



Manually generated representative image thumbnails

Brett Bausk and Ramin Samadani
Media Technologies Lab
HP Laboratories Palo Alto
HPL-2007-169
October 19, 2007*

image resampling,
image resizing,
image search and
browsing,
thumbnails

How would you best represent a high resolution original image given a fixed number of pixels for an image thumbnail? Is it possible to create better thumbnails than the standard computer-generated thumbnails that apply filtering and subsampling? In this report, these questions are addressed by manually generating thumbnails that attempt to well-represent the original high resolution images. Thumbnails are generated for several example images that contain high spatial frequency image textures that are difficult to represent using standard thumbnails. For each example image, we discuss: 1) the reasons the image was selected; 2) the manual steps used to generate the representative thumbnail; and 3) our informal observations about the results and the comparisons between the manual and standard thumbnails. Common themes distilled from these examples may serve as guidelines for developing automatic algorithms for generating improved thumbnails.

* Internal Accession Date Only

Approved for External Publication

© Copyright 2007 Hewlett-Packard Development Company, L.P.

Manually generated representative image thumbnails

Brett Bausk¹ and Ramin Samadani²

1) hp.com 2) HP Labs, Media Technologies Lab
Ramin.Samadani@hp.com

Abstract

How would you best represent a high resolution original image given a fixed number of pixels for an image thumbnail? Is it possible to create better thumbnails than the standard computer-generated thumbnails that apply filtering and subsampling? In this report, these questions are addressed by manually generating thumbnails that attempt to well-represent the original high resolution images. Thumbnails are generated for several example images that contain high spatial frequency image textures that are difficult to represent using standard thumbnails. For each example image, we discuss: 1) the reasons the image was selected; 2) the manual steps used to generate the representative thumbnail; and 3) our informal observations about the results and the comparisons between the manual and standard thumbnails. Common themes distilled from these examples may serve as guidelines for developing automatic algorithms for generating improved thumbnails.

1 Introduction

We believe there is an opportunity to develop new image resizing algorithms that better represent their high resolution originals. Image thumbnails are used in the majority of image products and applications, where they allow quick preview of the image content of the underlying high resolution images. For many years, standard thumbnails have been generated by filtering and subsampling the high resolution originals [1]. Although this process prevents aliasing and it preserves the image composition, it loses information about image quality and high spatial frequency textures [2]. The difficult images with high spatial frequency textures and content are the subjects of this study.

Recently, several non-traditional approaches to reducing image size have been developed [3, 4, 2]. By repeatedly removing paths that pass through uninteresting regions of an image, Avidan and Shamir [3] describe content-based image resizing. The removal of the paths, in effect, modifies the scene geometry and composition, providing an effective generalization of cropping [4]. An approach to thumbnail generation that we previously developed [2], does not modify the image composition, but rather uses the image content to generate thumbnails that better reflect the image quality of their originals. Subjective evaluations [5] of this latter approach has shown that these new automatically generated thumbnails better represent images with blur or noise. Images with high spatial frequency textures, however, are difficult examples for thumbnail generation.

The practical importance of thumbnails and the recent work in non-traditional image resizing leads us to pose a basic question: *what is the most representative thumbnail one can generate given a fixed number of thumbnail pixels?* This report does not discuss algorithmic solutions to this difficult problem, but rather discusses experiments in manual generation of image thumbnails using the popular Adobe[®] Photoshop[®] photo editing software. The experiments use images containing textures with spatial frequencies higher than the thumbnail sampling frequency in order to learn how to best generate thumbnails for these images. Unlike standard thumbnails, the manual approach to thumbnail generation implicitly takes into account perceptual and cognitive aspects. Review of the manual thumbnail generation process then helps identify some of these important aspects. This report concludes with some common themes, distilled from these experiments, that may serve as guidelines for developing new, automatic algorithms for generating representative thumbnails.

