IBM Research Report

Expanding the Reach of the Digital Camera

Chandra Narayanaswami, M. T. Raghunath
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

Expanding the reach of the digital camera

Chandra Narayanaswami, M. T. Raghunath
IBM TJ Watson Research Center
Email: {chandras, mtr}@us.ibm.com

Abstract

Digital cameras, first introduced to consumers in the late 1980s, have now become very affordable and are on the verge of outselling film cameras. 53 million digital cameras are expected to be sold in 2004. Many of these are integrated into cell phones. Some cars have started to include front and rear view cameras. Though digital cameras can hold a very large collection of high resolution images and have short and long range communication capabilities in the foreseeable future, some aspects such as their small display size and restricted input/output devices limit the potential value of such image collections.

We explore how the value of these cameras can be increased by forming symbiotic relationships with larger displays and other devices. We present some challenges that need to be solved on the way and how digital cameras may evolve to support such relationships.

1. Introduction

Within the last ten years digital cameras have evolved from being expensive and bulky devices providing low resolution images to affordable compact units capable of recording high resolution pictures. More recently, cellular phones that integrate digital cameras have far outsold regular digital cameras. We expect this trend to continue because of the value this combination provides. First, cell phones are the most ubiquitous portable devices around because voice communication is a need most people can relate to. Second, people enjoy the convenience of capturing high resolution digital images using a device they are already carrying. Some digital cameras have even started offering integrated WiFi capabilities for direct transfer of images. Recent trends on portable storage devices indicate that cellular camera phones can integrate several gigabytes of storage. Finally, cell phones can provide location information that can be automatically captured with the image.

While one may consider digital cameras as inevitable and straightforward replacements for film cameras, we have observed that the manner in which digital cameras are used is quite different. With digital cameras, people tend to take more pictures, and also take pictures in more situations than they do with film cameras. For example, since digital images do not require much physical space, some have converted children's art projects from school, and toys into compact yet accessible digital albums and discarded the space-consuming articles. A few have recorded the progress of a skin rash, a snake or dog that has bitten a person, on a digital camera to show a doctor. Other examples include snaps of flowers plants, birds, etc., on hikes for later comparison and classification.

Digital cameras have also become *memory aids* and *transcription devices* – people take pictures of where they parked their car at the airport so they don't have to spend time looking for their car in the lot at the end of a trip. Others have captured the license plate number of erratic drives and reported them to law authorities. Digital cameras are now also being used as *input devices* [6, 7]. Thus, a new class of images that we denote as *ephemeral images* is emerging. The lifetime of such images is more limited than pictures taken with film cameras.

There are several reasons to expect these types of applications to grow. Digital cameras are likely to be with the user more often, the cost of taking an extra picture is minimal, and plenty of images can be stored locally on the camera itself.

Even though the image resolution, storage capacity, and wireless connectivity of cellular camera phones are increasing, their physical size is dictated by human preferences essentially limiting the size of the integrated display. The advent of flexible displays may change the balance somewhat but such technology is in early stages at this point and even so, the size available in a truly portable form factor is going to limit the area of the rolled up display. Projection technologies are also being developed but at this point they consume too much power to be practical for most portable devices. Therefore, for the next few years, the size of the display on the camera will remain constrained. While the ability to immediately view pictures on the integrated display is valuable, only a limited amount of image detail can be shown on the integrated display of the digital camera. How can we address this limitation?

In a recent paper [12], the authors described a futuristic scenario where mobile computers establish symbiotic relationships with stationary devices in the environment in order to offer users a combination of the best attributes of both systems, namely personalization of content due to mobile computers as well as large, easy to read, high quality displays. Environmental displays become intelligent network objects that offer their services just like today's network printers. Mobile computers discover such displays, communicate with them to ascertain their characteristics, and securely transmit information.

The primary focus of this paper is the symbiotic mode of operation for digital cameras that include short and long

Table 1 Typical resolution of current displays and maximum possible resolution according to limitations on human visual acuity, using 20/20 human visual acuity, i.e., 1 minute of arc (adapted from [12]).

display type	user distance (inches)	typical width (inches)	typical width (pixels)	typical resolution (dpi)	maximum resolution[1] (dpi)	maximum width (pixels)
cell phone panel	10	1	100	100	350	350
PDA display	12	2	300	150	291	582
laptop display	16	10	1200	120	218	2180
desktop monitor	20	15	2000	133	175	2625
mtg. room screen	230	80	1200	15	15	1200

range wireless connectivity. We show how symbiosis between digital cameras and external displays can greatly enhance the application space of digital cameras.

