with a couple of new aspects) to make it possible (for the first time — to the best of our knowledge) to comprehend and familiarize large image collections.

The user can start with an overview of image parameters, focus only on a subgroup of users (e.g. female users), spot an interesting image and use it as a target. All similar images are selected and their parameters are highlighted. The user spots an outlier with respect to the averaged image hue (e.g. a green image) and discovers, for example, that images with similar color distributions are usually uploaded by male users. Using images and strict parameter specifications together as a filter offers unlimited possibilities to comprehend large image collections.

2 Related Work

Image collections grow rapidly — *Flickr* [4] exceeded 3 billions of images (Nov. 2008) with 5,000+ images uploaded every minute. There are other attempts to exploit large Internet image collections [12] but as far as we know we use a novel approach making it possible to comprehend and familiarize with such collections.

Roberts[11] provides an extensive report on the current state of the art of coordinated multiple views in exploratory visualization. Various techniques were developed to analyze series of numeric values. Hochheiser et al. [6] developed the *TimeSearcher* tool for visual analysis of time-dependent series of numerical values by displaying multiple graphs in the same display window. These graphs can be filtered by drawing *timeboxes*. Only graphs passing through all timeboxes are displayed. Konyha et al. [9] proposed the segmented curve view as another approach to analyze time series data. The series of values are represented by bars that are divided into multiple blocks. The color of each block becomes more saturated with the number of virtual graph lines passing through it.

Ahlberg and Shneiderman [1] describe the *Filmfinder* tool for exploring a large film database by interactive filtering using various criteria. By applying the dynamic queries approach to information filtering, a continuous starfield display of the films, and tight coupling among the components of the display, the *FilmFinder* encourages incremental and exploratory search.

Hyunmo and Shneiderman [8] developed *PhotoFinder*, a picture management application for flexible filtering of images by dynamic queries and previews. Bederson [2] described another tool, *PhotoMesa*, that uses quantum treemaps to arrange thumbnail pictures on the screen. Quantum treemaps improve the aspect ratio and reduce the waste of screen space.

Hu et al. [7] used Bregman Bubble Clustering for an algorithm to find dense clusters in image search results. They tested the algorithm using an image collection from *Flickr* [4] and made comparison with the *Flickr* search algorithm. When searching for common names, their clustering method has better results than the *Flickr* image search. Yang et al. [14] use semantic image segmentation in order to browse the image collections.

While all these approaches proved useful with respect to selected aspects of the here addressed (larger) challenge, we know of no approach which allows to utilize a multifaceted analysis, considering the images, image similarities, and image parameters altogether.

3 Image Collection Description

Conventionally, data analysis approaches, such as traditional statistics or OLAP techniques, assume the data to be of a relatively simple multi-dimensional model (simple with respect to the separate data dimensions). For complex data sets like image collections, it is necessary to provide an adequate data model. Each image is associated with a data point so the data set representing the image collection consists of data points (tuple values) of n dimensions. Each dimension represents a parameter describing an image. Parameter values can be categorical, numerical, and data series. The image itself is also a dimension in the data item.

There are several categories of parameters like external parameters (image caption/title; name, age or gender of the author, time the image is created, etc.), environmental parameters (file name, file size, file creation time, file modification time, etc.), content parameters (image format, pixels, etc.), and derived parameters. The derived parameters are derived (calculated) from the raw pixel data for example, hue, saturation, contrast, etc.

We analyze an image collection from *Flickr*. There are two groups of external parameters. The first group of six external parameters (flickr parameters) relates to the "use of an image" on the *flickr* site and includes *date taken* (the date when the image was taken), *number of comments*, *number of favorites*, *number of views*, *number of tags*, and *title*. The second group of ten external parameters (user parameters) relate to the user (image "owner") and includes *user name*, *gender*, *singleness*, *home town*, *location*, *occupation*, *image count per user*, *contact count*, *pro* (i.e. if the user has a pro account on Flickr), and *first date*.

