Content selection based on compositional image quality

Pere Obrador
Media Technologies Laboratory
HP Laboratories Palo Alto
HPL-2006-166
November 9, 2006*

no reference image quality, sharpness metric, document layout, document balance, photographic composition

Digital publishing workflows usually have the need for composition and balance within the document, where certain photographs will have to be chosen according to the overall layout of the document it is going to be placed in. i.e., the composition within the photograph will have a relationship/balance with the rest of the document layout. This paper presents a novel image retrieval method, in which the document where the image is to be inserted is used as query. The algorithm calculates a balance measure between the document and each of the images in the collection, retrieving the ones that have a higher balance score. The image visual weight map, used in the balance calculation, has been successfully approximated by a new image quality map that takes into consideration sharpness, contrast and chroma.

* Internal Accession Date Only
Published in and presented at the IS&T/SPIE 19th Annual Symposium on Electronic Imaging, 28 January - 1 February 2007, San Jose, California, USA. Approved for External Publication

Content selection based on compositional image quality

Pere Obrador

Hewlett-Packard Laboratories
Palo Alto, CA, USA
pere.obrador@hp.com

ABSTRACT

Digital publishing workflows usually have the need for composition and balance within the document, where certain photographs will have to be chosen according to the overall layout of the document it is going to be placed in. i.e., the composition within the photograph will have a relationship/balance with the rest of the document layout.
This paper presents a novel image retrieval method, in which the document where the image is to be inserted is used as query. The algorithm calculates a balance measure between the document and each of the images in the collection, retrieving the ones that have a higher balance score. The image visual weight map, used in the balance calculation, has been successfully approximated by a new image quality map that takes into consideration sharpness, contrast and chroma.

Keywords: No reference image quality, sharpness metric, document layout, document balance, photographic composition

1. INTRODUCTION

Creating a document from scratch is a hard task for non-experts, and even professional graphic artists will probably use some help when indexing large image databases from which to select an image to insert in the document. Effective ways to cluster and retrieve sets of relevant images to ease the document creation process is the main motivation behind this paper. This is a continuation of the work started in[1] where the proposed tool retrieves images from a database that will color harmonize with the document where the user wants to insert it/them.
Document layout balance has been researched for some time now[2,3]. This paper presents an attempt to help the user in indexing large image databases, by analyzing the document composition, and selecting appropriate images based on some simple balance and symmetry rules.
The presented approach is not trying to look at the photograph composition in itself; it rather tries to extract some fundamental image features that will help determine whether that photograph will work with the rest of the document.
A no-reference objective image quality map is presented. This takes into account sharpness, contrast and chroma features, each of them calculated on a region by region basis, eventually assigning a quality number to each of the regions in the image.
The presented image quality map is then used to calculate the visual weight map of such a photograph, and plug the result into a document balance measure.

This paper presents a method to help in the area of document creation where the user starts with a document, and needs to select a photograph from a large image collection, i.e., given a layout, automatically retrieve a photograph from an image database that will balance that selected layout, or, retrieve the top N photographs that would best balance that layout for the user to choose from.

2. PROPOSED METHOD

2.1. DOCUMENT BALANCE

Document balance is a very important aesthetic feature that graphic artists work towards in their creative process. It can either be achieved by symmetrical balance, which gives a feeling of permanence and stability, or by asymmetrical balance which creates interest[2]. Two main ways of defining balance are[3]: centered balance, where the center of visual weight is at the visual center of a page, and left-right balance, where the weight of an object on the left side of the page is matched by the weight of an object at the same vertical position on the right side of the page.

The visual weight of an object is defined as its area times its optical density, and this is also for photographs and graphics.

In this paper, a novel way to calculate a photograph's visual weight is presented, which can be used to better balance the document, in which image analysis and no-reference quality assessment are used to generate a quality map from which extract the visual weight.

2.2. IMAGE QUALITY MAP

As mentioned above, the visual weight of an object is defined as its area times its optical density, and in the literature[3] this has been applied to images too.