2. Symbiotic Displays in the Environment

Many researchers have explored the idea of using environmental displays to supplement the display capabilities of portable devices [10, 13]. The key observation leading to this approach is that every display has intrinsic limitations that create usage barriers. For example, human visual acuity imposes an upper bound on display resolution. Given that even people with perfect vision cannot resolve details smaller than one minute of visual arc angle for prolonged durations at comfortable levels of brightness, increasing display resolution beyond that point does not contribute significantly to improvements in the readability of shown information. Table 1 (adapted from [12]) shows the typical resolution of current displays and the maximum "meaningful" resolution as a function of visual acuity. To gain some perspective, exotic IMAX theaters currently use 20-30 Mpixels per 70mm frame, a 35mm film frame has the equivalent of 2Mpixels, and HDTV has a resolution of 2 Mpixels. Most computer displays have less than 4Mpixels today.

It is easy to see from Table 1 that technological advances will not improve portable displays to a state where viewing a large amount of detail becomes an easy task. The improvements in picture quality resulting from better capture resolutions are seldom observable on the integrated displays of digital cameras. To provide an extended set of applications, we expect future digital cameras to establish symbiotic relationships with large environmental displays on demand.

In addition to cameras, we expect several other mobile devices to leverage the services of network attached displays to view documents and other types of content that cannot be viewed easily on small displays.

We believe that a new class of intelligent network-connected displays will emerge to support such symbiotic relationships. These displays will have some or all of the following attributes:

- Displays are directly connected to the wired network infrastructure and are addressable on the network.
- Displays support several content formats such as ASCII, HTML, PDF, Postscript, JPEG, GIF, MPEG, Flash, etc., and are able to negotiate with other network devices, in a manner similar to how browsers express "accept" tags to web servers. Displays also express their capabilities in terms of pixel resolution, dimensions to mobile devices.
- Displays also support direct user interaction via keyboards, mice, touch sensitive screens, or other forms of gesture recognition to enable users to perform simple operations such as scrolling, pausing video, etc.
- Displays also support downloadable code such as JavaScript, Java applets etc., for richer user interactivity.
- Displays preferably support a short-range wireless network interface to communicate with mobile devices, and to make it easy for mobile devices to discover just those displays that are in the immediate vicinity.
- Displays may also support one or more wired interfaces such as USB or FireWire to connect directly to mobile devices.

Though network attached displays have a lot in common with network printers, we believe these displays will be less expensive to operate since there are no consumables such as ink and paper that need to be replaced on a regular basis. We are already seeing trends in this direction with the addition of Ethernet and WiFi interfaces to projectors. Some even come with software to display contents from a remote PC. In the future, we expect to see network displays become standard appliances that are deployed in various environments. Over time, we expect several different types of network displays, such as projection displays, TFT, OLED, etc., will become available, all supporting standard software and hardware

interfaces. The prices of such displays are also likely to drop with time, widening their deployment further.

We have started prototyping such network connected displays by enhancing standard projectors as well as our Everywhere Display research prototype [11]. Recently we also demonstrated an application scenario where our WatchPad [10] wristwatch computer and the Everywhere Display establish a symbiotic session [1].

3. Application Environments

We expect digital cameras to be used symbiotically in several environments, including homes, offices, shops, cars, airplanes, and trains. This wide range introduces several challenges. Some of these environments are friendlier than others. Some have better access to power, some have organization firewalls, and some are more private than others. This variability impacts both the design of the camera and the infrastructure required.