We have more than 60 derived parameters. Averaged image hue, saturation, lightness, and contrast are computed for each image. In order to gain deeper insight into such collections we subdivide images also into cells of a regular grid and compute average hue, saturation, lightness, and contrast values for each cell in this grid (Figure 1b). Those parameter values constitute data series in the data tuple. We subdivide images in four different resolutions (rectangular grid): 2×2 (four cells) 3×3 (nine cells), 4×4 (16 cells), and 5×5 (25 cells).

In order to allow a more semantically based approach to color features we introduce color names based on the work by Van de Weijer et al. [13]. In English, eleven basic color terms have been defined based on a linguistic study. The normalized amount of pixels of each color is individually computed for each image. To additionally explore the colors in images colorfulness is measured by computing the Earth Mover's Distance [3] between an "ideal uniform colorful" distribution and the measured color distribution of the image. Simple average hue can be problematic, especially for images with a high amount of various red tones (the simple average of hue will suggest that the image is blue) or with low saturation (black, white and grey tones will have almost random hue values). Therefore a vector-based, saturation weighted formula is used to compute the average hue [5]. Further, we explore the texture and graininess in the images by computing the three-level decomposition of the Daubechies wavelet transform for each color channel (H/S/L). We have also calculated depth of field indicator [3]. The Viole-Jones face detector is used to locate frontal faces on the images.

Each of the 26,000 images we used (downloaded from Flickr using the publicly available Flickr API) corresponds to a tuple which contains the described parameters and the image itself. We can, through interactive visual analysis, analyze relationships among various parameters and get insight into image collections.

4 An Iterative and Integrated Approach to Comprehension an Familiarization of Image Collections

Conventional but yet very effective views such as histograms, scatterplots, or parallel coordinates are used for numeric or categorical data such as the average contrast, average hue/saturation, or photographers' gender and marital status. Figure 1a shows a histogram of average lightness values. The lightness range is divided into 16 bins and each bar shows how many images have an average lightness of a particular range. The sum of all values is the total sum of all images in the collection. We can clearly see that the images with medium average lightness are the most frequent in our dataset. We could easily depict all other parameters using the same view. Such views give us an overview of the image collection but often they are not enough to reveal fine-grained details or to visualize values that are averaged over the entire image.

4.1 Segmented Curve View

Besides numerical parameters we also have data series parameters. Average hue, e.g., for subrectangles can be observed as a series of numbers, each representing the hue of a subrectangle. There are many ways of visualizing series data, but

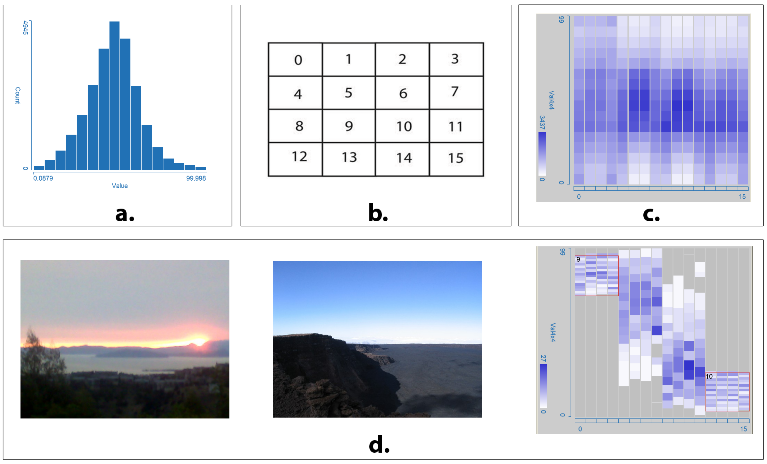

Fig. 1. a. A histogram showing the distribution of lightness values for an image collection consisting of more than 26,000 images. **b.** Subdivision of an image in 16 segments for partial parameter computation and corresponding cell numbering. **c.** A segmented curve view of lightness distributions in segments of all images in the collection. The x axis of the segmented curve view has 16 segments (16 subrectangles in b.) and each vertical bar is divided into 16 segments (bins). Bins are colored dependent on number of images having lightness in a particular segment in the range defined by a particular bin. **d.** A segmented curve view of lightness distributions. The first four bars represent the upper quarter of all images, the last four the lower quarter. We select images that are bright in the upper and dark in the lower part at the same time. Two images in the left illustrate typical images for such a selection.

