From Photography to Ubiquitous Capture Systems

Albrecht Schmidt, Bastian Pfleging, Christian Holz, and Lars Erik Holmquist

In the analog age, creating a photo involved two steps: taking a picture and then developing it. The user normally performed the first step using a camera, while a technician performed the second step in a lab. The invention of the charged-coupled device (CCD) lead to the development of digital cameras, which, just in the last decade, have replaced nearly all analogue cameras for consumer use. The improvements in digital camera technologies during this time have been astounding.

Cameras have become mobile computers with a capturing function. They include processing capabilities, wireless network connectivity, displays, and sensing. Digital cameras are often integrated into other devices, such as smartphones and cars. These developments, which are ongoing, allow for new use cases that go well beyond photos, picture frames, and printed albums.

Here, we reflect on these changes and how a ubiquitous computing technology replaced an "old" product—in this case, analog photography. The photography domain serves as an interesting use case of how innovations in ubicomp products can change an entire industry. At the same time, this domain also shows that innovations in a single product (the camera) can affect the entire ecosystem (how to store, access, and share photos).

A MULTITUDE OF IMPROVEMENTS

Camera sensors have made continuous progress on spatial resolution, size, speed, and sensitivity over the past decade, bringing cameras with over 40 megapixels to mobile devices (such as the Nokia Lumia 1020). Cameras now also provide increasing quality in the temporal dimension, which lets them capture photos in bursts—for example, 5–10 images per second at over 20 megapixels on high-end single-lens reflex (SLR) cameras—and over 1,000 frames per second at low resolution in consumer cameras (Nikon J1 offers 640 × 240 pixels at 400 frames per second (fps) and 320 × 120 pixel at 1,200 fps).

Along with increased resolution, cameras now include sensors, with an improved color and dynamic range. Photo applications can leverage these sensors to computationally integrate multiple photos taken with different exposure times into a single high-fidelity photo.

Computationally combining a series of photos has become a trend. An early example of this is GigaPan (http://gigapan.com), a camera robot with software that creates high-resolution images by combining many photos (stitching). In this procedure, photos are computationally aligned and colors are matched. Other applications use this approach to capture photos in poor light conditions or to remove objects from a photo. For example, Empty LA (http://emptyla.com) and the Empty America series (http://rossching.com/empty-america) are first examples of photos and videos of originally crowded scenes that don't show any passing pedestrians or cars. By manipulating a time-lapse series of photos, moving objects are discarded to create an atmosphere of an abandoned landscape. Although these examples have been created manually and for an artistic reason, we can imagine applications that automate this work.

A logical progression of 2D photography is capturing entire scenes in 3D. Following the tradition of analog cameras, manufacturers include two lenses to capture pictures from different viewpoints (for example, the LG Optimus 3D P920; also, 3D lenses are available for many camera types).

With light field cameras (such as the Lytro camera; www.lytro.com) photography moves from recording a 2D or 3D photo to capturing a light field. Using this approach, you can shift the focus point or perspective after the image is taken, because all the information is encoded in the "photo."

SENSING BEYOND THE VISUAL

The range of sensors that today's devices include lets cameras activate the trigger based on the image context or record contextual information in addition to the actual photo. For example, the StartleCam automatically takes a photo of moments of interests based on users' state of arousal, detected using physiological sensors.[1] Sensor data can also be used to augment photos with metadata—for example, Jennifer Healey and

Rosalind Picard presented a sensor box with an accelerometer and temperature sensor that was attached to a digital SLR.[2]

The most common metadata currently associated with photos are time, orientation, and location data, because most current cameras have an orientation sensor and more manufacturers are incorporating GPS into their cameras. Today's mobile phones have an even wider range of sensors, including camera sensors, compasses, and assisted GPS. Common use cases where sensors are used include capturing panoramic scenes that also records the direction in which the picture is taken. Additional metadata provides value to users by helping them navigate and search their photo collection for specific content.

