

Spatially Adaptive Rendering of Images for Display on Mobile Devices

Amit Singhal, Jiebo Luo, Christophe Papin, and Nicolas Touchard
Eastman Kodak Company
Rochester, New York

Abstract

Mobile imaging is enabling visual communication "any time, anywhere" to become a reality. Using today's wireless technology, consumers can use their mobile device to capture, review, share, and print pictures while "on the go." Apart from wireless communication issues, a key technical challenge is how to achieve best-perceived image quality, given the limited screen size and display bit-depth of most of the mobile devices in common use today. The limited hardware and displays need to be utilized in an intelligent and effective fashion to communicate image information to the mobile device user. In this paper, we present a spatially adaptive rendering scheme to generate visually enhanced images for display on mobile devices. More specifically, we have developed a method for adapting the image-rendering algorithm for different spatio-frequency regions of an image, based on subject content analysis. When trying to achieve a specific bit rate, compression ratio, color bit-depth, or image resolution, this method has the effect of maximizing the visual quality of regions that are likely to be important for the reconstructed image quality, as perceived by the user viewing the image on a mobile device. Experimental results using consumer pictures are shown to demonstrate the efficacy of the proposed system.

Introduction

Images are a means of communicating and sharing feelings of belonging between people. They allow people to stay emotionally connected while physically apart. While communication and sharing can take different forms, none is as powerful as the visual form. Videos and pictures are most effective in sharing, expressing, and remembering one's life. The continued progress in digital imaging, wireless and broadband connections, and improved mobile hardware has enabled new, richer ways for families and friends to share vacation pictures, babies' first steps, and graduation pictures, further fostering the feeling of connectedness. Figure 1 shows an overview of a typical mobile imaging scenario. A number of wired and wireless devices are connected to an imaging and media applications system via the Internet. The imaging and media applications system further comprises of a storage

area for video and images, a device discovery component that identifies the type of mobile device requesting access to the media, and an image-rendering engine to re-render the stored media for optimal display on the requested device.

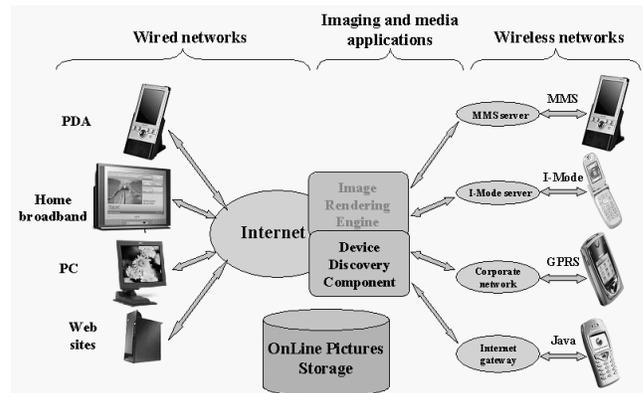


Figure 1. Mobile imaging "any time, anywhere."

In addition to allowing family and friends to stay connected, images and videos are also effective ways of communicating commercial information. In particular, wireless imaging is enabling a new way of performing many business tasks such as news agency photographers sending pictures from the world's hot spots out to hundreds of US newspapers; people reading news footage dramatically enhanced by the insertion of images and videos related to the events; insurance adjusters filing images of a burned house or a damaged automobile from the field; construction-company engineers sending pictures of a project back to the home office; or even marketing spies sending back images of rivals' products on display at a trade show. We expect that advances in mobile imaging technologies will have a significant impact in the business market and grow at a rapid rate.

While there have been many advances in mobile imaging and communication technology, major technical challenges still limit the applicability and usage of these devices. Some of these challenges include: limited connection bandwidth, sparse coverage area, hardware

capability limitations (power requirement, computing power, memory/storage, etc.), multiple communication standards, image rendering and display, image management, and, last but not least, consumer ease-of-use.