most visualization methods were designed for continuous data (mostly for time series). The segmented curve view [9] is one of the exceptions. Due to the discrete nature of the mapping domain (number of rectangles), the segmented curve view seems to be more appropriate here. It provides a comprehensive view of a large family of data series. It depicts the distribution of a dependent variable (e.g. lightness) for each value of an independent variable (subrectangle in our case) for every data point (image in the collection). For each discrete value of an independent variable (subrectangle) there is one vertical bar in the segmented curve view. This bar is subdivided into segments (bins) and the number of items (images) passing through a bin is depicted using a color scale. The total number of items in a bar equals the total number of images in the collection. Each image in the collection is subdivided into rectangles.

The segmented curve view for lightness ('V' in HSV color system) shows lightness in all subrectangles for all the images (Figure 1c). For example, the first bar in the segmented curve view depicts distribution of lightness in subrectangle 0 (upper-left corner) for the whole image collection.

The leftmost bar in the segmented curve view shows a relatively uniform distribution across the entire range of lightness in our data set. This is apparent from very similar shades of blue for all 16 segments of the bar. This shows that the number of images with light, dark and medium light segment 0 is approximately equal in our collection. Segments 1 and 2 have slightly less frequent dark values. Segment 3 is similar to segment 0. Segments 5, 6, 9, and 10 (center of the images) have significantly less light or dark colors than other segments.

Figure 1d shows another example. The upper parts of the first four bars represent the images that are bright in the upper quarter. The lower parts of the last four bars represent the images with the dark lower quarter. Figure 1d shows the selection of images that are bright in the upper quarter and dark in lower quarter at the same time. The selection was done interactively by the user. The view shows distribution in the selected parts of the bins only, and the rest is shown in gray. Two images from such a selection are shown in Figure 1d on the left.

Although the segmented curve view and the approach with several subrectangles might seem complex at the first sight, the gained additional information is significantly larger than the information from simple overall average histograms, as illustrated in Figure 1a. If we would have only overall average values per image, we could not observe and analyze distributions of parameters across images in the collection as demonstrated in Figure 1d.

4.2 Forms of Interaction

Interaction is the core of interactive visual analysis and plays the most important role in coordinated, multiple views systems. The basic idea of linking and brushing makes it possible to select a range of parameters in one view (e.g., specific lightness values) and to see which parameters the selected items have. The selection is traditionally called brushing in interactive visualization. Figure 2 illustrates linking and brushing.

The histogram on the left (Figure 2 top row) shows a user's selection (made by a mouse) of highly saturated images. A linked parallel coordinates view of eight further parameters is shown as well. The selected images are depicted in red and the context (all other images) is gray. Figure 2 shows the lightness distribution for 16 segments of the selected images on the right of the top row. The distribution has all bins covered. Figure 1c shows a different distribution in the segmented curve view for all images in the collection. Note many images with dark corners (dark blue lowest segments in bars 0, 3, 12, and 15).

The user is not limited to a single brush but can combine many brushes and use Boolean operations on them. Let us refine the previous selection now. If we add another histogram and additionally select only single users and only taken users (taken is Flickr terminology used for non-single users) we can explore the difference in lightness distribution of highly saturated images for these two groups (Figure 2 bottom row). Note how singles tend to have much darker corners!

