The Composition Context in Point-and-Shoot Photography

Daniel Vaquero and Matthew Turk
University of California, Santa Barbara*

With the recent popularization of digital cameras and cameraphones, everyone is now a photographer, and the devices provide new opportunities for improving the process and final results. While there has been research on what kinds of subjects users prefer to photograph, and what they do with the images once they are captured [Van House et al. 2005], no formal studies on the process of framing an image using a camera have been performed. To fill this gap, our study attempts to characterize the actions performed by users while framing photos using a point-and-shoot camera, in preparation for taking a photograph. This includes adjusting the camera’s orientation and point of view and triggering zoom and autofocus controls.

1 User Study

We implemented a camera application based on the Frankencamera architecture [Adams et al. 2010], running on a Nokia N900 smartphone. It mimics the interface of a point-and-shoot camera in “automatic” mode, with a digital viewfinder, automatic metering algorithms, and controls for zoom and autofocus. A sensor box containing accelerometers and gyroscopes is attached to the camera. The camera silently records viewfinder frames (at 25 fps) and their capture parameters, as well as inertial sensor data, for the immediate 18 seconds before the user triggers the shutter to capture a photograph. We call the viewfinder frames, their capture parameters, and inertial sensor data collected while the user is framing a photograph the composition context of the photograph.

We have recruited nine volunteers so far to participate in the user study. The study is in progress, and more users will be added. The sessions were conducted at the UCSB campus, with variations in times of day and weather conditions. Participants were provided with our camera and asked to take three pictures for each of seven different categories that represent common situations: an office environment, a close-up scene, a building, a sign, an open area, a posed picture of a person or group of people, and a moving subject. They were also instructed to try their best to capture compelling images, relying on their own sense of what makes a good picture. Once these 21 pictures were taken, the users were requested to take at least five additional pictures of scenes chosen at their discretion.

A preliminary analysis of the collected data indicates interesting characteristics. In this analysis, we considered only the composition context frames that overlap with the scene captured in the final photograph. For a total of 255 pictures, the average duration of a composition context was of 10.17 s. The camera was typically held still, with small movement due to handshake, during short intervals before capture, interrupted by sharp translations or rotations, due to attempts to adjust the composition; this effectively constitutes variations in field of view. 108 photos had exposure variations of at least one stop in their composition contexts due to autoexposure, 54 photos had at least 10 diopters of variation in focus due to autofocus, and zoom was used for framing 61 photos. Moving subjects were present in the composition context of 172 pictures.

2 Implications

We plan to use the results of this study to help create new camera functionality. From the user’s point of view, the camera would preserve the interface of current point-and-shoot cameras, and it would still be used in the same way. However, when users trigger the shutter, instead of obtaining a single photograph, they would also obtain additional photo suggestions from the same scene, created through automatic image alignment and combination of composition context frames captured under different parameters (such as panoramas, collages, extended dynamic range, selective focus, synthetic long exposure, synthetic panning, motion trajectories, and moving object removal), or selection of alternative views by optimizing computational aesthetics measures. As the variations in capture parameters in the composition context happen naturally in the point-and-shoot process, no additional input or familiarization with new interfaces would be required from the user. We expect this capability to expand the photographic abilities of casual and amateur users, who often rely on automatic camera modes, without changing the widespread point-and-shoot paradigm.

References