Users with large photo collections benefit from additional sensor data, which lets them detect interesting trends and patterns and derive new information. For example, Yahoo's TimeTraveler application uses the location and time stamps of Flickr photos to find the most popular locations in a city and calculates the time required to see them (http://travel.yahoo.com/timetraveler). With growing numbers of available geotagged photos, users can identify previously unknown patterns. For example, researchers derived the names and boundaries of neighborhoods from photo collections by mining manual user tags and geolocation in Flickr photos.[3]

While many existing photo apps on smartphones, such as Instagram or Flickr, offer artistic filters to augment captured photos, sensors could be used to complement the visual experience of taking a picture. An example is the Context Camera, which "captures the invisible" in addition to photos by sensing motion and sound to influence the photos' look. Picture taken in a loud environment would look different than the same scene in silence, so users become more engaged with the photo-taking experience and take photos in
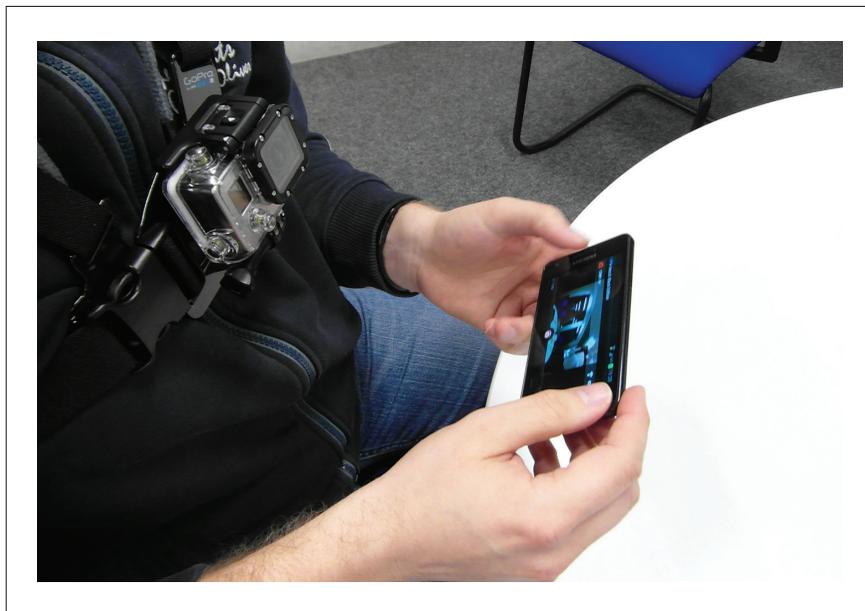

Figure 1. One current trend in cameras is to include networking, which separates the camera and user interface.

situations they otherwise might not have captured.[4]

SEPARATING COMPONENTS

We're currently observing a trend in which phones become alternative user interfaces for cameras. Figure 1 shows the typical use of a Wi-Fi-connected phone as the interface for a GoPro camera. The Sony QX10 and QX100 explores the idea of further separating the lens from the user interface for still photography by omitting a display on the camera element entirely. The phone provides interactive controls and connects to the lens using Wi-Fi.

As this trend progresses, we expect future cameras to start acting in concert. Wireless connectivity will continue to enter cameras and thus allow for remote operation and image and data transfer among a variety of devices. Linking networked cameras is the basic idea in the social video platform Vyclone (http://vyclone.com). It lets users simultaneously record multiple views of a scene with different cameras. The software then automatically syncs the multiple perspectives into one seamless stream.

NEW APPS AND OPERATING SYSTEMS

Smartphones are reintroducing the notion of scripting on cameras. Although the Kodak DC290 provided scripting capabilities in 2001 using the Digita Operating Environment (http://lisas.de/digita), programmable cameras didn't become popular. However, smartphones have changed this, with countless camera apps appearing in the App stores for iPhone and Android. Many provide functionality similar to original cameras and add filtering capabilities, image manipulation, and sharing options.

Many developers are intrigued by creating their "own camera" through software on a phone or on Android-based cameras (such as the Samsung Galaxy Camera and Polaroid SC1630 Smart Camera). Surprisingly, high-end SLRs still don't offer this option and don't provide frameworks and App Stores to create and share applications for this hardware. An example of an open source project that aims to extend camera capabilities through additional software is magic lantern (www.magiclantern.fm), which grew

Figure 2. Camera mounted to a Draganfyler X6 drone. (Photo courtesy of DraganFly; www.draganfly.com.)

out of users' desire to improve the camera's video capabilities. Magic Lantern targets audio and video limitations of Canon SLRs. Similarly, the Canon Hack Development Kit (http://chdk. wikia.com) enhances the capabilities of consumer cameras with features from SLRs and with scripting capabilities.