Image as a Brush. The nature of images, their complexity and our difficulty to visualize images from parameters require additional interaction. Even though we get deeper insights by utilizing subrectangle values and advanced linking and brushing, it is still not enough. It is impossible to correctly mentally visualize an image only from the parameters. The user still prefers the image itself. We have to include the original images in the analysis. The ability to treat image in the

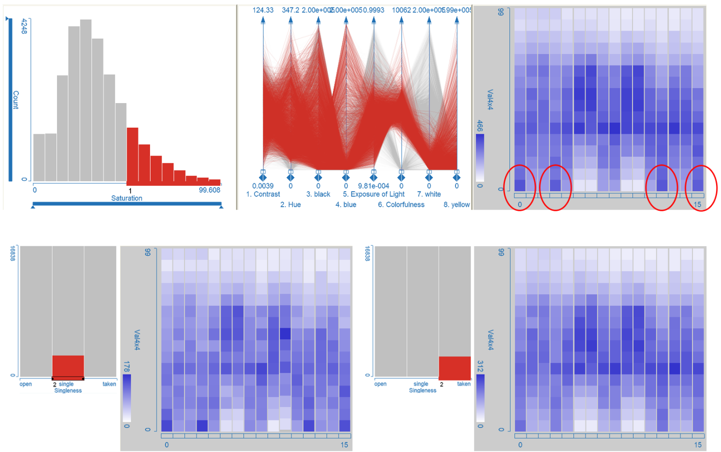

Fig. 2. Top Row: User brushes images with high average saturation selecting only the right hand side bins in the saturation histogram shown on the left. We can see the distribution of eight other parameters for the selected images in the parallel coordinates view in the middle. Finally, the segmented curve view on the right shows a more detailed view on lightness of selected images. Note the differences between this view and Figure 1c. Note also, many images with dark corners (bins in red ellipsis) for images with high saturation. Bottom row: If we refine selection now differentiating between single and taken users (Flickr terminology) we can see the difference in lightness distribution. The dark corners are more probable by single users.

same way we treat other parameters (brushing and linking) provides a unique opportunity to analyze large image collections. We provide a linked picture view showing the original images. Now the user can see not only the parameters of selected images, but the images themselves. When we see the selected images together, some similarity which can not be seen in the parameter space might become visible. The picture view is not only passively linked, it can be used for brushing, too.

The idea is to let the user select an image, the system then finds similar images, and images that are similar enough define a brush now. Since all views are linked, the user can see the parameters of the selection in the parameters views. This is a completely new way of brushing. The user does not select the exact parameters but a target image. Refining the search with additional parameter brushes, image brushes, and their combinations provides unlimited possibilities.

In order to realize such a system we needed a definition of image similarity. There are many content based image retrieval systems, techniques and similarity measures that can be used. We have used the Visual Image Query [10] based on color layout similarity. It serves as a tool to allow us brushing by image selection and can be easily replaced or enhanced with any other available technique. The query algorithm is only a tool for us, it is not a topic of this paper.

When we select an image as a target image, the underlying algorithm is too complex for us to predict which images will be selected. In contrast to conventional brushing in the parameter space where the exact parameter limits are specified, using an image as a brush represents a fuzzy form of brushing. The user hints a color layout by image selection and the system responds with a new selection. This adds a different quality to this way of brushing, and makes it particulary convenient for image collections analysis. The interplay between brushing in the parameter space, and brushing by the selection of a target image creates a unique system for analysis and exploration of large image collections.

Selections can contain a lot of images so it may not be possible to display them all using a limited display space. We display as many thumbnails as possible in the view and allow the user to see the next and previous page with thumbnails. The user can specify the criterion for image sorting and influence the order of the resulting images (e.g., display highly saturated images first). Any of the available parameters can be used for sorting. In this way some hidden relations can be discovered, and image collections can be understood better.