When the graphic artist inserts an image into a document[13], the center of gravity towards which he is expecting the document viewer to look at has less to do with the image optical density than, rather, the image composition and which areas of the image are in high quality for the human eye to linger and examine. This is specially true in high quality photography and design magazines, billboards, even websites, where low depth of field images are heavily used.

An image quality map is presented below, where the areas with higher quality will be assumed to have higher visual weight when balancing such image with a document.

Image segmentation[4] is performed on the input image (see figure 2.b), so that each region is assigned a certain quality level[5]. The rational behind this is that if all the pixels in a region belong to the same object, it can be assumed, in a normal situation, that all these pixels will have similar objective quality. This will generate a map with multiple regions, each of them having its own quality.

The quality map is defined below as the combination of three different maps (sharpness, contrast and chroma) as shown in figure 1.

The focused saliency map[6] (FSM) is the first step towards the final sharpness map used in this approach. Most of the energy in the FSM corresponds to the object in focus, while a large amount of the energy of the out of focus region is removed efficiently. This provides for an excellent starting point to effectively segment the low depth of field images mentioned above, and it also does a quite good job at detecting different levels of sharpness (without being an accurate sharpness measure as explained below).

In recent subjective visual tests[7] on Just Noticeable Blurriness, it was found that the human visual system has different responses to blur /sharpness at different contrast levels. The resulting non-linear function of the contrast (figure 2.c) can be factored into the sharpness measure to be used, and in this approach it is applied to the FSM (figure 2.d). This results in a much more homogeneous FSM on the sharp object.

In order to separate the in focus regions from the out of focus regions, a bilateral filter is applied on the FSM (figure 2.e), and further morphologically filtered (figure 2.f) in order to generate an in focus mark map that can be intersected with the segmented image. For each of the resulting regions an approximate sharpness value can be derived from the contrast corrected FSM (figure 2.d). A more accurate sharpness map is obtained by combining the contrast corrected

FSM with a more accurate sharpness measure[8,9], where each of these additional measures are also corrected with the non-linear contrast function.

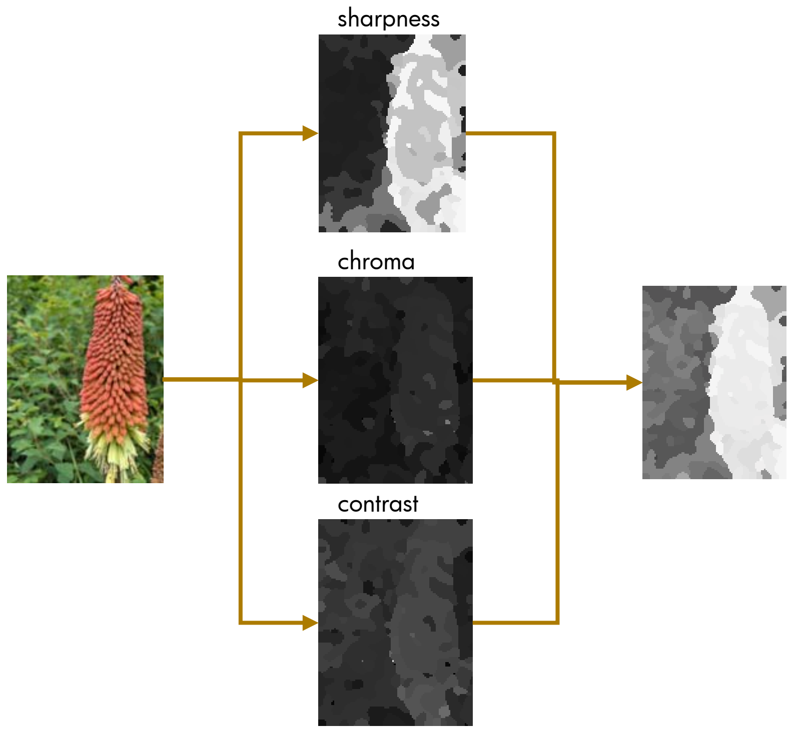


Figure 1. Combination of the sharpness map, contrast map and chroma map into the final quality map.