In particular, conflicting requirements for wireless imaging impose further challenges on some of these technologies in terms of displayed image quality.¹ Many color image output mobile devices are not capable of displaying all the colors in an input image because they must be stored in a memory buffer with a reduced bit-depth. It may also be desirable to represent an image using a reduced bit-depth, in order to reduce the amount of bandwidth needed for transmission or the amount of memory needed to store an image. In the early years, many computers used an 8-bit color representation to store an image that was to be displayed on a soft-copy display, such as a CRT or an LCD screen. Such representations allow only 256 unique color values. This is significantly less than the 16,777,216 possible color values associated with a typical 24-bit color image. This problem has attracted renewed interest with the recent boom of cell phones and PDAs (personal digital assistants). The problem is made more acute by the severely limited display size, in addition to limited display bit-depth. Most mobile devices have small-sized display screens, a result of form-factor limitations associated with the small size of these devices. While the smaller screen size makes it imperative that the displays have the highest color bit-depth possible to achieve good image-rendering results, physical limitations on the size of these devices does not permit a high-resolution display or a color bit-depth higher than 8 bits in total (instead of 8 bits per color channel) in most cases. Because most mobile imaging usage scenarios call for wireless access to multimedia, the constraints on bandwidth necessitate small image file sizes that can be hard to achieve without significant levels of compression. All of these aggravations not only affect image perception but also severely hamper ease of use when it comes to image-related applications on mobile devices.

In this paper, we focus on innovative ways of making digital imaging easy and effective on mobile devices. We begin by reviewing the current device/channel capabilities supporting image display, revealing the practical issues and motivating the potential technology solutions. In particular, we describe a region of interest-based schemes for spatially adaptive rendering of images for display on a mobile device, which results in a perceptually improved user experience.

Region-of-Interest Detection

Mobile devices are currently limited in terms of their display capabilities. Most have display screens with less than 8 bits of color information, high viewing flare, and small display resolution. Image attributes readily perceivable by a human observer in a high-color bit-depth, high-resolution display (such as a SVGA monitor) cannot be readily seen in reduced bit-depth, reduced-size images

rendered on mobile devices with limited display capabilities.

In addition, the advances in mobile devices such as PDAs (e.g., Palm products and pocket PCs) and cell phones usher in a variety of new mobile computing applications. Given today's hectic life style, people are increasingly attracted to the appeal and benefit of mobile computing—accessing and manipulating your data from anywhere at anytime. From the early innovative applications of the address book and scheduler, mobile computing is rapidly moving into territories unimaginable just a couple of years ago. While laptop computers are becoming closer matches to their desktop counterparts in terms of computing power, their size, weight, and power consumption are still a great hindrance to mobile computing. The computing power of PDAs has reached a point that can enable people to start thinking about more computationally intensive mobile computing tasks such as image processing.

In order to allow imaging applications to run effectively on small-sized displays, it is necessary to render the image in a manner that makes it visually preferential for the imaging application. This, in turn, makes it necessary to accord preferential rendering treatment to regions of interest in an image. As an example, in the context of user-assisted red eye correction, it is important to render the face and eye regions at the highest level of detail possible, while rendering other regions at lower resolution to stay within the constraints of a mobile computing system. As another example, in the context of a real-estate application, the regions of interest are the architectural details of a listing rather than the people present in the image. Thus, for image display, it is often desirable to render the main subject of an image at higher resolution, preserving details and color (if possible), than the background regions of an image.

Regions of interest can be obtained via user interaction or facilitated by an automatic region-of-interest detection system. The user-defined regions of interest can be generated offline or acquired online in an interactive setting. The automatic main-subject detection system² can be used offline to generate the regions of interest in an image. Similarly, a skin³ or face⁴ detection algorithm can also be used generate region-of-interest masks that preserve the color and spatial details in people present in images. A region-selective rendering scheme that preserves edge details and colors in the detected regions of interest can be used to create a visually enhanced (but similarly constrained in terms of file size, number of colors, etc.) reduced bit-depth, reduced-size version of the higher bit-depth, high-resolution original image. If the user uses a mobile device to mark regions of interest in an image, the re-rendering operation can be performed in the mobile device (if it has enough CPU power and memory) or sent via a limited bandwidth communication link to a central image processing server which returns a re-rendered image with enhanced main subject regions.

The system proposed addresses the need to provide an image-rendering system that is capable of (1) acquiring a high resolution, higher bit-depth image, where (2) the regions of interest may be detected automatically via image understanding algorithms, or (3) the regions of interest may be selected by the user via a mobile device, and (4) and generating a reduced resolution, reduced bit-depth image by spatially varying the rendering of the original image, based on the detected regions of interest. In order to satisfy constraints such as file size or visual appeal, the rendering process can be repeated by modifying the region-of-interest map or the rendering scheme until the constraints are satisfied. The regions of interest in an image may also change, depending on the imaging application. As an example, an image containing a person in front of a house may have the person as the region of interest for a personal digital albuming application, but the house as the region of interest for a real estate application.