5 Demonstration

We illustrate interaction possibilities (parameters, images, sorting) of our approach using a general purpose interactive visualization tool enhanced with the image view and image brushing capabilities.

First we investigate if there is a pattern in parameters of popular images, and if there is a difference between popular images of single and taken photographs. If we find that there is a difference this information can be used in, e.g., a marketing campaign. If target customers are single, one could use images with parameters that appeal singles.

We use parallel coordinates to select parameters and see the results. Parallel axes are used and the values belonging to one image are connected with a polyline, depending on image parameters. There is a polyline for each image in the collection. Figure 3 top left shows parallel coordinates with multiple selections.

We select images with a large number of comments, views and tags. We also select taken users (left set of images) and single users (right set of images). There is a large difference in the lightness and saturation distribution between these two groups. Popular images of taken users are much more saturated. There is a large number of unsaturated images in the singles' group (dark rectangles at the bottom of the segmented curve view in the right hand side set.) The lightness distribution in the taken users' images is also more regular. There are many images with medium lightness. The parameters depicted in lower parallel coordinates are similar, and finally the picture view shows 25 images form each groups with most comments. The saturation characteristic can be clearly seen in the images. The interpretation of the results, i.e. why popular images of single users tend to have low saturation is out of scope of this paper. We intend to analyze these results with sociologists in order to interpret the results.

We now restrict our selection in Figure 4 to those images that are favorites of many users. We examine the histograms to find a correlation between the popularity of images and the author's gender. We also examine if the images of

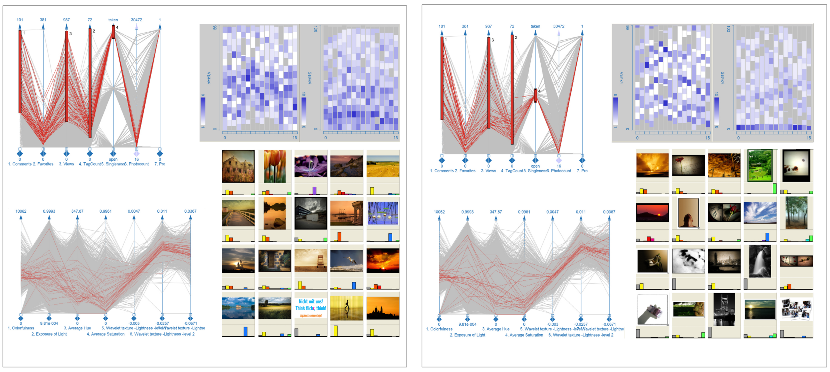

Fig. 3. Images with many comments, views, and tags are selected using parallel coordinates (numbers next to the selections are used to identify various brushes only.)Left images show additional selection of taken users, and images of single users are shown on the right. The distribution of lightness and saturation is shown in segmented curve views. The 25 images with most comments from each group are shown in the picture view. Different saturation distribution and different mood in the images.

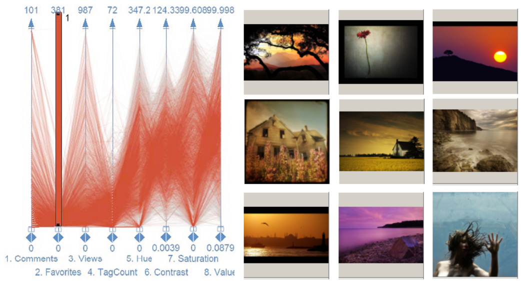

Fig. 4. Left: In a parallel coordinates view we can brush pictures, that have been selected by many users as their favorites. This way we can concentrate on popular images. **Right:** The most popular pictures of the result set.

singles or taken authors are more popular. The histogram shows the exact values of popular images in relation to all images, so we can conclude, that about 7% of all pictures taken by women are popular, whereas only 5% of pictures taken by men are popular. That means pictures taken by women in this case have 38% higher chance of becoming favorites. The influence of an author's singleness is smaller, but still measurable. The single users' pictures have about 4.9% chance to become popular. If the image author is in a relationship, the chance is 6.3% or about 29% more. Figure 4 shows the most popular pictures from our collection.