3.1. Camera Image Capacity

Camera	Year, Max price	Image Resolution and Typical Size	Display Size (mm), pixels	Max Memory Available, Image Capacity
Ricoh RDC-1	1995 $1800	768x480 100KB	51x31, 72K	24MB, 240
Kodak DC210	1997 $600	1152x864 300-400KB	37x27	64MB, 160
Kodak DC290	1999 $800	1792x1200 600-800KB	41x28	128MB, 160
Sony DSC-P10	2003 $600	2592x1944 2MB	31x23, 123K	2GB, 1000
Samsung VGA1000 Cellphone	2003 $299	640x480 100KB	34x42 20.4K	

Table 2: Camera specifications over years

The size of the images captured by digital cameras has increased at a steady pace as shown in Table 2 in the last few years due to higher resolution imaging chips. The number of images that could be stored in the camera did not increase significantly during these early years since increased storage capacity was neutralized by increasing image resolution. As technology improved, several cameras became physically smaller as shown in Figure 1 and as a result the physical size of the display in those cameras also shrank.

Due to limitations of the human eye we do not expect the resolution of the captured image and the resolution of the integrated camera display to continue to grow significantly. Many consumers find that the current image capture resolutions are adequate because they can already print high quality enlargements. In contrast, we can expect camera storage capacities to continue to double every

several years effectively increasing the number of pictures that can be stored on the camera. The introduction of the IBM MicroDrives with at 340MB capacity in a compact flash form factor began this trend. Atomic force microscope (AFM)-based data storage technologies such as Millipede [16] can have a data storage density of a 125 GB per square inch -- ten times higher than the densest magnetic storage available today. Due to this difference in the rates of growth of image and storage size, we eventually expect users to retain large numbers of images in their cameras.

Figure 1: Cameras from 1995-2003

Figure 2: 512MB-40GB capacity portable storage

For example, a 40GB storage device can hold 20,000 images, each 2MB in size. A VGA (640x480) resolution thumbnail that can be used to view the image on the camera display takes about 100KB. Thumbnails for 10,000 images will take only one GB. Assuming a progressive user takes about twenty pictures a day and half of them are ephemeral ones that get deleted automatically, the storage device can hold images taken over a five year period. If the camera retains only thumbnails when the actual image is transferred out of the camera, a 40GB

storage device on the camera could hold 400,000 VGA resolution thumbnails!

Thus, relatively soon, some users will be able to retain images taken over several years on the camera, along with annotations. Other users will be able to retain thumbnails and annotations for all pictures captured over several years. As users upgrade to newer cameras they may wish to copy the entire contents of their current camera to their new camera.

Regardless of camera capacity, we expect that users will often copy their images from their cameras to home computers or to web repositories primarily to safeguard against the possibility of losing the camera. There are already several providers such as Ofoto, Yahoo, Shutterfly, etc., who offer network accessible storage and photo printing services. Such repositories also enable users to share their pictures with others and also obtain prints when necessary.

Just as MP3 players have enabled music lovers to carry huge collections of their favorite music, technology trends will allow consumers to carry a large collection of images on their person. In order to unleash the full potential of such portable image collections, we need to empower users to view and share images easily.

3.2. Direct two-party symbiosis

When users wish to view their images on intelligent environmental displays of the type described earlier, there are a few different ways in which this task can be accomplished depending on whether the high resolution image is present on the camera itself, or in some web repository or in both places.

If the high resolution images that the user wants to view are stored on the camera itself, the simplest approach is to transfer the image directly from the camera to the display as shown in Figure 3. However, in order to retain privacy and security in settings that may be less private than the user's home, the images are previewed on the camera display before being sent for viewing on the larger display. So this procedure differs from what we do today – i.e., just connect the camera to the larger display and view everything there.

First the camera and the display establish a session after the camera discovers the display and its capabilities and mutual authentication if required. The user then scrolls through the image thumbnails using the built-in display on the camera and picks one to view on the environmental display. The camera then sends the higher resolution image to the display. It is possible that the environmental display does not support as high a resolution as the actual image captured by the camera. In this case the camera may resize the image to the resolution supported by the display so as to reduce the amount of data that is transferred.