2 Image selection

This section describes the images used in the experiments in this report. The images are best viewed on a display since they are not tuned for printing. Textured images are used because, based on prior experience [5], they are difficult to represent as thumbnails. Image textures may be regular or stochastic [6, 7] and one example from each category was selected. The image in Figure 1 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 3 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.

Images where many different textures appear were also used. Figure 5 shows an indoor scene showing fine, crafted decorations found on a church wall. This image is selected to represent one common class of images, where fine, man-made textures, such as the writing seen near the bottom, are clear in the originals but lose details in the thumbnails. In addition, it is hard to

distinguish the different image elements from the thumbnails of these dark images. Figure 7 shows an example of an image that was selected because it contained many different natural textures, including grass, bushes and trees.

Finally, the image shown in Figure 10 shows an image where the subject of interest, the cat, only covers a small portion of the entire image. Images with small subjects are very difficult to represent and a special type of *picture-in-picture* thumbnail was generated for this example.

For each of these 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. Except where noted, the thumbnails are 128 pixels wide and 96 pixels high. The original images are shown at about 70% of their correct size in order to fit this document. It was confirmed that the appearance of these smaller images is similar to the images viewed on a display at full scale.

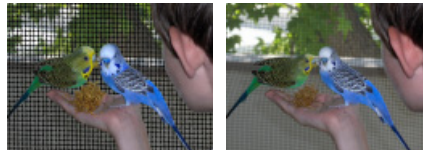


Figure 1: Manual thumbnail shown on the top left, and standard thumbnail shown on top right for the original shown on the bottom. The figures in this report are best viewed on a display.



Figure 2: The modified original before thumbnail generation.

3 *Birds*

Figure 1 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.

3.1 Process used to generate manual thumbnail

The manual thumbnail for this image, and most of the examples that follow, was generated by first creating a modified high resolution image (for the input image in Figure 1, this modified image is shown in Figure 2) 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 1 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. The screen texture was enlarged so that it would represent a perceptually similar texture 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.

3.2 Comparison with standard thumbnails

The thumbnail on the top right of Figure 1 is the standard thumbnail generated by filtering and subsampling. The screen door is not well preserved in the standard thumbnail since its spatial frequency is higher than the thumbnail sampling frequency, and it is thus removed by the antialiasing filter [1]. On the other hand, the manually generated thumbnail on the top left of Figure 1 shows the repetitive structure of the screen door at a larger scale than the texture would appear if properly scaled. Our conjecture, which may be tested with user studies, is that this manual thumbnail would be found more representative of the original, and that the scale distortion would not be objectionable.

In the image of Figure 1 the birds are the subject of interest. In addition to the enlargement of the screen, the birds were also enlarged in the manual thumbnail. This enlargement is very noticeable in Figure 2, the *preprocessed* 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 1. In a real application only one thumbnail would be used to represent the original. We conjecture that the manual thumbnail would be preferred as a representation of the original, and the subtle enlargement of the birds would in fact be beneficial. In a way, the enlargement attempts to preserve the saliency, or attention paid to the birds when viewing the original image.

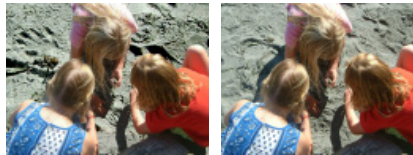


Figure 3: Manual thumbnail on the top left, and standard thumbnail on the top right for the original on the bottom.

4 Sand



Figure 4: The modified original before thumbnail generation.

The *sand* image in Figure 3 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.

4.1 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.

4.2 Comparison with standard thumbnails

There are noticeable differences between the standard thumbnail shown on the top right of Figure 3 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 4 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 5: Manual thumbnail shown on the top left, and standard thumbnail shown on top right for the original shown on the bottom.

5 Church

The *church* image in Figure 5 shows the artwork on the wall of a church. In this example, typical of indoor images of man-made decorations, there are fine structures with low image contrast, such as the writing seen in the bottom center of the figure.

