Representative Image Thumbnails: Automatic and Manual

Ramin Samadani, Tim Mauer, David Berfanger, Jim Clark and Brett Bausk

Media Technologies Laboratory
HP Laboratories Palo Alto
HPL-2008-6
January 24, 2008*

image resizing, image browsing, thumbnails, blur, noise, subjective studies

Image thumbnails are used in most imaging products and applications, where they allow quick preview of the content of the underlying high resolution images. The question: "How would you best represent a high resolution original image given a fixed number of thumbnail pixels?" is addressed using both automatically and manually generated thumbnails. Automatically generated thumbnails that preserve the image quality of the high resolution originals are first reviewed and subjectively evaluated. These thumbnails allow interactive identification of image quality, while simultaneously allowing the viewer's knowledge to select desired subject matter. Images containing textures are, however, difficult for the automatic algorithm. Textured images are further studied by using photo editing to manually generate representative thumbnails.

The automatic thumbnails are subjectively compared to standard (filter and subsample) thumbnails using clean, blurry, noisy, and textured images. Results using twenty subjects find the automatic thumbnails more representative of their originals for blurry images. In addition, as desired, there is little difference between the automatic and standard thumbnails for clean images. The noise component improves the results for noisy images, but degrades the results for textured images. Further studying textured images, the manual thumbnails were subjectively compared to standard thumbnails for four images. Evaluation using forty judgments found a bimodal distribution for preference between the standard and the manual thumbnails, with some observers preferring manual thumbnails and others preferring standard thumbnails.

Internal Accession Date Only Approved for External Publication

To be published at Human Vision & Electronic Imaging, Jan 2008, San Jose, CA

Representative Image Thumbnails: Automatic and Manual

Ramin Samadani[1], Tim Mauer[2], David Berfanger[2], Jim Clark[2] and Brett Bausk[3]
1) HP Labs (Ramin.Samadani@hp.com), 2) HP Rainbow Image Science and 3) hp.com

ABSTRACT

Image thumbnails are used in most imaging products and applications, where they allow quick preview of the content of the underlying high resolution images. The question: "How would you best represent a high resolution original image given a fixed number of thumbnail pixels?" is addressed using both automatically and manually generated thumbnails. Automatically generated thumbnails that preserve the image quality of the high resolution originals are first reviewed and subjectively evaluated. These thumbnails allow interactive identification of image quality, while simultaneously allowing the viewer's knowledge to select desired subject matter. Images containing textures are, however, difficult for the automatic algorithm. Textured images are further studied by using photo editing to manually generate representative thumbnails.

The automatic thumbnails are subjectively compared to standard (filter and subsample) thumbnails using clean, blurry, noisy, and textured images. Results using twenty subjects find the automatic thumbnails more representative of their originals for blurry images. In addition, as desired, there is little difference between the automatic and standard thumbnails for clean images. The noise component improves the results for noisy images, but degrades the results for textured images. Further studying textured images, the manual thumbnails were subjectively compared to standard thumbnails for four images. Evaluation using forty judgments found a bimodal distribution for preference between the standard and the manual thumbnails, with some observers preferring manual thumbnails and others preferring standard thumbnails.

1. INTRODUCTION

For many years, standard thumbnails have been generated by filtering and subsampling their high resolution originals. [1] Although this process prevents aliasing and preserves the image composition, it loses information about image quality and high spatial frequency textures.[2] Several recent papers offer interesting, non-traditional approaches to reducing image size.[3–6] Even though the algorithms differ, they each modify image composition while resizing, generating better thumbnails. The papers address images with different types of content such as web pages,[3] documents[5] and photographs.[4,6]

Automatic thumbnails that we previously developed[2] do not modify the image composition, but rather better reflect the image quality of the originals. Current, standard thumbnails do not distinguish between high and low quality originals. Both sharp and blurry originals appear sharp in the thumbnails, and both clean and noisy originals appear clean in the thumbnails. This leads to errors and inefficiencies during image selection. Our new thumbnails provide a quick, natural way for users to identify images of good quality, while allowing the viewer's knowledge to select desired subject matter. Figure 1 shows examples of the automatic thumbnail results, where the originals are cropped in order to be shown at the correct scale. These images are not tuned for the print process and they are best viewed on a display. A benefit is that the new thumbnails are natural to interpret; there is no learning necessary to understand the blur and noise in the new thumbnails. The alternate approach of automatic image ranking by quality[7] is extremely difficult because the knowledge about the subjects of interest resides with the user, not with the image. On the other hand, with the new thumbnails the user can quickly check, for example, whether the subject of interest is in focus.

Section 2 reviews the algorithm that automatically generates new, quality-preserving thumbnails and describes a noise generating component that is much faster than the original.[2] Section 3 describes the subjective evaluation of the automatic thumbnails and Section 4 discusses the findings of the subjective evaluation.

Difficulties encountered with images containing textures prompted a basic question: *What is the most representative thumbnail one can generate given a fixed number of thumbnail pixels?* Section 5 does not discuss algorithms for this difficult problem, but rather presents experiments in manual generation of image thumbnails using the popular Adobe ® Photoshop® photo editor. To learn how to generate thumbnails for textured images, the experiments use images containing textures with spatial frequencies higher than the thumbnail sampling frequency. Sections 6 and 7 describe the steps used to generate the manual thumbnails for two images containing textures. Unlike standard thumbnails, the manual approach to thumbnail generation implicitly takes into account perceptual and cognitive aspects. Section 8 distills common themes from the manual thumbnail experiments that may in the future provide guidelines for developing new algorithms for generating representative thumbnails for textured images. Section 9 describes the results of a subjective study comparing the standard and manual thumbnails. Section 10 offers conclusions about the automatic and manual thumbnails.