The connection between the camera and the environmental display could be over a short-range wireless link or through a wired connection such as USB. From a communication semantics perspective the camera is the master and the display is the slave, even though from a physical cabling perspective the camera may be a USB slave.
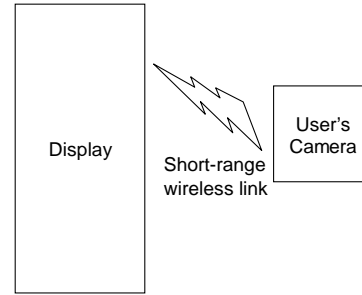

Figure 3: Direct Symbiosis

3.3. Fetching images from a web repository

The direct transfer approach works only when the higher resolution images are actually on the camera itself. It may be slow if the link between the camera and the environmental display is a low bandwidth link. Further transferring images in this manner consumes battery power on the camera, which is a precious resource.

When the images are present in a web repository, it may be better to enable the display to fetch these images from the web repository directly. This approach, shown in Figure 4, may be faster if the bandwidth between the display and the web server is higher than the direct link between the camera and the display. More importantly this approach does not move large image files over the battery powered wireless link from the camera. We expect that the user will still view the thumbnails privately on the small display in the camera before deciding which ones to view on the large display. Image selection will still be a problem if there are hundreds or thousands of thumbnails on the camera. We present approaches to mitigate this problem later.
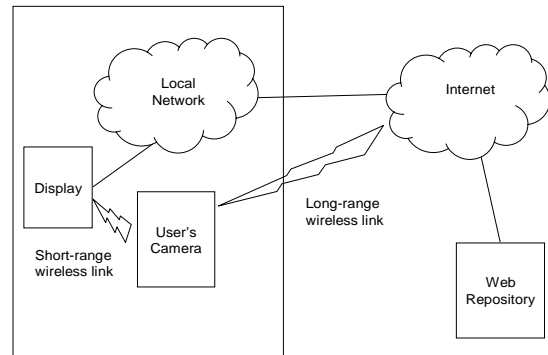

Figure 4: Indirect Symbiosis

However, once an image is selected the camera can direct the display to fetch the image from the web

repository. For this approach to work, the camera will need a way to authorize the transfer of the image from the web repository to the display. The camera can connect to the web repository over its long-range wireless connection and request the repository to create a freshly computed ephemeral token that the camera hands off to the environmental display over the short-range wireless interface. The environmental display then obtains the image from the web repository on behalf of the camera, proving its authority by supplying the token handed to it by the camera. The web repository rejects any requests without valid tokens. The entire mechanism may be simplified if the camera can send to the display a URL that includes both the image identifier as well as the authorization token. The web repository will usually not be able to initiate the image transfer to the display since the display is usually behind some firewall. The short and long range wireless capabilities of the camera are used in conjunction to use the public display to download and view selected images without compromising the user's entire image collection.

This same approach of fetching from a web repository can be used to share images. For instance, one may receive a message from a friend who has uploaded a collection of images that he wishes to share. The message may include thumbnails that arrive at the user's mobile device. While scrolling through the thumbnails in the message, the user can instruct the mobile to display selected thumbnails on the larger display. At this point the user's camera should delegate its authority to the display to enable it to fetch the images from the friend's web repository. Optionally, a copy of the image can be transferred from the display into the camera over the short range link.

4. Challenges

We consider some of the design issues in the following sections, namely the range of the wireless interface, the type of storage, security, usability, and power management issues.

4.1. Image management problem

Image management is going to be a problem with cameras with large amount of storage. This differs from storing thousands of songs on an MP3 player because songs have descriptive metadata such as names and artists, whereas images have to be annotated by the user. Since thousands of images can be stored, easy and effective naming schemes have to be developed instead of just sequentially numbering them. Voice based input could be used for naming groups of pictures. Users could assign names to locations and then the camera could convert the GPS coordinates of the image and organize them into viewlists associated with that location.

People do not typically annotate images because it is cumbersome. Several approaches are possible to make this task simpler [4, 9, 14, 17]. The first is to make annotation

simpler at time of capture by allowing voice labels that can later be converted to searchable text with the assistance of other devices such as PCs. Automatic means to capture the context in which a picture was taken could be central to building organized albums. Cameras can easily be enhanced to automatically record several parameters, such as location, the identity of the photographer captured using biometrics, etc. A different approach would be to build annotations on another device and transfer them to the camera. At present a large fraction of cameras have no notion of accepting image content or metadata from other devices. This shortcoming would have to be addressed as well.