TAKING A NEW PERSPECTIVE
Because users operate traditional cameras through the physical viewfinder, such cameras take photos sharing the user's perspective. Digital cameras, in contrast, enable completely new perspectives. The Canon G1, as an early example, introduced (in the year 2000) a swivel display, which lets users more easily take photos from the floor, from overhead, or of themselves.

Mobile cameras have taken this concept even further. Due to their size, such cameras can be mounted onto moving objects, such as balloons or flying drones (see Figure 2). Manually controlled "flying cameras" are a first step, and we imagine exciting photos being taken by cameras following a person engaged in an activity, such as when climbing a wall. Waterproof cameras, such as the Nikon AW1 or the Olympus Tough series, also offer new perspectives.

Strapping cameras to animals, as demonstrated in Cat@Log,[5] brings about an unfamiliar perspective and takes photography to places where humans rarely

go. Of course, for this use case, devices must be rugged and provide metadata, such as location, so viewers can better understand the photos.

PRESSING THE SHUTTER: ANALOG AND DIGITAL
Analog photography directly links the shutter button to exposing the film to light, thereby defining the precise moment when the picture is taken. Cameras with autofocus and automatic exposure often have a two-level shutter; the first level causes the camera to prepare the capture, setting the aperture and focus. Pressing down to the second level then takes the photo.

In digital cameras, the shutter button remains the central user interface element, but its meaning has broadened. Some digital cameras incorporate a two-level shutter reminiscent of analog cameras. At the first level, the camera starts continuously filling a buffer with images. At the second level, it takes the picture. However, digital cameras additionally enable a shutter lag of zero or even a negative value (known as the "pre-capture mode" on the Samsung Smart Camera EX2F). The picture that "is taken" is then retrieved from the buffer, which might be the photo from the exact moment when the person wanted to trigger. This reactiveness creates a positive user experience, but experienced photographers, who will expect a lag, need to adapt to the new

usage model. Similarly, the Nikon Smart Photo Selector Mode stores a selection of images around the capture moment based on composition and motion.

If we take this concept further—for example, by assuming that the camera continuously records full resolution pictures at all time—the shutter button could become only a marker in a video stream. This would let users move through the stream temporally, similar to zooming in on a photo.

Automatic triggering is also available on many cameras or with add-on hardware. Examples are computer vision algorithms on the camera that trigger the capture (for example, take the photo when something enters the photo or starts smiling).

CROWDS AS SENSORS
Online photo repositories now contain countless photos for popular locations, such as the Golden Gate Bridge or Eiffel tower, including many directions and angles. Assuming we have access to these photos and their meta information, a shutter press could simply record the location, direction, angle, and time of day, and the camera would then retrieve photos matching those details. In particular, this approach might make sense when the conditions for taking the photo aren't optimal, such as the Golden Gate Bridge covered in fog. The camera could then virtually "see through" the fog.

SHOEBOX OF DIGITAL PHOTOS?
In contrast to the speed of development in digital cameras, the storage of photos remains difficult for most users. A recent discussion with computer science students interested in photography revealed that most have no sophisticated long-term strategies for storing and organizing their photos and that most of them had lost some digital photos.