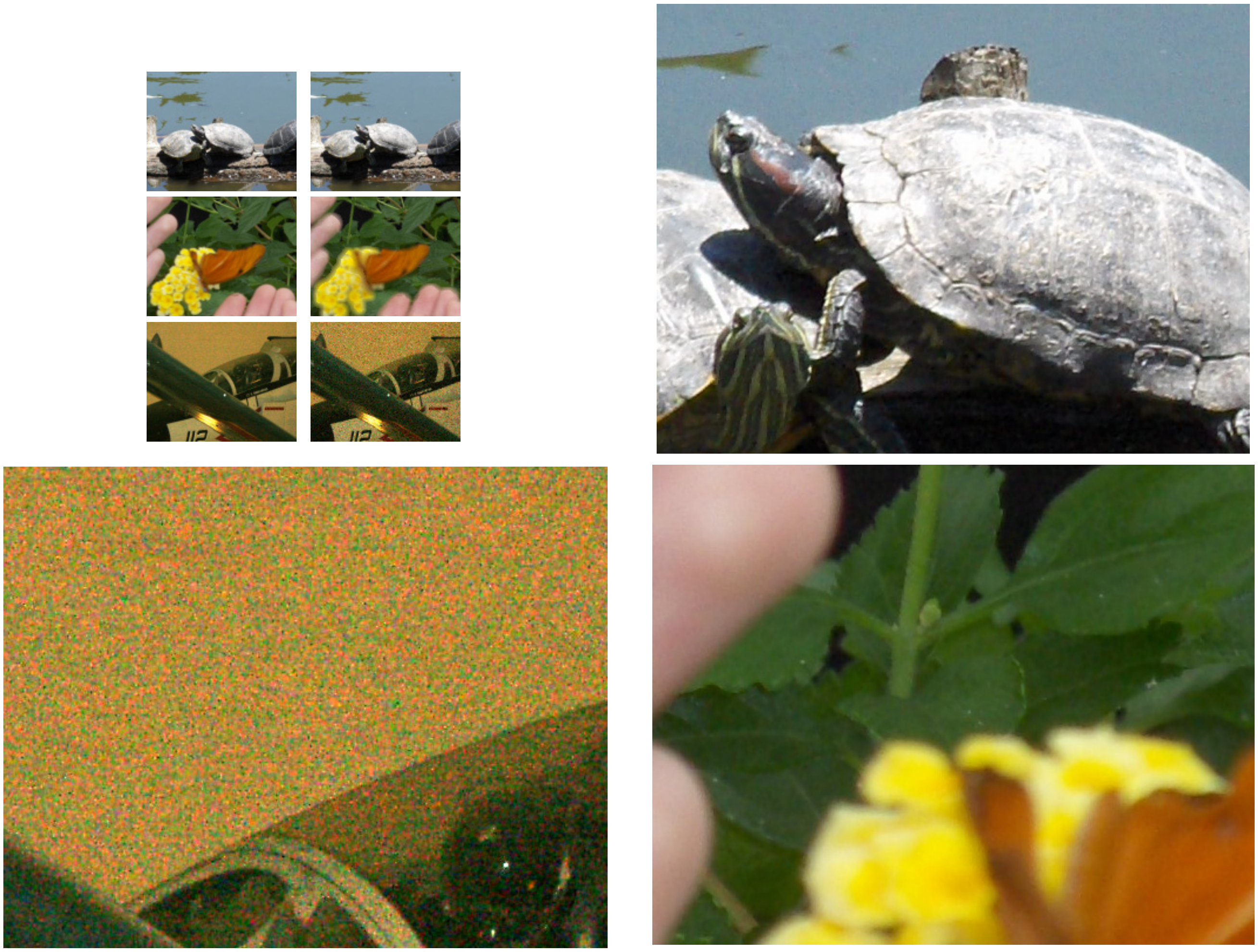

Figure 1. Standard thumbnails (left column) and automatically generated, new thumbnails (right column) for three original images (cropped to fit at the correct scale, shown surrounding the thumbnails). The images in this paper are best viewed on a display.

2. ALGORITHM FOR AUTOMATICALLY GENERATED REPRESENTATIVE THUMBNAILS

For simplicity of notation, images are considered column stacked vectors. [1] The image model used is

$$\boldsymbol{d} = B\boldsymbol{c} + \boldsymbol{n}. \tag{1}$$

In this equation, the vector $\boldsymbol{c}$ represents an ideal image captured with infinite depth of field. The matrix B represents, in general, a space-varying blur corresponding to unintended distortions such as camera shake, motion blur or misfocus, and $\boldsymbol{n}$ represents an additive, perhaps correlated, noise. Well taken digital photographs will not have unintended distortions. In this case, the noise $\boldsymbol{n} = 0$. But the matrix B may not be the identity, since it still represents the space-varying depth of field blur. In the special case of infinite depth-of-field, $B = I$, and therefore $\boldsymbol{d} = \boldsymbol{c}$.

This work takes advantage of prior work on the very difficult problems of image denoising [8] and blind deconvolution,[9] where the goal is to recover, in Equation 1, $\boldsymbol{c}$ from $\boldsymbol{d}$. The goal of this work, however, is to generate a low resolution thumbnail $\boldsymbol{t}_n$, not the exact reconstruction of high resolution $\boldsymbol{c}$. This changes the requirements of our component algorithms. For example, our solution works well with both shake and defocus blurs, by applying an appropriate space-varying

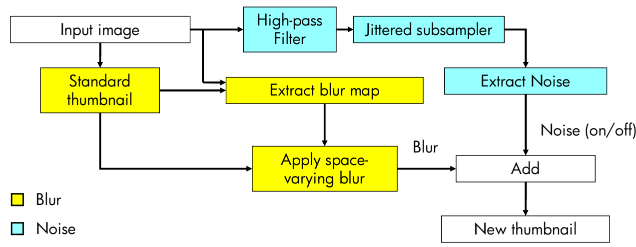

Figure 2. To generate the automatic thumbnail, the standard thumbnail is modified with a space-varying blur and an estimated noise is optionally added.