In the next section, we describe a system that uses region-of-interest detection to generate spatially adaptive rendered images for display on mobile devices. We also present some results of the comparison of these spatially adaptively rendered images versus those rendered using a uniform scheme.

Spatially Adaptive Image Rendering

A mobile imaging system has to carefully maintain sufficient image quality according to both the device characteristics and the network bandwidth restrictions. The proposed method provides a region-of-interest- (ROI) based rendering scheme that efficiently decreases the output image size while preserving important areas in the image. Two kinds of ROIs have been identified in the prototype system: face and textured areas. Depending on the application, other ROIs (such as architectural edges for a real estate application) may be more pertinent. Dedicated processing is separately applied on background and foreground (ROI), according to the specified output format (i.e. indexed image or not). ROI-based processing involves adaptive filtering and quantization with adjustable error diffusion.

ROI Extraction

As mentioned earlier, we have selected two types of ROIs for our prototype system, faces, and textured areas. For face detection, we use a Bayesian classifier based on the *maximum a posteriori* (MAP) criterion to delineate rough areas that are likely to contain faces (denoted as FD areas).⁵ Further refinement of these rough areas is necessary to automatically and precisely extract the whole face surface in order to avoid generating artifacts around ROI/background boundaries. Moreover, the neck has to be rendered similarly to the face when connected. This is accomplished by coupling an iterative skin detector (SD) to the FD classifier. The SD is run a first time with selective parameter values providing only limited flesh-color areas. A connected component-labeling scheme involving a

combination of morphological closing and dilation operations eliminates false alarms and small isolated skin areas while enclosing eyes and mouth regions. False alarms typically occur for flesh-colored background regions (e.g., reddish pixels) located outside or slightly within FD areas. They can also be parts of human bodies such as hands or legs that are assumed less important than faces. Flesh-colored areas are iteratively increased by running the skin detector with wider parameter settings while keeping the total number of skin pixels within FD areas lower than a threshold (75%). Results are depicted in Figs. 3 and 5.

Image simplification such as spatial smoothing can cause unacceptable visual artifacts within textured areas. Thus, for improved image perception, it is necessary to preserve them from severe processing. The key idea in our scheme is to render large uniform areas such as sky or sea with only one color. A thresholded version of the amplitude of the intensity spatial gradients is used to compute such ROIs. Once again, the derived maps of textured areas are regularized by means of morphological operators. An example of texture preservation is shown in Fig. 7. Realistic classifications of large, non-textured regions can be performed by applying specialized algorithms such as grass⁶ or sky⁷ detectors.

Rendering Scheme

To generate an image for display on a mobile device, we have developed a two-stage rendering scheme (see Fig. 2) involving an enhancement step followed by a ROI-based adaptation step.

The scheme successively performs image enhancement (color balancing, noise reduction, gamma correction, etc.), image resizing and sharpening (eventually including a cropping operation), ROI-based image-rendering, auto rotation according to the client display, and, finally, image compression. The ROI-based image-rendering step depends on the output image format (indexed or not).

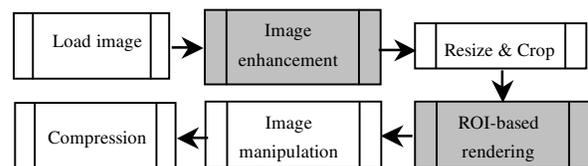


Figure 2. ROI-based image-rendering scheme.

Content-Based Dithering

For display on most mobile devices, the image needs to be rendered with fewer colors than are present in the input image. This can be done by applying a dithering method, e.g., the Floyd-Steinberg algorithm (FS).⁸ However, such error diffusion algorithms produce visually noticeable artifacts (“wormy” textures in highlights and shadows) as a result of the dot placement choices. However, low-contrast and low-resolution screens

available on cell phones or PDAs can reduce the influence of these artifacts.

For optimal viewing preference, we have chosen to adapt the error diffusion rate according to the content. The standard FS algorithm with a high error diffusion rate is applied within ROI (faces) to preserve details, while a lower error dispersion ratio (25%) is used in the background areas. Fig. 3 presents comparisons between 8-bit rendered 208 x 176 images quantized with a 3-3-2 color palette (Fig. 2 without the ROI-based processing stage) and ROI-based rendered images. The image size was reduced by 22% and 21% (top row and bottom row), respectively. For better visibility of the effect, Fig. 4 shows zoomed results (on a different image). The effect of the content-based dithering is clearly visible on the sweater and the background regions in the image in Fig. 4. Results on a large database are shown in the Experimental Analysis section.