Since users may keep copies of images on other devices as well and edit image albums on them, the images on the camera may not be synchronized with the images on the other device. Thus standard synchronization techniques may have to be adapted for digital cameras. If images are deleted on the web repository, the corresponding thumbnails have to be deleted on the camera as well.

As some images are expected to be ephemeral in nature, a mechanism to specify when those images can be automatically deleted needs to be provided.

Cameras may be replaced every few years, and therefore users will need to transfer images from the old camera into the new camera just as we transfer data between computers today. Adjustments may have to be made to account for the differences in the cameras such as image resolution viewfinder size and resolution.

Techniques to manage large collections of images on PCs [3] can be applied here as well.

4.2. Quick image/thumbnail search and retrieval

Clearly, as the amount of storage in cameras increases, quick image retrieval and image management will become a problem.

Several approaches are possible for aiding image retrieval. The first and preferred one is to use the input controls and software on the environmental display to compose a query and overcome the input limitations on the camera. The camera could send JavaScript code to the environmental display and present a form where the user selects search criteria. The search criteria includes all the metadata captured with the image, such as ranges for date, time, location, annotation sub-text, aperture, focus and flash settings, photographer identity, privacy levels, retention settings, etc. Other criteria such as, "Find images similar to the one selected," can be provided using research in image similarity [5]. The query is then sent back to the camera in an encoded form for completing the search. Another approach is to use embedded voice recognition technology to specify the query. A limited vocabulary recognition engine is sufficient for most parameters except for annotation strings. If our camera

incorporates a phone, use of voice for interaction is natural.

The fallback option is to use the camera input controls to search for the images. For example in the search mode, the flash, scene, privacy and retention dials could be positioned appropriately and the date range can be selected using the display on the camera. In our experience, specifying location is more difficult than most of the above search parameters we have listed. A map interface similar to the one available on world clocks to select time zones, with zoom capability works best to specify a location for our purpose. A combination of the multiple modes can be effective.

4.3. Security and privacy issues

Given that cameras may have thousands of personal images it may be necessary to encrypt the images or at least provide password and biometric means to protect the images from unauthorized access. Borrowing from USB storage devices that include fingerprint sensors and partition the flash into a public and a private part (see the USB key in Figure 2), we can divide the camera storage into public and private areas. The private section can be accessed only when the right biometric or password is provided. It may be useful to be able to indicate at the time of capture what level of privacy is associated with each image.

Since the camera will interact with unknown displays, it must be able to carry out challenge/response sequences to establish identity and authenticity of the display. Techniques described in [2] may also be applied.

The display on the camera serves two purposes in our case. The first is to preview images privately before deciding to view them on the larger public display. The second is to display information from the environmental display to verify its credentials and trustworthiness.

Network attached displays will also need to enforce access controls, verifying that only authorized users are able to use the display. If the display has a short range wireless interface or a direct wired connection, it may permit anyone who is able to connect to it via these connections to use the display. If the display is on an access controlled network, the display may permit anyone who is on the same network as the display to use its services.

Also the displays will need to support the notion of a user session where one authorized user has control of the display and any one else wishing to use the display has to wait until the session ends. User's must have the same level of comfort with such displays will not retain their content after use as they have with printers and speakers today.

4.4. Network and power issues

For maximum flexibility, the camera must include a wired, a short range wireless and long range wireless communication capability. The advantages of a wired connection include higher speed, lower power, and simplicity since device discovery becomes easier. However, wired connections often require users to carry cables and are prone to hardware malfunction due to incorrect attachment or mechanical fatigue. Moreover, with some displays such as ceiling mounted projectors it may be difficult to make wired connections. Short range wireless is useful to establish trust relationships also transfer images from the camera to the display. A long range wireless channel allows the camera to fetch images from other sources and send it to the display over the short range wireless channel. It also helps the camera obtain a fresh ephemeral token that grants permission to the display to images from the user's network accessible image repository. Long range wireless channels consume a fair amount of power for large images and also are slow. The long and short range channels can be used to establish trust between the camera and the display even if the camera cannot implement high strength security algorithms. The long range channel can be used to fetch a response to a challenge by communicating with a proxy the camera trusts.