Gaussian blur. The details of the applied blur kernel is not critical to our results. Similarly for noise level, we do not need an extremely accurate estimate, but rather a rough, fast one may be sufficient.

The new thumbnail, $\boldsymbol{t_n}$, is generated by starting first with the standard thumbnail, $\boldsymbol{t}_s$, which is clean even for distorted input images. To this standard thumbnail, blur and noise are applied to correspond to the blur and noise in the original,

$$\boldsymbol{t}_n = \boldsymbol{t}_b + \boldsymbol{n}_t = B_t \boldsymbol{t}_s + \boldsymbol{n}_t. \tag{2}$$

Figure 2 shows the automatic thumbnail algorithm. To avoid detection errors, we wish the process to be the least as possible *detection-based*. For example, global blur detection is avoided, but rather a pixel by pixel blur estimation is applied. The resulting space-varying blur can represent, among other possibilities, locally varying depth-of-field. The blur portion of the algorithm has a single gain parameter, ρ. In order to be fast, as little processing as possible is done at the high resolution image. Most of the noise calculations described in this paper occur at the low thumbnail resolution. In the resulting algorithm, there is a single noise gain parameter. In addition, the noise calculation can be turned on or off.

The blur component computation remains the same as previously described.[2] The *Extract blur map* block results in a two dimensional thumbnail resolution blur map with estimates of the amount of blur at each thumbnail pixel and the block *Apply space-varying filter* applies a filter based on this blur map. This local computation accounts for depth-of-field as well as other undesired blurs. The blur map is determined without identification of the type of blur. The assumption is that users will not be able to easily distinguish between different types of blur in the thumbnail.

The local amount of blur is computed by noting that the image edge profiles differ between sharp and blurry images. At an edge, for example, the profile of a blurry high resolution image will be more gradual than its corresponding low resolution standard thumbnail, $\boldsymbol{t}_s$, whose profile will be steeper.[10] Applying successively larger blurs to $\boldsymbol{t}_s$ will cause its edge profile to become more gradual, and to correspond better to the blurry original. To have the system work with various image features, and not just edges, the computation is based on *pixel range* (difference between maximum and minimum pixel values in a spatial neighborhood) to determine the local image profiles.

Subsequently, a new algorithm for the noise generation component was developed with most of the processing operations occurring at thumbnail resolution. This is accomplished by exchanging the order of subsampling and noise generation so that the subsampling occurs first, as shown in Figure 2. Adding random jitter to the subsampling breaks up potential moire from any residual image textures that may erroneously appear in the noise image. The noise residual from a modified wavelet based soft thresholding denoiser,[11] known as *VisuShrink*, is used. Even though the *Extract noise* block involves non-linear processing, multirate signal identities,[12] and approximations for the VisuShrink variance estimator, allow exchange of the order of subsampling and noise extraction. The computation time, on a laptop with an Intel T2500 2GHz processor with 1 GB of RAM, for converting an original 1280x1024 image to a 128x102 thumbnail now takes .14 second instead of the original 2.25 seconds.

3. SUBJECTIVE EVALUATION OF AUTOMATICALLY GENERATED THUMBNAILS

The algorithm was first tested informally with several hundred images and it was found effective for blurry and noisy images. There were, however, differences between the standard and new thumbnails for textured images. By turning off the noise processing, corresponding to term $\boldsymbol{n}_t$ in Equation 2, the noise term was found to account for the differences in the textured images (as well as the noisy images). This is due to the currently used noise algorithm, which does not always

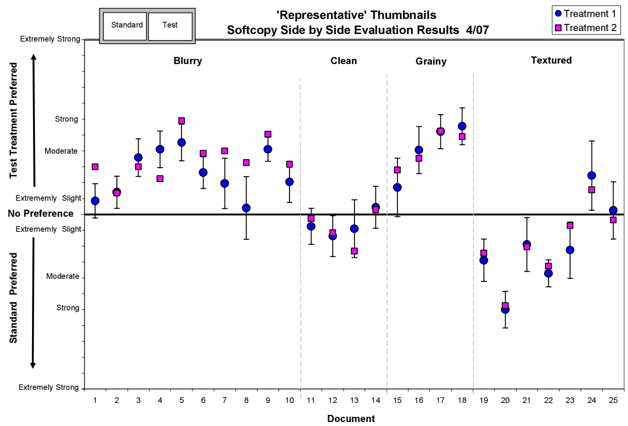

Figure 3. Plots of results for 96x96 pixel thumbnails for 25 images

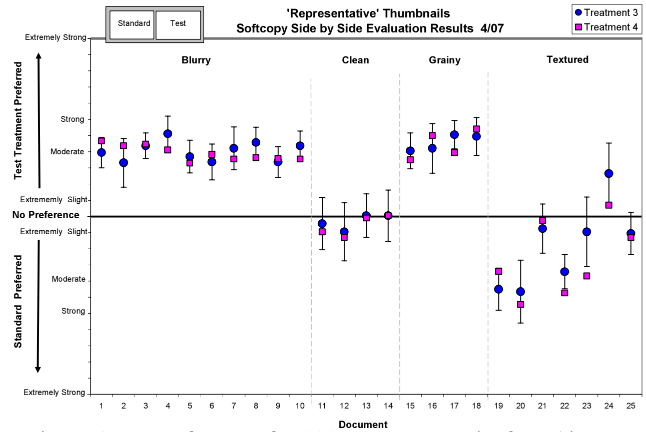

Figure 4. Plots of results for 300x300 thumbnails for 25 images

distinguish between image noise and texture, both of which contain high spatial frequencies. These findings identified four categories of input images, *Blurry*, *Clean*, *Grainy* (as used here, the term grainy is equivalent to noisy) and *Textured* for further study. Sections 3.1 and 3.2 discuss a softcopy subjective evaluation conducted using these four image categories.