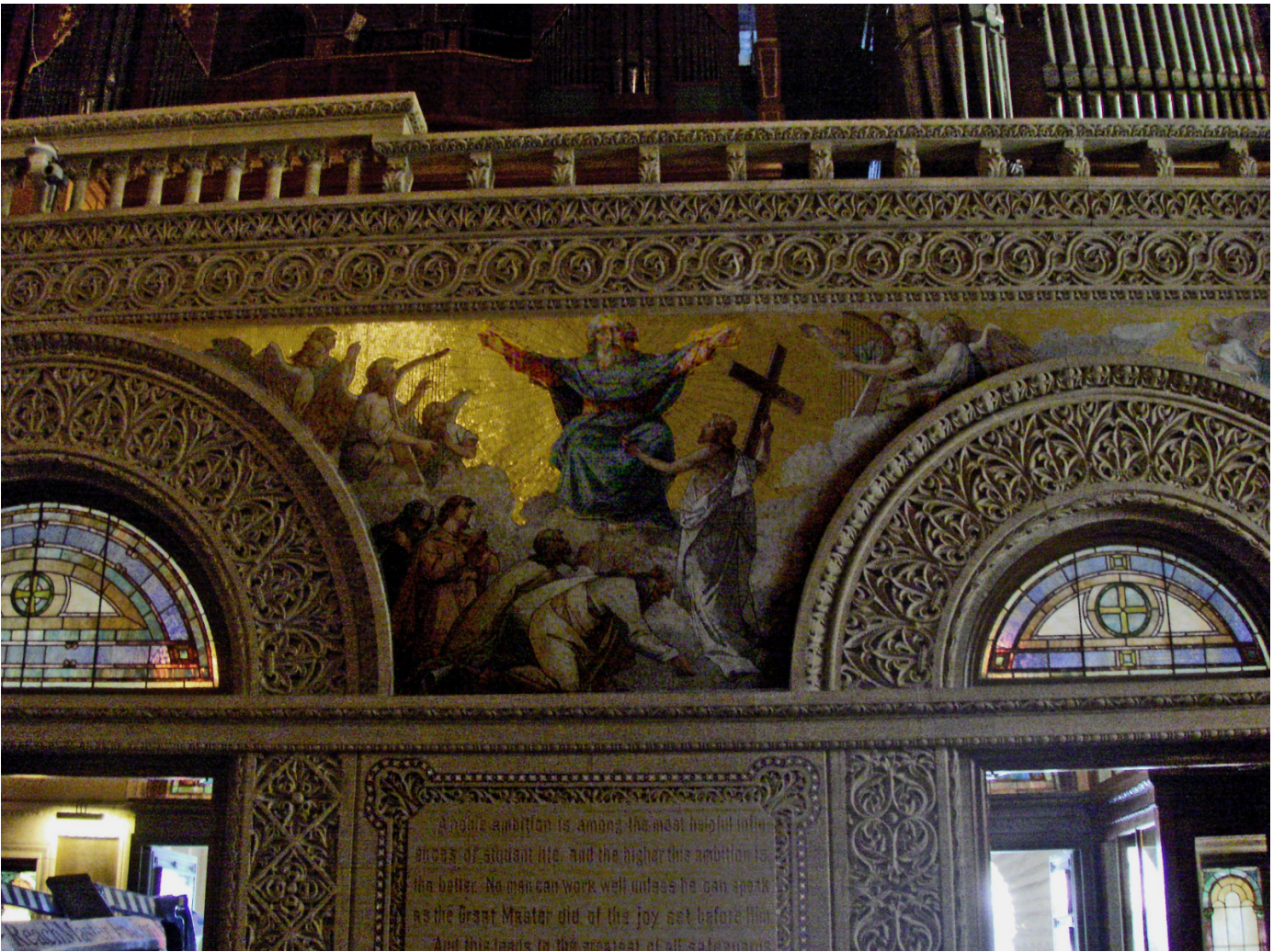


Figure 6: The modified original before thumbnail generation.