In low depth of field situations, the described sharpness map is enough for the presented approach to work well. In other situations some more assumptions need to be made in order to guess where the document viewer's attention will be set. Visual attention or saliency models are being used extensively in the literature[10], which do a good job at guessing exactly that. In this region based approach, though, it is not straightforward to extend the saliency models. Instead, a simpler local contrast measure is implemented, which assumes a certain image size and viewing distance. This is calculated on a region by region basis, and a contrast map (figure 1) is generated.

A third and final map is generated by calculating the local chroma content in each of the regions. This map is created after[11], where the relation between perceptual image quality and naturalness was investigated by varying the colorfulness and hue of color images of natural scenes, and concluding that human observers prefer more colorful images.

Both the contrast and chroma map complement the sharpness map, but never dominate it, i.e., if a region is out of focus, there is so much contrast and chroma can add to the final quality map (see figure 1).

2.3. DOCUMENT BALANCE BY MEANS OF THE QUALITY MAP

Balancing photographs and documents has been tried in[1], where color harmony was accomplished by looking at the relationship between image regions and document regions, both colorimetrically and spatially, i.e., region with harmonizing colors can be weighed higher if they are close together or opposing in the document, and hence balancing one another.

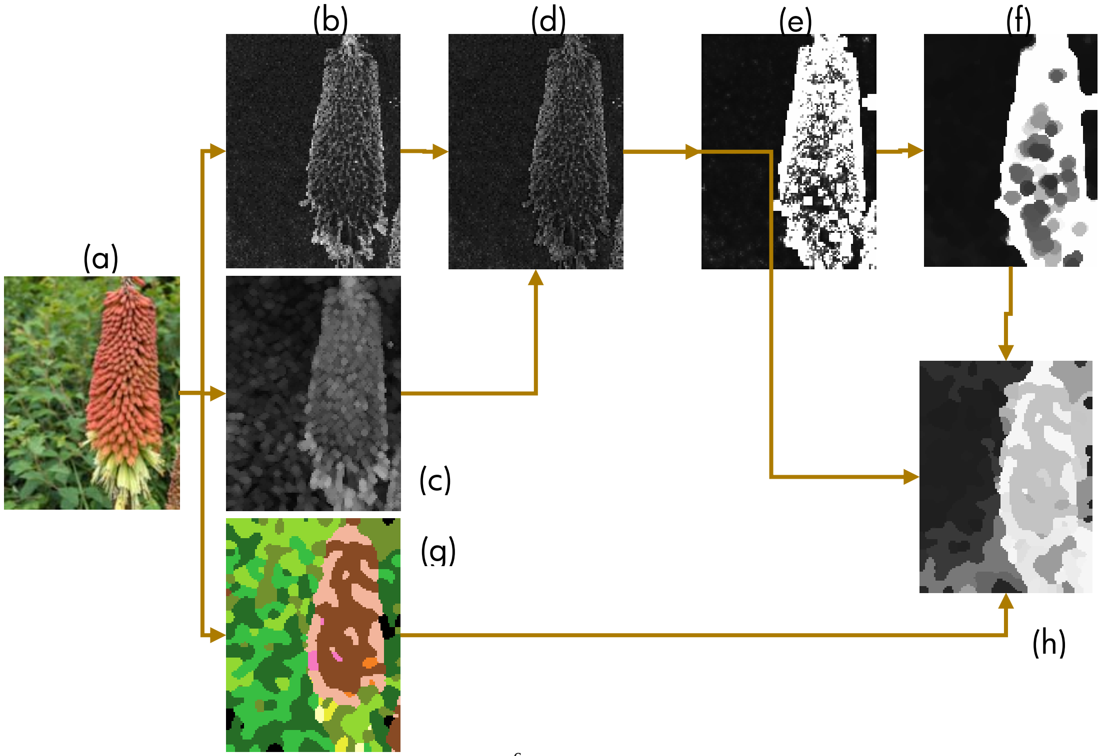


Figure 2. (a) original image, (b) FSM as described in[6] , (c) contrast map, (d) contrast corrected FSM, (e) bilateral filtering, (f) morphological processing, (g) region decomposition), (h) final sharpness map.