3.1 Evaluation Method

The evaluation consisted of comparing a standard thumbnail vs. a candidate treatment thumbnail for best representation of the original image. Two thumbnail sizes were included, one with 96x96 pixel bounding box and one with 300x300 pixel bounding box. Two thumbnail treatments, corresponding to different blur parameters ρ, were tested for each size. For the 96x96 thumbnails, the blur gain values used were $\rho = .85$ for *treatment 1* and $\rho = 1.25$ for *treatment 2*. For the 300x300 thumbnails, the values used were $\rho = 1.25$ for *treatment 3*, and $\rho = 1.5$ for *treatment 4*. The image suite consisted of a total of twenty five photos, divided into the four categories: blurry, clean, grainy, and textured. A total of one hundred pair were judged in a softcopy evaluation. The judges were asked to determine which thumbnail version of a pair best represented the original full image. Twenty judges participated.

The equipment and experimental conditions are now described. The monitors used are HP L2335 Active Matrix TFT's (thin film transistor), which have a 23 inch diagonal viewing screen length and a native resolution of 1920 x 1200 pixels and a pixel pitch of 0.258mm. The monitors were calibrated just prior to the testing using the Monaco Optix 2.0 software and sensor produced by X-Rite. The overhead room lighting was turned off for this evaluation. There were a few task lamps turned on across the room to provide enough light for the judges to see to walk safely to the test cubicle where the softcopy monitors are stationed.

The observers were selected from a pool of HP employees at the Vancouver site that have satisfactorily completed color vision discrimination testing and observer orientation training. They are experienced at evaluating image quality primarily because of involvement in the evaluations that are conducted weekly. They are not considered *experts*, but are believed to predict the general consumer response. Periodic external evaluations have validated the calibration of the HP internal testing with real customers. The observer pool is diverse in gender and age. Special care is taken to keep the observers unwitting with regards to the source of the samples and the objectives of the test. Image scientists, color scientists and other imaging professionals are excluded from the observer pool.

Thumbnail treatments were positioned to appear on the right and left side randomly. Image samples were presented to each observer in a different random order. This technique distributes any *start-up* or *fatigue* effects over different samples. Observers were presented with a full version view of an image in softcopy and then two candidate thumbnail versions side by side. They were asked to indicate which thumbnail version was the most representative of the full-size original and indicate the degree of their preference using the scale provided. They were able to toggle the screen back and forth between the full image and the thumbnails. After recording their response, the next image set would load and the observer would proceed until all 100 samples were evaluated.

3.2 Results

Graphs of the results for the 96x96 thumbnails are shown in Figure 3 and the results for the 300x300 thumbnails are shown in Figure 4. The data on each graph are grouped by image category. The results are now described by image category.

For the *blurry* images, for the small thumbnails (96x96) treatment 2 produced slightly better representations of the original than did treatment 1, and both treatments were better than the standard. For the large thumbnails (300x300) there was no difference between treatment 3 and 4 performance, both were significantly better than the standard, and the result was consistent across the document suite and on average better than the 96x96 results. For the *clean* images there was no difference in the treatment performance for either size, and no significant preference between standard and treated thumbnails (that is to say, the treatments didn't *break* the clean images). For the *grainy* images no difference in performance between the thumbnail treatments for either size, and for both sizes the treated versions were more representative than the standard thumbnails. For the *Textured* images the treatment applied to the thumbnails in most cases was not as good a representation of the original as the standard thumbnail. This was true for both sizes of thumbnails. In general, the treatment added *speckle*. In the case of the bird image with a screen door in the background, shown in Figure 5, and also shown as Document 20, the second textured plot in Figure 3, significant and non-representative distortion was visible. Treatment 4 on the 300x300 size was slightly worse than Treatment 3.

Overall, the treatments appear to work well for blurry and grainy images, do not harm clean images, and do harm the textured images. For the blurry images, testing new values of ρ may provide for *better* representation of the original than the treatments tested provided.

4. DISCUSSION OF EVALUATION OF AUTOMATICALLY GENERATED THUMBNAILS

The results of the study, even without careful algorithm parameter tuning, are encouraging. At both resolutions, users find that the new thumbnails: 1) better represent the blurry and noisy images; and 2) are not significantly different than standard thumbnails for the clean images. For most of the textured images, however, the users prefer the standard thumbnails.

Figure 1 shows thumbnail comparisons and originals for three examples used in the tests, including ones with noise and blur. The original image on the top right shows an example where the new thumbnails and standard thumbnails are indistinguishable for a good quality image. In the image on the bottom right, the hands, yellow flowers and red butterfly are out-of-focus, as is seen in the new thumbnail, but not the standard thumbnail. The image on the bottom left is noisy, as seen in the new thumbnail, but not the standard thumbnail.

Figure 5 shows two examples of textured images, with high spatial frequency content. The current noise component of our thumbnail algorithm does not distinguish well between noise and texture. For example, in the top right original image of Figure 5, the screen door in the background has repetition frequency above Nyquist. The screen door disappears (at 96x96 bounding box), as expected, using the standard thumbnail shown in the top left of Figure 5. With the new thumbnail, however, the screen door appears as a patterned noise. Neither thumbnail, in this case, accurately represents the original but for this image users prefer the standard thumbnail. This image corresponds to document 20, the second textured plot in Figure 3.