5.1 Process used to generate manual thumbnail

In this image, there are a number of details that are evident in the original, but are lost in the standard thumbnail. In the original, there is text that, although in shadow, is clearly visible. The primary focus of the image is the painting on the wall. Again, in the original the painting is in shadow, but the subject is clear. Lastly, there is a great deal of texture in the archway that is lost in the standard thumbnail. In order to bring these details back, a mask was used to select the archways and the text inscription. Curve levels were adjusted so that the detail was clear and bright, but not so much to change the character of the scene. Another mask was used to select the painting. Curve levels were then adjusted in order to bring out more of the painting details. Finally, the man in the painting was enlarged by warping the layer to give him greater prominence without exaggerating the proportions too much.

5.2 Comparison with standard thumbnails

In the standard thumbnail shown in the top right of Figure 5 the details are less apparent compared to the original. In the manual thumbnail, shown on the top left of the figure, more details are seen, including the text near the center bottom of the image. The manual thumbnail also better reflects the lightness of the original. This example is typical of indoor scenes with low contrast, where the filtering used to generate the standard thumbnail loses the details that are visible in the high resolution original. The increase in contrast is seen in the modified high-resolution image, shown in Figure 6, from which the manual thumbnail was generated. In addition, examining Figure 6 shows that the figure in the center, facing the crowd, has been enlarged in order to

be more apparent at the thumbnail resolution.

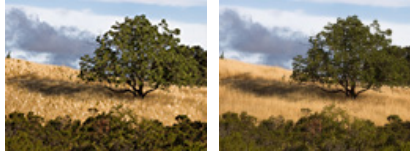


Figure 7: Manual thumbnail shown on the top left, and standard thumbnail shown on top right for the original shown on the bottom.

6 *Field*

The *field* image in Figure 7 shows a number of natural textures in the grass, trees and bushes.

6.1 Process used to generate manual thumbnail

During creation of the manual thumbnail, the standard thumbnail shown on the top right of Figure 7 was first examined. In the standard thumbnail, most of the leaf and grass structures are lost. A multiple-step process was used to enlarge and enhance



Figure 8: The modified original before thumbnail generation.

the contrast of the grass so that the grass texture would remain visible at thumbnail resolution. First, a temporary layer was formed using a representative portion of the grass and its size increased ten-fold. The shadow detail was extracted from the temporary layer, and this new layer was darkened. Next, the highlights were extracted from the temporary layer, and this layer was lightened. The enhanced highlight and shadow layers were then combined and superimposed on the original full-resolution image. This process maintained the overall shading gradient of the hill while increasing the grass texture. To maintain the tree shadow cast onto the grass and the grass texture, the original shadow detail was separated into another layer that was overlaid as the top layer. In addition to the grass edits, the tree was also separated into a layer, and its contrast and saturation was increased.

6.2 Comparison with standard thumbnails

Comparing the manual thumbnail shown on the top left of Figure 7 with the standard thumbnail shown on the top right shows clear differences between the two thumbnails. In particular, the grass texture of the standard thumbnail is not very prevalent compared to the original image shown in the bottom. The manual thumbnail definitely shows more prevalent textures, reflecting the textures that appear in the full resolution image. The manual thumbnail grass texture seems much coarser in proportion to the rest of the thumbnail, whereas the grass texture in the original full-resolution image seems more refined.



Figure 9: Picture-in-picture thumbnail shown on the top left, and standard thumbnail shown on top right for the original shown in Figure 10.

7 A Special Picture-in-Picture Thumbnail

In this section, a special technique is described for generating thumbnails for images where the subject matter is small in proportion to the entire image size. The image used is shown in Figure 10, where a cat in a box is seen.