Color balance is not enough, as stated above, since the main object of interest (in high quality) is the area that should be weighed higher in the image visual weight map. The current approach combines both techniques when querying an image database, i.e., both the balance query and the color harmony query (see figure 6).

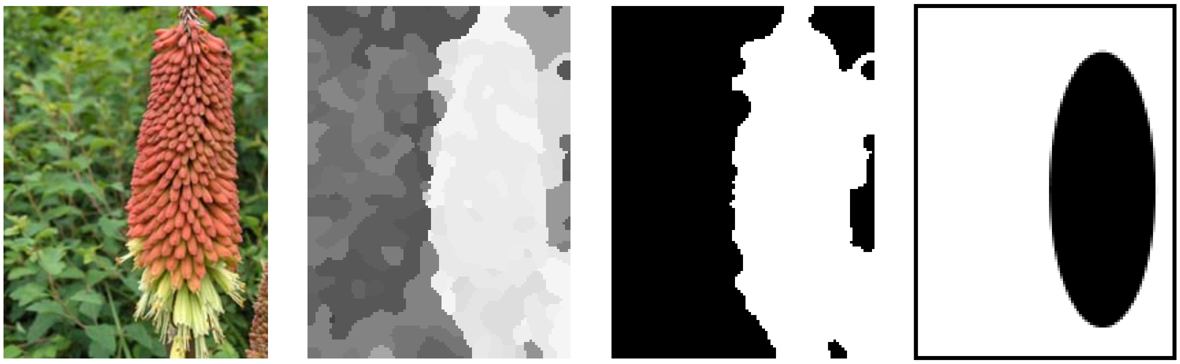
In order to perform the query, a simple model for the image quality map is presented below, which can be extracted and stored with each of the images as metadata for fast indexing.

2.3.1. Image quality map abstraction

Once the image quality map is obtained (see section 2.2 above), it has to be abstracted into an easy to use model for fast querying and retrieval. The current approach thresholds the quality map (see figure 3), and the resulting region/s are approximated by an ellipse.
Two different thresholds have been experimented with. First, a fixed threshold for all images, and second, an adaptive threshold to the image quality map content. Each has advantages and disadvantages.
In the case that the image collection is known to have only high quality pictures, having an adaptive threshold makes most sense, since it is known beforehand that at least there is a high quality region in each image. This takes care of some artistic soft focus photographs, and/or abstract photography.
In the case of a consumer photo collection, such assumption cannot be made since excellent photos coexist with very bad ones. Therefore a fixed threshold is necessary, meaning that those worse shots will never be retrieved in a balance query.

The resulting thresholded image is then approximated by an ellipse, (centroid plus spread or axes). Best results are obtained if a quality map weighted centroid is used.

Both centroids and spreads are expressed in percentage of width and height. This allows for querying both landscape and portrait photographs with little added complexity.



(a) (b) (c) (d) (e)

Figure 3, (a) original, (b) quality map, (c) thresholded quality map, (d) map abstract representation by an ellipse

2.3.2. Document visual weight map abstraction

In order to perform a fast query, a simple abstraction of the document's visual weight map is needed[12]. For this reason, the centroids and size in pixels for each of the objects in the document are calculated (see figure 4.b). All these are then combined to find the centroid for the whole document visual weight map, and its spread.

(a) (b) (c) (d) (e)

Figure 4. (a) original document, (b) extracted document visual weight map consisting of multiple document objects, (c) mirrored visual weight map in photo area, (d) retrieved image quality map from the collection, (e) retrieved image.

For all objects in the document visual weight map:

$$centroidX_{documentWeight} = 100\frac{\sum_j x_j M_j}{documentWidth\sum_j M_j}; \quad centroidY_{documentWeight} = 100\frac{\sum_j y_j M_j}{documentHeight\sum_j M_j}$$

Where (x_j,y_j) is the centroid coordinates for document object *j*, and M_j is the number of pixels belonging to this document object *j*. Again, a percentage of width and height are calculated to ease the query process (query both portrait and landscape photographs).