For most of the textured images, including ones containing spraying water and sand, users preferred the standard thumbnail. The textured image original for which users preferred the new thumbnail is shown in the bottom left of Figure 5. For this image, it appears that the new thumbnail better represents the carpet texture whereas the texture is not apparent in the standard thumbnail. The subjective results for this image correspond to document 24, the sixth textured plot shown in Figure 3.

5. MANUALLY GENERATED REPRESENTATIVE THUMBNAILS

The difficulties encountered with images containing textures prompted experiments to manually generate representative thumbnails for the class of images containing textures. A fixed number of thumbnail rows and columns were specified. Then, four images were selected, and given the fixed thumbnail dimensions, Adobe ® Photoshop® was used to generate a good representation of the high resolution originals. Due to space limitations, sections 6 and 7 describe the steps for manual thumbnail generation for only two of the images. The remaining two textured images and another, picture-in-picture example, are described in HP Labs technical report HPL-2007-169.[13] Image textures may be regular or stochastic[14,15] and one example from each category is presented here.

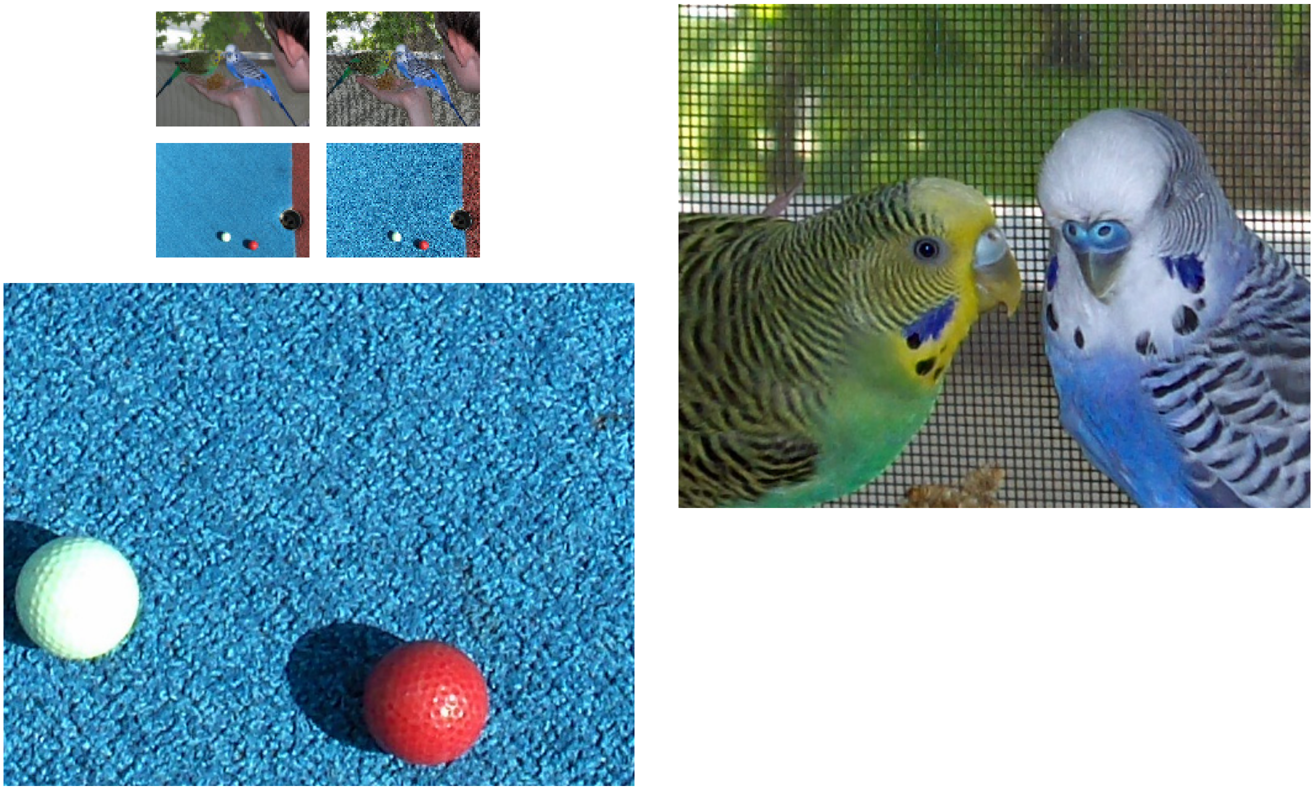

Figure 5. Standard thumbnails (left column) and automatically generated thumbnails (right column) for two *textured* images (cropped to fit at the correct scale, shown surrounding the thumbnails).

The image in Figure 6 shows birds in front of a screen door. The regular texture of the screen door has spatial frequencies higher than the thumbnail sampling frequency. The image in Figure 8 shows girls playing on a beach. This image was selected because of its complex, stochastic sand texture. This image also contains coarser, more complex textures such as the hand-print seen towards the top left. For these two images, the following sections describe the steps used to manually generate the thumbnails. Then, comparisons are made between the manual thumbnails and the standard (filter and subsample) computer-generated thumbnails. A thumbnail size of 128 pixels wide and 96 pixels high was selected. The original images are shown at about 70% of their correct size in order to fit this document. The appearance of these resized images is similar to the images viewed on a display at full scale.