If we restrict our selection to even more popular images (Figure 5), the differences between male and female authors further increase, while the differences between single and taken users decrease a little bit. We can view images taken by female and male authors separately, and look for differences. Figure 5 shows

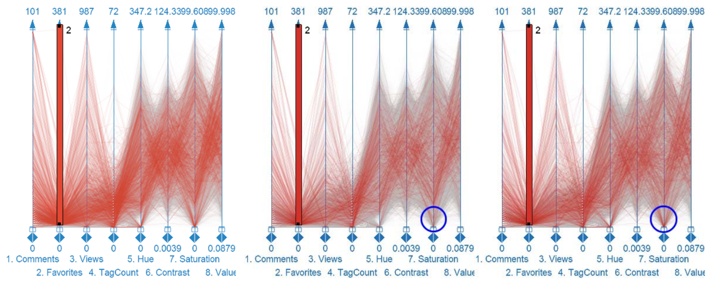

Fig. 5. Left: We can resize our brush and cut away some of the pictures that are less often chosen as favorites. **Middle:** Popular images from female authors. **Right:** Popular images from male authors. The blue circles emphasize pictures with very low saturation. These are greyscale images which are more often taken by male authors.

a comparison of the parallel coordinates view of female and male authors. The blue circle indicates an obvious difference in the saturation axis. Male authors seem to prefer low saturated photos, that are favorites of other users.

After applying a brush to the saturation axis, we view the resulting images and check if the trends we found before are still valid. Both trends are now reversed. For greyscale images (low saturation), photos taken by men have 30% higher chance to become favorites. Single authors outperform taken ones by 22%.

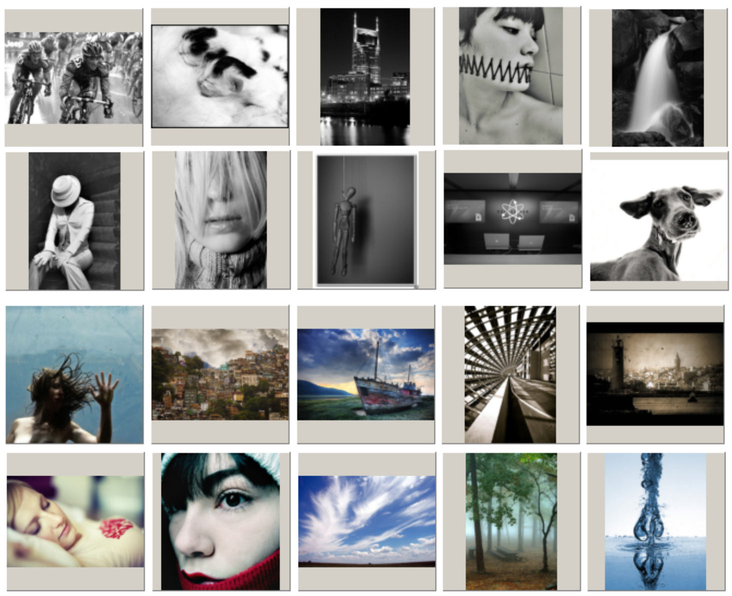

Fig. 6. Top: These greyscale pictures have been selected as favorites by many users. **Bottom:** These popular images (subset shown) have a medium saturation and are more often taken by female authors.

Fig. 7. The user provided a landscape image as brush. The selection was further narrowed to male and female users, and interestingly, in our image collection, female users (a.) tend to upload landscapes with some additional objects in the image.

When selecting medium saturated pictures however, we can reproduce the trends we found in the beginning. Pictures taken by female authors now have a 66% higher chance to become popular. The chances of pictures taken by the users in a relationship is more than 42% higher than for the singles. The most popular pictures with a medium saturation can be seen in Figure 6 bottom.