5. Camera Architecture

In order to meet the requirements for the different settings, current digital cameras have to be augmented with features from other mobile devices. Our hypothetical cell-phone camera designed for effective symbiosis integrates a high resolution imager, a phone with a wide area 3G capability as well as a short-range wireless interface, low power storage, high density storage, and a location tracking mechanism. A wired USB connector for fast transfer of data while powered by a host device is also included.
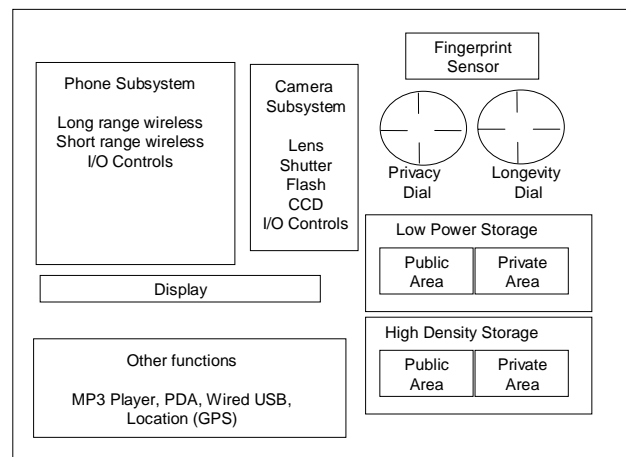

Figure 5: Camera Architecture

The storage segments have a public area and a private area. The contents of the private area are shared only when suitable authentication is provided. A fingerprint

sensor in addition to optional passwords is used for authentication. The fingerprint sensor is also used to record the identity of the photographer. The location sensing mechanism is used to record the location where the image was captured. In addition several image parameters such as the aperture, light condition, flash status, etc., are captured in the image [9]. A high density drive with a capacity of several gigabytes is included. This drive also has a private and public area. This drive is used as a secondary storage device and the flash is used as the primary storage device. The camera also includes additional dials to record the level of privacy of an image and this may determine which section the image gets stored in. Example choices for who can view the images could include immediate family, friends and family, and then everyone. Another dial on the camera is used to indicate how long the image is expected to be kept. The privacy and longevity values can be changed after the picture is taken if necessary.

The display on the camera is as large as feasible given the form factor. As explained earlier, the resolution of the display will be substantially smaller than that of the image captured by the camera. The camera includes voice input for recording annotations and embedded voice recognition technology for accessing images.

One of the problems with integrated devices is that if the device is turned off, all functions of the device become unavailable. For example if the cell phone needs to be turned off in an airplane for example, the camera function becomes unavailable as well. In order to overcome this situation and to save energy we believe independent power controls need to be provided for the camera, the phone, storage, and other subsystems.

6. Applications

We now consider several examples of how the overall architecture we described can be applied.

6.1. Collaborative image viewing

We first consider the example of sharing an image album with guests at a party in your house. The objective is to show the guests the pictures from your album but not show them any private pictures that could be embarrassing. The images are stored on a home PC that is connected to a projector through a high speed link. Our camera is used as a remote control.

An initial connection phase establishes connections between the camera, the PC and the projector. Issues around network boundaries and authentication are simplified since we are in a friendly environment where the devices trust each other.

In this situation the camera has a thumbnail view of the images along with the annotations and links to the actual images on the home PC. Even though the camera may have the complete images, it may be faster to fetch the images from the home PC and sustain audience interest.

The user can look at the thumbnail on the camera and decide the image is suitable for public viewing and tap on the thumbnail. The application on the camera then shows any notes the user might have associated with the image such as the date and circumstances under which the image was taken, e.g., in 1996 when Antonio was 4 months old. Private audio annotations may be delivered to the camera headphone directly. While the image is being shown on the projector, the user may advance the viewer on the camera and queue up other images to show. The user may skip any images he does not wish to show this audience.

In order to simplify the image selection process, the user may create *viewlists* ahead of time and then just launch the viewer. In this case the content is pre-fetched ahead of time if possible to reduce the delays in image loading. The viewlist can be updated on the camera based on comments received from the audience. If someone in the audience needs a copy of the image and is on the address book in the camera, a simple action can email the image from the home PC to that person.