6. *BIRDS* EXAMPLE

Figure 6 shows the *birds* image, where the screen door, an example of a regular texture, is difficult to represent at thumbnail resolution. The screen door cannot be displayed with accurate scale in the thumbnail since the screen repetition is at a smaller interval than a single thumbnail pixel.

6.1 Process used to generate manual thumbnail

The manual thumbnail for this image, and for the other examples processed, was generated by first creating a modified high resolution image (for the image in Figure 6, this modified image is shown in Figure 7) from which a representative thumbnail was generated by filtering and subsampling.

During creation of the manual thumbnail, the standard thumbnail shown on the top right of Figure 6 was first examined. In comparing the standard thumbnail to the full-size original image there were two noticeable differences that were immediately evident: 1) the screen pattern in the original was entirely absent from the standard thumbnail; and 2) The emphasis on the primary subjects, the two birds, was diminished. In order to regain the lost details, the elements of the image were separated into layers in Photoshop. The screen texture was enlarged so that it would represent a perceptually similar texture

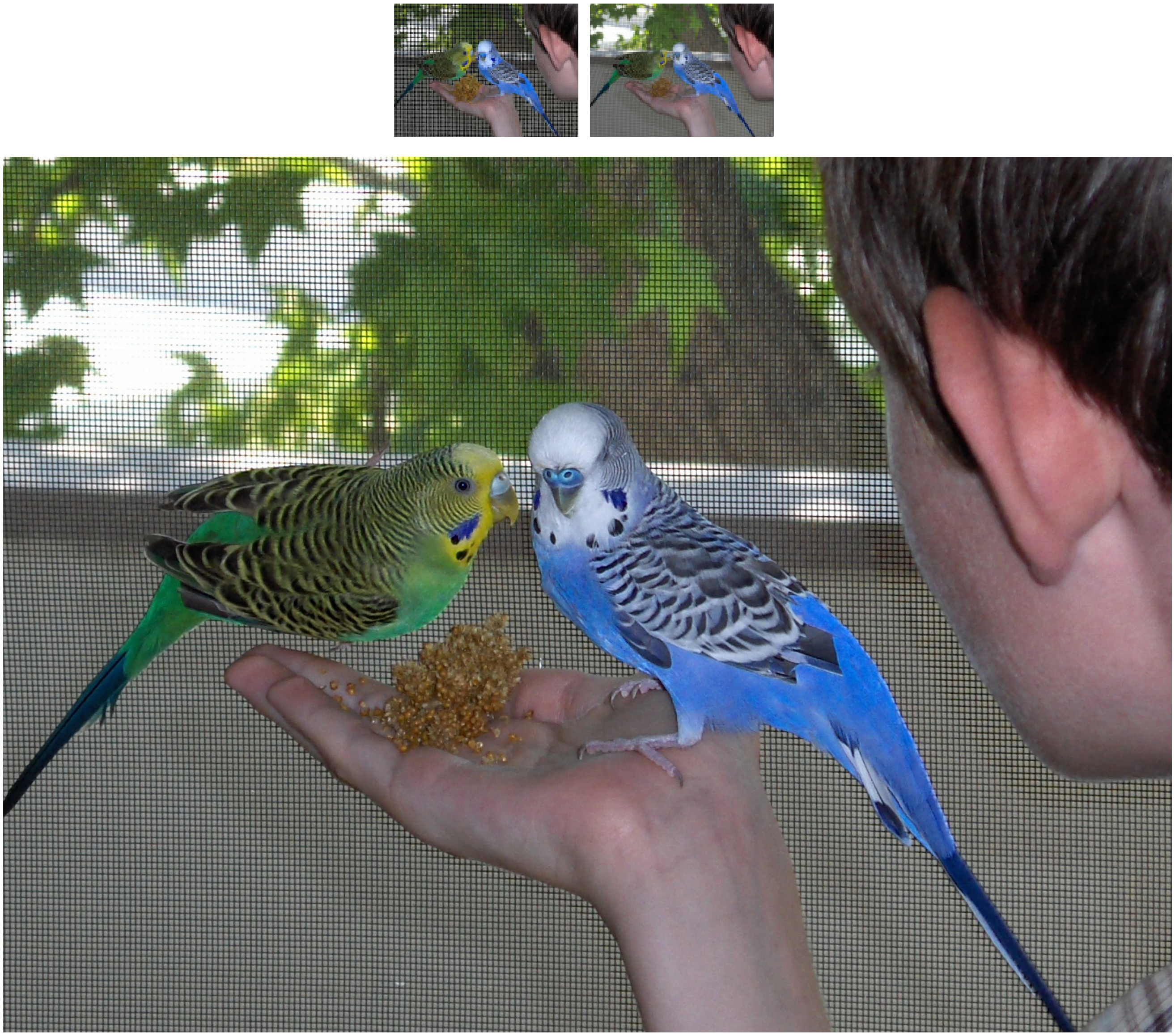

Figure 6. Manual thumbnail shown on the top left, and standard thumbnail shown on the top right for the original shown on the bottom. The original is shown at about 70% scale.

when resized to the thumbnail scale. The saturation of the colors of the birds and the bird feed was increased to compensate for the loss of saturation in the standard thumbnail. The birds' heads and the birds' feed were enlarged while maintaining the distance between the birds.

6.2 Comparison with standard thumbnails

The thumbnail on the top right of Figure 6 is the standard thumbnail generated by filtering and subsampling. The screen door is not preserved in the standard thumbnail since its spatial frequency is higher than the thumbnail sampling frequency, and it is removed by the antialiasing filter.[1] On the other hand, the manually generated thumbnail on the top left of Figure 6 shows the repetitive structure of the screen door at a larger scale than the texture would appear if properly scaled.