7.1 Process used to generate manual thumbnail

To generate the thumbnail for the example shown in Figure 10, the subject in the image (cat in a box) was identified and the cat's face was cropped to form one component of the thumbnail. The contrast and saturation of the cropped face was increased to retain more detail at thumbnail resolution. The size of the cropped region at thumbnail resolution is 78 pixels wide and 52 pixels high. In the second step a second component of the final thumbnail was formed: a standard thumbnail of the entire image was generated at a size of 78 pixels wide and 104 pixels high. The *zoomed in* cropped face was placed at the top of the thumbnail for the entire image, to form the manual thumbnail seen on the left of Figure 9. The size of this manual thumbnail is 78 pixels wide by 156 pixels high. The standard thumbnail, of size 96 pixels wide and 128 pixels high is shown on the right of Figure 9. The total pixel count is kept the same for these two thumbnails, even though the aspect ratios of the two thumbnails differ.

7.2 Comparison with standard thumbnail

When looking through an unfamiliar collection, the thumbnail should provide a good understanding of the subject of the image. The user may be searching for images of a particular subject, in this case a cat, even if the image is not perfect. For the example shown in Figure 10 it is unclear from the standard thumbnail shown on the right of Figure 9 that there is a cat in the box. This situation often occurs when the subject of interest is small in proportion to the whole image. In these cases, depending on thumbnail size, the subject may not be clearly visible in the thumbnail. The manual *picture-in-picture* thumbnail on the left of Figure 9, however, shows that there is a cat in the image.

8 Common themes and findings

This section summarizes some of the general findings distilled from the experiments in manually generating representative thumbnails. Both the techniques used, and how the image content affects the thumbnail generation process are described in this section. It is hoped that some of the issues discovered during the experiments here may be further tested by more careful subjective evaluations, and that the methodology used in generating the manual thumbnails may influence the development of automatic algorithms for image resizing.

It is known to be impossible, by traditional filtering and subsampling, to represent image textures that have higher spatial frequencies than the thumbnail sampling frequency. It is impossible to preserve both the local image texture structure and the scaling accuracy. 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 manual 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 1 may show the *appearance* of the screen even though it is depicted at an inaccurate scale.

Many images have a clear subject of interest, one 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 10: Original cat image corresponding to the thumbnails in Figure 9.

Figure 2 but it is also seen in the hand-print shown in the processed *sand* image shown in Figure 4 and in the processed *church* image shown in Figure 6. For the example shown in Figure 10 the subject is very small and in this case it is not possible to enlarge the cat without introducing large geometric distortions in the scene. For this example, the alternate approach of picture-in-picture thumbnail was used.

There are also images, particularly of dark indoor scenes with fine details where the contrast of the details is lost in 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 [1] as well as the lighter surrounds of the thumbnails. In this case, compensating for the

change in contrast improves 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 or the picture-in-picture technique. Some of the techniques used seem to adjust for known perceptual contrast sensitivity changes with spatial frequency [8]. For example, the 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.

The choice of techniques used may need to 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 who is familiar with a collection will use the thumbnails as a reminder of particular images to use for their current tasks. In this case, spatial or scale distortions in the thumbnail may be undesirable.

It is interesting that the manual thumbnails often were generated by changes that on the one hand may be considered dramatic and on the other hand may be considered subtle. For example, we *drastically* enlarged the screen door repetition period in Figure 2. On the other hand, the manual thumbnail subtly reflects the *impression* of the screen door in the high resolution image. 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] Anil 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*, 2007, p. 10.
- [3] Shai Avidan and Ariel Shamir, "Seam carving for content-aware image resizing," in *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, New York, NY, USA, 2007, p. 10, ACM Press.
- [4] B. Suh, H. Ling, B.B. Bederson, and D.W. 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] R. Samadani, T. Mauer, D. Berfanger, J. Clark, S Lim, and D. Tretter, "Honest image thumbnails: Algorithm and subjective evaluation," Tech. Rep. HPL-2007-88, HP Labs, June 2007.
- [6] A.A. Efros and T.K. Leung, "Texture synthesis by non-parametric sampling," *International Conference on Computer Vision*, vol. 2, no. 9, pp. 1033–1038, 1999.
- [7] A.A. Efros and W.T. Freeman, "Image quilting for texture synthesis and transfer," *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 341–346, 2001.
- [8] Brian Wandell, *Foundations of Vision*, Sinauer Associates, 1995.