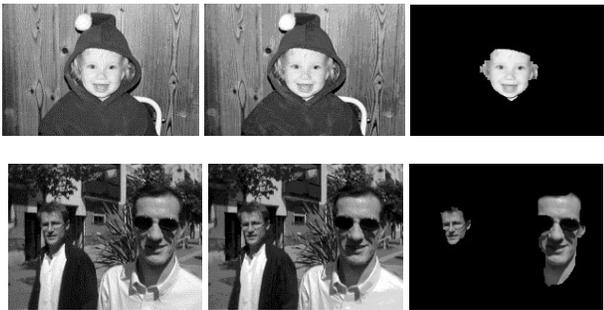


Figure 3. Content-based dithering version of the ROI-based image rendering (center) versus standard image rendering (left column). Right column depicts detected faces.

Looking at Fig. 3, we can see that a comparison between the left and the central column shows that faces are preserved, while background regions undergo degradations causing flat areas of unnatural looking approximate colors. However, on a mobile device display, the user's major interest is focused on the faces and this degradation of color in the background regions is not found to be too objectionable.



Figure 4. Details of a content-based dithered image (right) versus a standard rendered image (left column). Compression gain reaches 37% because of the uniform background.

The content-based dithering step is activated only when an indexed output image (PNG, Gif formats) is request.

Content-Based Adaptive Filtering

We have also investigated a ROI-based rendering scheme dedicated to non-indexed output image formats such as JPEG. Color reduction (as in the indexed case) is not employed for producing a JPEG image. In this case, the image would not be in accordance with the color palette of the display, and the quantization may generate gross artifacts in the JPEG output. Instead, a sigma filter⁹ is applied to the image to reduce the file size. The sigma filter reduces image noise and excessive details, sharpens regions, preserves edges, and retains thin lines. We use a fairly large window (15 x 15) and a value of sigma equal to 30.



Figure 5. Content-based adaptive filtering of a ROI-based image rendering (center) versus standard image rendering (left column). Right column depicts detected face.

Figure 5 shows the effect of content-based adaptive filtering for a non-indexed image format. The central column represents a ROI-based, filtered image and shows that areas with rather low details (e.g., creased jacket and curtain patterns) are smoothed. In these examples, the respective file size savings are 23% and 13% (for the image on top and bottom for an image resolution equal to 208 x 176). Again, for better visibility, Fig. 6 shows zoomed results (on a different image). Adaptive smoothing is clearly visible on the sweater. Complete results are shown in the Experimental Analysis section below.



Figure 6. Details of a content-based, filtered image (right) versus a standard rendered image (left column). Compression gain reaches 26% for this image because of the uniform background.

The content-based adaptive filtering step results in a lower compression gain than the dithering step used for indexed image formats. However, it does not result in as many noticeable visual artifacts.

Texture-Based Adaptive Filtering

If an image does not contain significant face regions, we can still achieve some compression gain without excessive loss in perceived image quality by using a texture ROI for content-based rendering. A texture-based, rendered image is shown in Fig. 7. Again, we make use of a sigma filter for content-based adaptive filtering. Regions with a low level of texture details, such as the sky, are greatly smoothed, while areas with a high level of details are preserved. In this example, the sky region becomes uniformly white, resulting in a further compression gain of about 7%.

In general, we have observed that face-based rendering provides higher compression gains than texture-based rendering. Significantly, a joint exploitation of these two types of information seems to offer the most benefit. Indeed, textured areas are likely to appear in most of the scene, while faces tend to occur in limited areas.



Figure 7. Texture-based image rendering (center) versus standard image rendering (left column). Right column depicts detected textured areas.

Experimental Analysis

This section presents an experimental analysis of our ROI-based rendering scheme on a database of 100 consumer images with face content. The images have been obtained from various sources, including 1 to 4 MP digital cameras and VGA PhoneCams. The rendering scheme depicted in Fig. 2 is used to generate the ROI-based rendered images. GIF images are quantized to an image-dependent, 8-bit color palette (3-3-2). We first present results of the rendering scheme as a function of the output resolution, followed by the influence of the input image content on the image file size.