In Figure 6, the birds are the subject of interest. In addition to the enlargement of the screen, the birds were also

Figure 7. The modified original of Figure 6 from which the manual thumbnail was generated. Shown at about 70% scale.

enlarged in the manual thumbnail. This enlargement is very noticeable in Figure 7, the *modified* high-resolution image from which the manual thumbnail was generated. It is also detectable in the side-by-side comparison of the thumbnails in Figure 6. In a real application only one thumbnail would be used to represent the original. In a way, the enlargement attempts to preserve the saliency, or attention paid to the birds when viewing the original image.

7. *SAND* EXAMPLE

The *sand* image in Figure 8 shows children playing on a beach. In this case, the sand is the stochastic high spatial frequency texture that we would like to represent. Also appearing in the sand are coarser, more regular structures such as the hand-print that make this an interesting, complex test case.

7.1 Process used to generate manual thumbnail

The manual thumbnail for this example was created by identifying a recognizable pattern in the sand. This was enlarged to a scale that was visible when resampled to thumbnail resolution. In addition, the contrast and saturation were increased so that the downsampled thumbnail would have a similar aesthetic to the original image. One characteristic that we desired to reproduce, but to date have had no success with, was the sparkle in the girl's hair. This was laborious and difficult with the tools available in Photoshop. The manual thumbnail does not reflect the attempts to preserve the sparkle in the hair.

Figure 8. Manual thumbnail shown on the top left, and standard thumbnail shown on the top right for the original shown on the bottom. The original is shown at about 70% scale.

7.2 Comparison with standard thumbnails

There are noticeable differences between the standard thumbnail shown on the top right of Figure 8 and the manually generated thumbnail shown on the top left of the figure. The differences are in the representation of the sand. The sand shown in the manual thumbnail appears rougher in texture while the sand texture is smoother in the standard thumbnail. Also, clearly recognizable in the new thumbnail is the hand-print seen near the top left of the thumbnail. This hand-print appears at the correct scale in the standard thumbnail shown on the top right, and at this small size, it is not noticeable. Figure 9 shows the modified high-resolution image from which the manual thumbnail was generated. It is clear from this image, that the contrast was increased and the size of the sand was enlarged in order to give a more representative appearance when filtered and subsampled to form the manual thumbnail.

Figure 9. The modified original of Figure 8 from which the manual thumbnail was generated. Shown at about 70% scale.

8. COMMON FINDINGS FOR MANUALLY GENERATED THUMBNAILS

This section summarizes the common findings from the experiments in manually generating representative thumbnails. Both the techniques used, and the effects of image content on the manual thumbnail generation are described. The lessons learned from the manual thumbnails may influence the future development of automatic algorithms for image resizing.

Traditional filtering and subsampling cannot represent image textures with higher spatial frequencies than the thumbnail sampling frequency. No method can preserve the accuracy of both the local texture structure and its scale. In the manually generated thumbnails, a common theme was to relax the scaling accuracy in order to preserve the local texture structure. The textures were first enlarged at the original image resolution in order to be subsequently visible in the thumbnails. The inaccuracies in scale may not matter as much as preserving the texture characteristics. For example, the thumbnail on the top left of Figure 6 may show the *appearance* of the screen even though it is depicted at an inaccurate scale.

Many images have a clear subject of interest that would be agreed upon by most viewers of the image. For these images, depending on the size of the subject of interest (and the size of the thumbnail), the subject may not be as salient in the thumbnail as it is in the original. A subtle enlargement of the subject attempts to *match* the saliency of the subject in the thumbnail with the saliency of the subject in the original. This is seen most clearly in the processed *bird* image shown in Figure 7 but it is also seen in the hand-print shown in the processed *sand* image shown in Figure 9.

There are also images, particularly of dark indoor scenes [13] with fine details where the contrast of the details is lost in

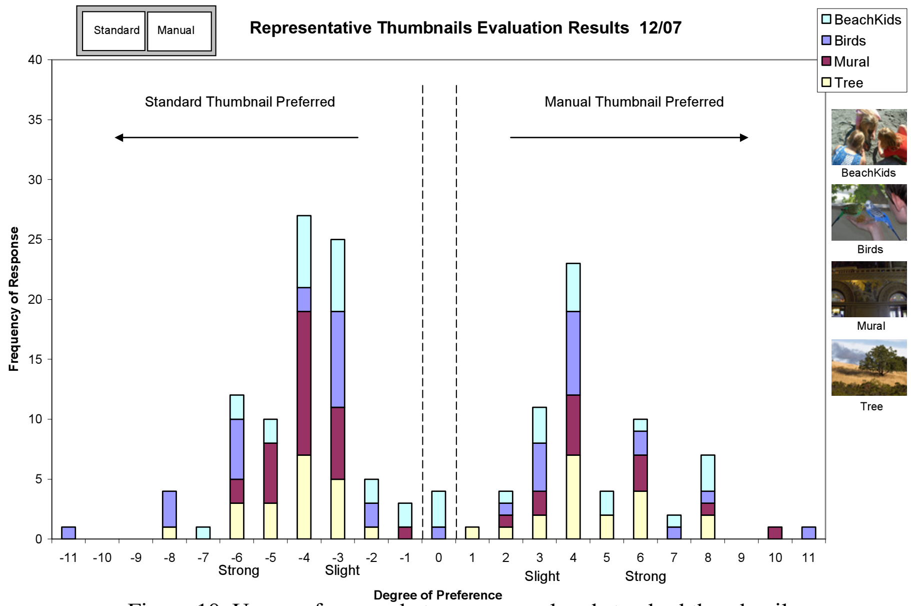

Figure 10. User preferences between manual and standard thumbnails

the thumbnails. It is possible that the perceived loss of the details is due to complex factors including the change in human contrast sensitivity with spatial frequency[16] as well as the lighter surrounds of the thumbnails. In this case, compensating for the change in contrast may improve the perceived results for the manually generated thumbnails.