2.3.3. Balance measure between image and document

The hypothesis that was made is to approximate the presented image quality map as the image visual weight map. This quality map is then compared with the document's visual weight map, and a measure of balance between the two is calculated.

The image collection needs to be queried based on a certain balance criteria that the graphic designer or user need to specify. In this implementation only two criteria are possible:

1. Left Right symmetrical balance: implemented as a horizontal symmetry of the document visual weight map (see figure 4). In this case the centroid for such query would be: $centroidX_{query} = (100 - centroidX_{documentWeight});\quad centroidY_{query} = centroidY_{documentWeight}$
2. Centered symmetrical balance: implemented as a center symmetry of the document visual weight map: $centroidX_{query} = (100 - centroidX_{documentWeight});\quad centroidY_{query} = (100 - centroidY_{documentWeight})$

This modified query centroid (figure 4.c) is the one that will be compared with the image quality map centroid. The spread of the document visual quality map is not changed.

In the experiments it was seen that the balance was reduced roughly inversely proportional to the square of the centroid distance between the mirrored document weight map and the image quality map. The spread of the respective maps was less relevant, and the measure was set to be inversely proportional to its difference (see below).

Balance measure between image and document query:

$$balance\_measure_i = \frac{regionQuality_i}{1 + centroidDist_i^{\,2} + sigmaDist_i} \qquad \text{equation 1.}$$

Where regionQuality is an optional term, and is basically a 2D integral of the desired map the user wants to weigh into the equation, i.e., if the user wants to weigh chroma in the high quality region, this term would add up all the chroma values in the chroma map (see figure 1) in a region under the ellipse abstracting the mirrored document's visual weight map. See figure 8 for an image retrieval example using regionQuality, where the chroma has been factored in. If the regionQuality term needs to be used, then the whole quality map needs to be stored as metadata, increasing the needed storage size as well as computation time;

The other terms in the above formula are:

$$centroidDist_i = \sqrt{(centroidX_i - centroidX_{query})^2 + (centroidY_i - centroidY_{query})^2}$$

and

$$sigmaDist_i = \sqrt{(sigmaX_i - sigmaX_{query})^2 + (sigmaY_i - sigmaY_{query})^2}$$

Where ($centroidX_i$, $centroidY_i$) are the coordinates of the high quality ellipse weighted centroid for image *i*, and $sigmaX_i$ and $sigmaY_i$ are the spread of such ellipse for image *i*; and $sigmaX_{query}$ and $sigmaY_{query}$ are directly proportional to the spread of the document visual weight map.

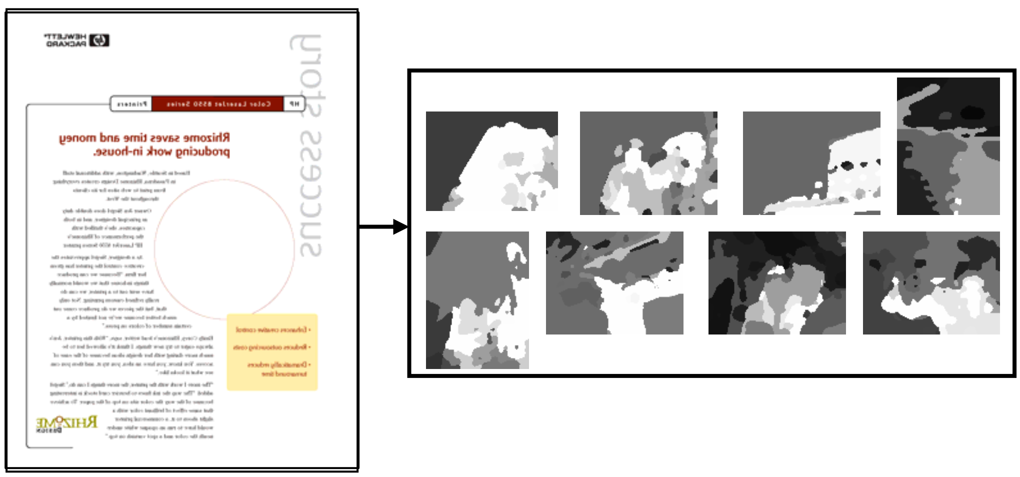
For a particular document, after abstraction, the centroid and spread of its visual weight map is used to index the image collection. The balance measure is calculated for every image in the collection, and either the image with the highest score, or the set of N images with the highest scores, are retrieved and presented to the user.