Finally, we show an example with a target image. A landscape image is used to look for landscape images. The resulting pictures are shown in Figure 7. As expected, we got a lot of landscape images. Some of them were pure landscapes and some of them included people, houses, or other objects. We now easily refine the search in the parameter space to see if this difference is related to some parameters. After short exploration we found out that in our collection landscape images uploaded by female users (Figure 7a) often contain some additional objects and those by males usually do not (Figure 7b).

The described examples show various possibilities of proposed approach. Interplay of image space and parameter space and possibility to iterative drill down to the wanted information using various interaction represent a novel and exciting way of understanding large image collections.

6 Summary and Conclusions

We describe a novel approach to analysis of large image collections that provides additional insights about image collection parameters, such as metadata or derived parameters. We can explore image collections and find patterns in image parameters while maintaining the focus on visual impression. It adds new qualities to image collection analysis and understanding and facilitates better comprehension and familiarization. It can be used in numerous scenarios for marketing research, sociology research, etc., or just by common users to get insight into image collections. The process is fun, fascinating and intuitive, both for domain experts or researchers, as well as for regular users "navigating on the sea of images." We have introduced a novel concept of image brush for interactive visual analysis. Future work will focus on increasing the scalability of the implementation and new input (e.g. multi-touch) and output (high-resolution) technologies.

References

1. Ahlberg, C., Shneiderman, B.: Visual information seeking using the FilmFinder. In: ACM CHI 1994 Conference Companion, pp. 433–434 (April 1994)
2. Bederson, B.: Quantum treemaps and bubblemaps for a zoomable image browser (2001)
3. Datta, R., Joshi, D., Li, J., Wang, J.: Studying aesthetics in photographic images using a computational approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 288–301. Springer, Heidelberg (2006)
4. Flickr photo sharing community (2007), `http://www.flickr.com`
5. Hanbury, A., Kropatsch, W.G.: Colour statistics for matching in image databases. In: Vision in a Dynamical World 27th AGM Workshop, pp. 221–228. Österr, Arbeitsgemeinschaft für Mustererkennung (2003)
6. Hochheiser, H., Shneiderman, B.: Dynamic query tools for time series data sets: timebox widgets for interactive exploration. Information Visualization 3(1) (2004)
7. Hu, Y., Yu, N., Li, Z., Li, M.: Image search result clustering and re-ranking via partial grouping. In: IEEE Inter. Conf. on Multimedia and Expo (2007)
8. Kang, H., Shneiderman, B.: Visualization methods for personal photo collections: Browsing and searching in the photofinder. In: IEEE International Conf. on Multimedia and Expo (III), pp. 1539–1542 (2000)
9. Konyha, Z., Matkovic, K., Gracanin, D., Duras, M.: Interactive visual analysis of a timing chain drive using segmented curve view and other coordinated views. In: Coordinated and Multiple Views in Exploratory Visualization, Zurich (2007)
10. Matković, K., Neumann, L., Siglaer, J., Kompast, M., Purgathofer, W.: Visual image query. In: SMARTGRAPH 2002: Proceedings of the 2nd international symposium on Smart graphics, pp. 116–123. ACM Press, New York (2002)
11. Roberts, J.C.: State of the art: Coord. & multiple views in exploratory visualization. In: Coord. and Multiple Views in Exploratory Visualization, Zurich (2007)
12. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3d. ACM Trans. Graph. 25(3), 835–846 (2006)
13. van de Weijer, J., Schmid, C., Verbeek, J.: Learning color names from real-world images. In: Proceedings of IEEE Computer Vision and Pattern Recognition (2007)
14. Yang, J., Fan, J., Hubball, D., Gao, Y., Luo, H., Ribarsky, W., Ward, M.: Semantic image browser: Bridging information visualization with automated intelligent image analysis. In: IEEE Visual Analytics Science And Technology (2006)