In summary, the techniques used to generate the thumbnails included pixel level adjustments to increase contrast, texture scaling to enlarge textures before downsampling to thumbnail resolution, saturation increase, and subject size increases using image warping. Some of the techniques used seem to adjust for known perceptual contrast sensitivity changes with spatial frequency. For example, Photoshop level adjustments and the saturation increases, may be consistent with required changes due to contrast sensitivity. Other techniques, such as the enlarging of the subjects or the textures seem related to more complex cognitive effects such as the saliency of subjects at different resolutions.

Which techniques are used should depend on the user's familiarity with the collection. Every user needs to be able to judge the content of the image, including *what* or *who* the subject is, color, clarity, composition, exposure, texture, contrast and many other factors. However, a user familiar with a collection will use the thumbnails as reminders of particular images to use for their current tasks. For these applications, spatial or scale distortions in the thumbnail may be undesirable.

9. SUBJECTIVE EVALUATION OF MANUALLY GENERATED THUMBNAILS

The methodology described in Section 3.1 was also used for pairwise comparisons between manual versus automatic, manual versus standard and standard versus automatic thumbnails for textured images. The user instructions were: "Please indicate which thumbnail version is the most representative of the full-size original and indicate the degree of your preference using the scale provided." Forty judgments were made for each of four images.

Comparison of manual thumbnails with automatic thumbnails found preference for the manual thumbnails. There was no preference, on average, between manual and standard thumbnails. The histograms of observer preference between manual thumbnails and standard thumbnails is shown in Figure 10. The figure shows the results are bimodal. The observers almost never selected a neutral preference between the thumbnails. For a given image, some observers preferred the manual thumbnails and others preferred the standard thumbnails.

The study was conducted in two sessions. The observers in the second session were also instructed to "Please provide a comment indicating why you chose one thumbnail over the other." For the *birds* image, comments about the screen background were made both by observers that preferred the manual thumbnail and by observers that preferred the standard thumbnail. Observers that preferred the manual thumbnail found the screen a better representation of the high resolution original. Observers that preferred the standard thumbnail commented that the screen shown in the manual thumbnail seemed too *strong*. For the *sands* image, many comments were about the sand textures. Some users preferred the standard thumbnail texture representation and others preferred the manual thumbnail textures. It would be interesting to conduct further tests with different texture contrast levels in the manual thumbnails to see if this modifies user preferences.

10. CONCLUSIONS

The subjective evaluation of the automatic thumbnails shows that the blur component of the algorithm is robust and it may always be turned on with improved results. The noise component of the algorithm, however, improves the noise images but degrades the textured images. The decision to use the noise component requires further testing with the expected image mix for the particular application. For the subjective evaluation described here, roughly equal numbers of the different image categories were used to better test the different cases. How often noisy images and textured images occur for a given application may help determine whether the noise component should be used. A better noise generator that better separates between noise and texture, may allow the noise component to always be turned on without degrading textured images.

The manual thumbnails for textured images show mixed results for preference of the manual versus standard thumbnails. The observers seem to base preference on the strength of texture depicted in the manual thumbnails. Future studies may determine if there is a texture contrast setting that makes the manual thumbnails preferable to the standard ones.

Insights into different aspects of human perception, such as changes of contrast sensitivity with scale, the effects of surround on contrast and attention and other cognitive mechanisms may provide additional guidelines for generating thumbnails that offer a representative impression of their high resolution originals.

REFERENCES

1. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Inc., 1989.
2. R. Samadani, S. Lim, and D. Tretter, "Representative image thumbnails for good browsing," in *IEEE International conference on image processing*, **II**, pp. 193–196, Sep 2007.
3. A. Woodruff, A. Faulring, R. Rosenholtz, J. Morrison, and P. Pirolli., "Using thumbnails to search the web," in *Proceedings of CHI*, pp. 198–205, Apr 2001.
4. B. Suh, H. Ling, B. Bederson, and D. Jacobs, "Automatic thumbnail cropping and its effectiveness," *Proceedings of the 16th annual ACM symposium on User interface software and technology* , pp. 95–104, 2003.
5. K. Berkner, E. Schwartz, and C. Marle, "SmartNails: display- and image-dependent thumbnails," in *Document Recognition and Retrieval XI, Proceedings of SPIE*, **5296**, pp. 54–65, Dec. 2003.
6. S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. on Graphics* **26**, 2007.
7. Y. Ke, X. Tang, and F. Jing, "The design of high-level features for photo quality assessment," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, pp. 419–426, 2006.
8. A. S. Wilsky, "Multiresolution markov models for signal and image processing," *Proceedings of the IEEE* **90**, pp. 1396–1458, August 2002.
9. D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing mag.* , pp. 43–64, May 1996.
10. R. Ferzli and L. Karam, "No-reference objective wavelet based noise immune image sharpness metric," in *IEEE ICIP 2005 Proceedings, Vol. 1*, pp. 405–408, Sept. 2005.
11. D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. on Inf. Theory* , pp. 613–627, Dec. 1995.
12. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Inc., 1993.
13. B. Bausk and R. Samadani, "Manually generated representative image thumbnails," Tech. Rep. HPL-2007-169, HP Labs, Oct 2007.
14. A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," *International Conference on Computer Vision* **2**(9), pp. 1033–1038, 1999.
15. A. Efros and W. Freeman, "Image quilting for texture synthesis and transfer," *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* , pp. 341–346, 2001.
16. B. Wandell, *Foundations of Vision*, Sinauer Associates, 1995.