3. RESULTS

The experiments were performed on 850 personal images, i.e., consumer type images. For this reason, a fixed threshold was used for the quality map abstraction (section 2.3.1.).

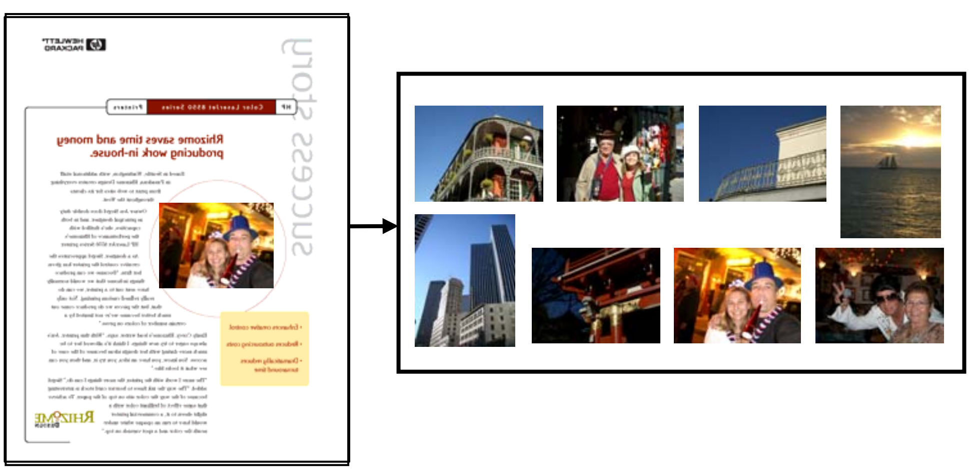
The results presented in this section were generated querying the collection with Left Right symmetrical balance: implemented as a horizontal symmetry of the document visual weight map.

Figure 5 presents the retrieved results using a certain document with its visual weight centered towards the left third. Notice the weight towards the right third of the image quality maps. Figure 6.b presents the actual top 8 retrieved images with such a document query. And figure 6.a presents the top retrieved image after applying the color harmonization query on the balance query results in figure 6.b. The algorithm managed to find pretty good result both for balance and color harmony within the relatively small 850 image collection.



(a) (b)

Figure 5. (a) document query, (b) top 8 quality maps retrieved from the collection



(a) (b)

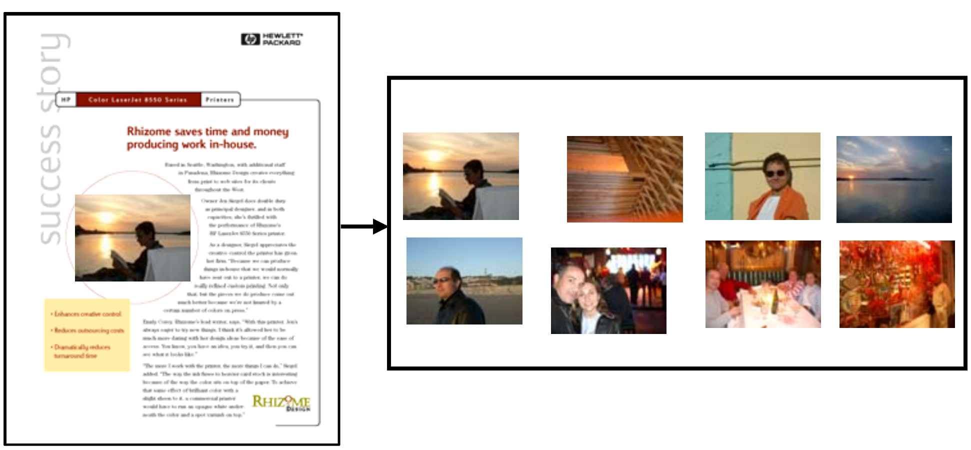
Figure 6. combination of balance and color harmony queries. (a) top retrieved image after performing the color harmony query[1] on the results of the balance query. (b) Top 8 retrieved images with the balance query.