6.2. Service applications

In our experience with home improvement projects we noted that several visits are required to the hardware shop in spite of taking lists of what was needed and making visual note of what is needed. Upon examining the nature of the visits carefully, we determined that some visits could have been avoided if we had a carried a picture of the carpet or tile of plumbing fixture that we needed to replace. The pictures could then be compared with the replacement parts to determine if the part available in the stores was suitable or not. Figure 6 shows an example. In other cases the image on the camera can be shown to the store assistant for further help in locating the necessary part. Another critical piece of information conveyed by the pictures is relative size. If the image records the auto-focus distance accurately, some size estimates can be made.

Figure 6: Furnace controls

We first look at the case where the complete image is stored in the camera and that the hardware shop has a kiosk where the images could be viewed in greater detail. A standard PC with image viewing software and either the ability to accept USB connections or wireless camera connections is sufficient. Once the image of the item is uploaded to the store kiosk, automatic image recognition techniques may be used to determine what the user is looking for and direct the user to the location of the item.

This example is applicable in many other scenarios. When asked by your hairdresser how you would like to have your hair cut, do you wish you had a full resolution picture of your desired style that you could show? A picture of the pantry, refrigerator and other shelves taken before grocery shopping could show how much of a particular item, e.g., milk can, is left. This process can be automated if refrigerators and pantries took pictures every day and automatically uploaded them to the camera. Picture viewing kiosks at the grocery store will allow the user to view the images in more detail. Such images have limited long-term value but it could be useful to have a time stamp on the image so the user can extrapolate and determine how much of an item might be left. In fact we believe older images can be overwritten when newer images are available.

Recently, one of our colleagues captured a picture of a dessert at a restaurant that looked appealing. When asked the reason, she replied she may show it to the waiter while placing an order at a future visit to the restaurant. Grocery lists could become image viewlists instead of just textual lists. Answer to some questions from service technicians fixing home equipment can be provided if images such as the one shown in Figure 6 are in the camera. Another example is taking pictures of rental cars before driving off the car lot. Two purposes are served – the first is to record any existing damages to the car and the second is to recognize your car more easily while on the trip.

When the date and location of the image have to be proved, we believe cameras can embed this information in the image and then sign them.

6.3. In vehicle symbiosis

Many people spend a non-trivial amount of time either in a car, train or airplane. How do digital pictures fit in these situations?

Typically cars are private spaces and in many cases the user is busy driving the car. Airplanes, trains and buses are public transportation and typically users are not in charge of driving and are in the midst of several people.

Let us consider travel in a car as an example. Many cars today have navigation consoles that show maps and other information. In addition, several high end cars have rear view cameras that show what is behind the vehicle when the driver is backing out. Similarly, some cars have front view cameras that monitor lane markers and signal the driver if he is straying away from his lane erratically.

Some cars have started including Bluetooth wireless communication within the car to facilitate operation of cellular phones without using hands. Voice based operation of some of the car's subsystems are becoming available as well. How can all the features in a modern car be exploited by users with digital cameras?

We built a prototype system [9] in 1997 that annotated pictures captured with a Ricoh-RDC1 digital camera with location metadata from a GPS that allowed us to augment textual directions with turn by turn images. The basic idea was to provide pictures of intersections, landmarks, exit signs, etc., along with textual directions. The images were captured ahead of time manually and added to an image database. The images were annotated with GPS coordinates and also with directional information. When directions between two places were computed, appropriate images would be fetched from our database and be included as hyperlinks in our system.

Given the developments in cars today, we can conceive that our camera could have retrieved images of appropriate intersections, signs, etc., before a trip and supply them to the onboard display. The short range wireless interface would be used to transfer the images from the camera-phone to the navigation console at the start of the trip. The navigation system could then show the pictures at appropriate junctures in the trip (see Figure 7 and Figure 8). The limited amount of detail that is visible in Figure 9 makes our point about the need for larger displays.