Output Resolution Influence

Table 1 depicts the increase in compression gain for a ROI-rendered image versus a standard version of image rendering. The output image resolutions are representative of current screen resolutions that can be encountered in mobile devices in use worldwide. We have selected 4 representative output resolutions available on cell phones and PDAs.

As noted previously, content-based (CB) dithering demonstrates higher compression gains than content-based adaptive filtering. However, visual artifacts are more noticeable in the former case. This last point is, however, highly dependent on the background content (see Table 2). As expected, the compression gain goes up with increasing output resolution as more image areas can be subjected to simplification. Greatest compression gain is obtained for the newest PDA screens with a 320 x 240 resolution display.

Table 1. Performance of the face-based image rendering scheme vs standard image rendering for different output image resolutions.

Ouput Resolution	CB Dithering (GIF)	CB Filtering (JPEG)
320 x 240	32%	24%
206 x 176	30%	18%
132 x 176	29%	16%
101 x 80	24%	11%

Table 2. Compression gain for the ROI-based, image rendering scheme for 3 sets of 15 images with different levels of detail within the background regions.

Level of Bkgd Detail	CB Dithering (GIF)	CB Filtering (JPEG)
Low	52%	22%
Medium	32%	12%
High	12%	10%

Image Content Influence

As one might expect, the compression gain is significantly dependent on the size of the ROI within the image. However, we cannot derive a generic function of compression gain versus ROI size. Indeed, image content within background regions also has a major influence on the resulting image size.

To illustrate this point, we have identified (in a heuristic manner) 3 sets of 15 images with low, medium, and high level of background details. Results for the compression gain over a non-ROI-based scheme are shown in Table 2. The image output resolution was set to 208 x 176. Significant compression gains are achieved for images with simple, uniform, background regions. Even for complex images with high level of detail in the background, we are able to achieve a compression gain of 10-12 %. Most significantly, a set of human observers was shown the images generated using an ROI-based and a non-ROI-based rendering scheme. In almost all of the cases, the human observers either preferred the ROI-based rendering or had no preference between the two. Thus, the use of an ROI-based scheme can improve the perceived image quality by allowing us to transmit a higher quality image as a result of the compression gains achieved.

Conclusions

In this paper, we have described a system for efficient and easy mobile imaging enabled by region of interest detection and spatially adaptive image rendering. We have demonstrated the usefulness of using important areas such as faces or texture in an image-file, size-reduction process. It enables the image-rendering engine to optimize quality within regions that are visually significant to the viewer. Our method provides a further compression gain of about 10% to 50%, depending on the output image format and

resolution. Greater compression gain is obtained for larger PDA screens than the smaller cellular phone screens. While the content-based (CB), adaptive filtering step provides lower compression gains than the CB dithering step, it results in less noticeable visual artifacts. Preservation of textured areas provides about a 7% reduction in image file size.

References

1. J. Luo, A. Singhal, G. Braun, R. T. Gray, N. Touchard, and O. Seignol, Displaying images on mobile devices: capabilities, issues, and solutions, *Proc. ICIP* (2002).
2. J. Luo, S. Etz, A. Singhal, and R. T. Gray, Performance Scalable Computational Approach to Main Subject Detection in Photographs, *Proc. SPIE Hum. Vision Electron. Imaging* (2001).
3. M. J. Jones and J. M. Rehg, Statistical Color Models with Applications to Skin Detection, *Proc. CVPR 99*, (1998).
4. H. A. Rowley, S. Baluja, and T. Kanade, Neural Network Based Face Detection, *IEEE Trans. PAMI*, (1998).
5. H. Schneiderman, A statistical approach to 3D object detection applied to faces and cars, *PhD thesis*, CMU-RI-TR-00-06, Carnegie Mellon University, (2000).
6. N. Serrano and J. Luo, Grass Detection in color image using wavelet feature and support vector machines, unpublished result.
7. A. Singhal and J. Luo, Hybrid approach to classifying sky regions in natural images, *Proc IS&T/SPIE 15th Symp. Electron. Imaging*, Santa Clara, USA, (2003).
8. R.W. Floyd and L. Steinberg, An adaptive algorithm for spatial gray-scale, *Proc. Soc. Inf. Disp.*, **17**, pp. 75-78, (1976).
9. J. -S. Lee, Digital image smoothing and the sigma filter, *Computer Vision, Graphics and Image Processing Vol. 24*, pp. 255-269, (1983).