Figure 7 presents a balanced document, where the photograph has to lie in the center of it. In this situation it is expected to retrieve images with a quality map with its centroid close to the center of the image.

(a) (b)

Figure 7. (a) top retrieved image with the balance query. (b) Top 8 retrieved images with the balance query.

Figure 8 presents the query results with a document balanced towards the right third. The regionQuality factor in equation 1 was used, and it integrated the chroma map for this specific example. The results should have high quality towards the left third, and also high chroma content around that left third.



(a) (b)

Figure 8. balance query with *regionQuality* performed on the chroma map. (a) top retrieved image with the chroma map weighted balance query. (b) Top 8 retrieved images with the chroma map weighted balance query .

These results confirm our hypothesis, and therefore it is safe to use the image quality map as the image visual weight map when performing document balance analysis.

4. CONCLUSIONS AND FUTURE WORK

A new image retrieval method has been presented, in which the document where the image is to be inserted is used as query. The algorithm calculates a balance measure between the document and each of the images in the collection, retrieving the ones that have a higher balance score. The image visual weight map, used in the balance calculation, has been successfully approximated by a new image quality map that takes into consideration sharpness, contrast and chroma. This retrieval method has been successfully combined with the color harmony retrieval method presented in[1].

Future work needs to be done in adding extra visual saliency features to the image visual weight map, which may solve some of the problems encountered with evenly sharp images.

The color version of this paper: please refer to the HP-Labs web-site (http://www.hpl.hp.com/techreports/).

REFERENCES

1. P. Obrador, "Automatic color scheme picker for document templates based on image analysis and dual problem", EI126, proceedings of Digital Publishing Conference, 2006, San Jose.
2. H. Balinsky, "Evaluating interface aesthetics: a measure of symmetry", HP-Labs Technical report
3. S.J Harrington, J.F. Naveda, R.P. Jones, P. Roetling and N. Thakkar, "Aesthetic measures for automated document layout", ACM Symposium on document engineering, 2004.
4. P. Obrador, "Multiresolution Color Patch Extraction", Electronic Imaging, VCIP 2006, San Jose.
5. P. Obrador, "Method and system for image quality calculation", Filed Patent Docket HP 200503391-1.
6. H. Li and K.N. Ngan, "Unsupervised segmentation of defocused video based on matting model", ICIP 2006, Atlanta.
7. R. Ferzli and L. J. Karam, "A Human Visual System-Based Model for Blur/Sharpness Perception," 2nd International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, 2006.
8. Shaked, D. and Tastl, I. , "Sharpness measure: towards automatic image enhancement", ICIP 2005. 11-14 Sept. 2005, Volume: 1, On page(s): I- 937-40
9. S.H. Lim, J. Yen and P. Wu, "Detection of ill-focused Digital Photographs," in Proceedings of the HP TECHCON 2004, Orlando, FL, June 2004
10. L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", IEEE transactions on pattern analysis and machine intelligence", vol. 20, no.11: 1254-1259.
11. H. de Ridder, F. J. J. Blommaert and E. A. Fedorovskaya, "Naturalness and image quality: chroma and hue variation in color images of natural scenes", Proceedings of SPIE -- Volume 2411 Human Vision, Visual Processing, and Digital Display VI, April 1995, pp. 51-61
12. Chao, H., Fan, J., "Layout and Content Extraction for PDF Documents." In proceeding of IAPR Int. workshop on Document Analysis System, 2004.
13. P. Bonnici and L. Proud, "Designing with photographs", Rotovision SA, 1998.