The common approach to secure photos seems to be a mix of sharing through social networking sites, and storing photos on hard-drives, on DVDs, and using online back-up facilities. Large cloud storage space could unify this but
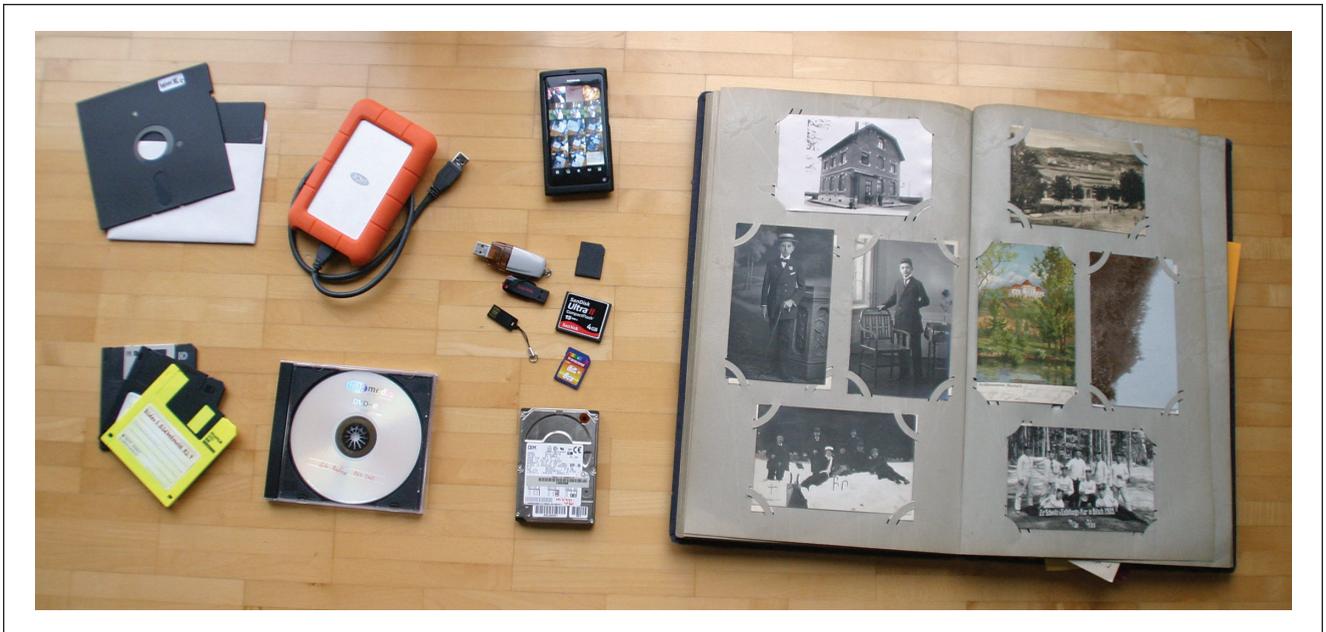
Figure 3. The short-lived digital storage solutions are a stark contrast to a traditional album of photos.

requires users to fully trust the quality of the service. Although such services seek to establish trust, the recent discussions around users' privacy in online systems might negatively affect users' willingness to move personal photos online.

Linked to the question of how to store photos is how to share them with others. Photo storage sites as well as social networks typically let users share photos with family and friends. However, maintaining control over the pictures in the long term is far from simple. Safekeeping personal photos is a central issue, because photos are often connected to people's memories. Difficult questions arise, such as "Who will inherit my photos when I die?" and "How will the transfer be done?" What's the digital equivalent to a shoebox or album full of photos (see Figure 3)?

This reflection on the change from analog to digital photography and the discussion of upcoming trends highlights the effect of pervasive computing technologies on well-established products and services. Looking around us, there could be similar changes ahead in

other industries. Computing is driving innovation in traditional products and redefining their capabilities. ℙ

REFERENCES

1. J. Healey and R. Picard, "StartleCam: A Cybernetic Wearable Camera," *Proc. Second Int'l Symp. Wearable Computers* (ISWC 98), IEEE, 1998, pp. 42–49.

2. P. Holleis et al., "Adding Context Information to Digital Photos," *Proc. 5th Int'l Workshop on Smart Appliances and Wearable Computing* (ICDCSW 05), vol. 5, IEEE CS, 2005, pp. 536–542; doi: http://dx.doi.org/10.1109/ICDCSW.2005.33.

3. B. Thomee and A. Rae, "Uncovering Locally Characterizing Regions within Geotagged Data," *Proc. 22nd ACM Int'l Conf. World Wide Web*, ACM, 2013l http://dl.acm.org/citation.cfm?id=2488500.

4. M. Håkansson et al., "More than Meets the Eye: An Exploratory Study of Context Photography," *Proc. Fourth Nordic Conf. Human-Computer Interaction*, ACM, 2006, pp. 262–271.

5. K. Yonezawa, T. Miyaki, and J. Rekimoto, "Cat@ Log: Sensing Device Attachable to Pet Cats for Supporting Human-Pet Interaction," *Proc. Int'l Conf. Advances in Computer Entertainment Technology*, ACM, 2009, pp. 149–156.

Albrecht Schmidt is a professor of human-computer interaction at the University of Stuttgart. Contact him at albrecht.schmidt@vis.uni-stuttgart.de.

Bastian Pfleging is a researcher and PhD candidate in the human-computer interaction group at the University of Stuttgart. Contact him at bastian.pfleging@vis.uni-stuttgart.de.

Christian Holz is a research scientist in human-computer interaction at Yahoo Labs. Contact him at christianh@yahoo-inc.com.

Lars Erik Holmquist is an independent researcher and an innovator. Contact him at lars.erik.holmquist@gmail.com.