How will images of intersections be compiled? One could look at how input from thousands of users has allowed the creation of CDDB, a database of music album information that is available to millions of users around the world. Since music lovers receive a lot of information while converting CDs to MP3 format, they are willing to enter some information for the benefit of the community. Potentially, a similar model will work for creating an image database. Cars with cameras could capture images at important locations and then send them to driver's cell-phone camera. The images will contain the time of day, the direction of travel, location, and perhaps even the lane from which they were taken. The driver's could then upload the images to the navigational image catalog. This catalog can be used by routing software to retrieve images along with turn by turn directions. Fancier approaches will retrieve images that were taken in a similar seasonal setting, i.e., if directions are sought in September, images taken in the September timeframe will be retrieved for the turns. Time of day and weather conditions can be taken into account similarly. One practical issue we experienced is that since images are captured while the camera is moving, fast shutter speeds have be used to reduce motion blur. Other related work discussed in [4, 8, 15] can be combined as well.

Figure 7: View before a turn

Figure 8: View closer to the turn

Figure 9: Camera display view

7. Conclusions

We presented how some of the technology trends and limitations in human abilities may influence the evolution of the digital cameras and how they will be used even more extensively. We believe the hardware challenges will be addressed by the consumer electronics industry in the foreseeable future. The infrastructure may be much slower to adapt and certain segments will have to lead the way. We expect these expanded uses for digital cameras to slowly make their way into the lives of people worldwide.

8. References

1. S. Berger et al., "Using Symbiotic Displays to View Sensitive Information in Public," IBM Research Report, RC23274, July 2004.
2. D. Balfanz, et al., "Talking to Strangers: Authentication in Ad Hoc Wireless Network," in Proc. 9th Network and Distributed System Security Symposium (NDSS'02), 2002.
3. A. Girgensohn, J. Adcock, M. Cooper, J. Foote, and L. Wilcox, "Simplifying the Management of Large Photo Collections," Human-Computer Interaction INTERACT'03, IOS Press, 2003; pp.196-203.
4. L. Wenyin, S.T. Dumais, Y.F. Sun, J.J. Zhang, M.P. Czerwinski, B. Field, "Semi-automatic image annotation," In Proceedings of Interact 2001, Eighth IFIP, July 2001.
5. M. Flickner et al., "Query by Image and Video Content, " IEEE Computer, 28(9), pp. 23-32, 1995.
6. R. Kjeldsen, A. Levas, and C. Pinhanez, "Dynamically Reconfigurable Vision-Based User Interfaces," in 3rd International Conference on Vision Systems (ICVS'03), Graz, Austria, 2003.
7. A. Madhavapeddy, D. Scott, R. Sharp, E. Upton, "Using Camera-Phones to Enhance Human-Computer Interaction," In Adjunct Proc. of Ubicomp 2004.
8. M. Naaman, A. Paepcke, and H. Garcia-Molina, "From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates," 10th International Conference on Cooperative Information Systems (COOPIS), 2003.
9. C. Narayanaswami, S. Kirkpatrick, "System and methods for querying digital image archives using recorded parameters", US Patent 6,504,571, IBM Corporation, Filed May 1998, Issued Jan 2003.
10. C. Narayanaswami, et al., "IBM's Linux Watch: The Challenge of Miniaturization," IEEE Computer, vol. 35 (2), pp. 33-41, 2002.
11. C. Pinhanez, "The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces," in Proc. of Ubicomp 2001.
12. M. Raghunath, C. Narayanaswami, and C. Pinhanez, "Fostering a Symbiotic Handheld Environment," IEEE Computer, vol. 36 (9), pp. 56-65, 2003.
13. S. J. Ross, et al., "A Composable Framework for Secure Multi-Modal Access to Internet Services from Post-PC Devices," in Proc. of WMCSA, 2000.
14. R. Sarvas, E. Herrarte, A. Wilhelm, and M. Davis, "Metadata Creation System for Mobile Images", In Proc. of ACM MobiSys 2004.

15. K. Toyama, R. Logan, A. Roseway, P. Anandan, "Geographic Location Tags on Digital Images," Proceedings of the 11th ACM International Conference on Multimedia, pp.156-166, 2003.
16. P. Vettiger et al., "The "millipede" - nanotechnology entering data storage," IEEE Transactions on Nanotechnology, pp 39-55, March 2002.
17. K. P. Yee, L. Swearingen, K. Li, M. and Hearst, "Faceted Metadata for Image Search and Browsing," Proceedings of the Conference on Human factors in computing systems, ACM Press; pp. 401-408.