Evaluation of the metric performance of mobile phone cameras

Armin Gruen and Devrim Akca
Institute of Geodesy and Photogrammetry, ETH Zurich, CH 8093 Zurich, www.photogrammetry.ethz.ch

With the availability of very affordable and good resolution CCD/CMOS cameras and other off-the-shelf devices, including compact computers (of laptop type) the potential of building efficient mobile, low-cost, high-performance systems has substantially improved. Among various hardware and software components mobile phones constitute an interesting component for image data acquisition for obvious reasons: They are inexpensive, light and handy and have CMOS cameras integrated of currently up to 10 Mpixels image format.

Their usability for many applications has recently been realized. Some of the most prominent applications are the character/text recognition, facial animation, human face identification, image-based context awareness, content providing to Information Systems.

In spite of the availability of a broad diversity of applications, the metric capabilities and characteristics of mobile phone cameras have not been investigated so far. In 2004, Sharp Corporation developed a 2 Mpixel CCD camera module with 2X optical zoom and auto-focus function (Figure 1a) intended for use in mobile phones. In 2005, they released two new camera modules (Figure 1b) with a 3 Mpixel CCD chip. One year after, Samsung announced a 10 Mpixel camera phone (Figure 1c) at CeBIT exhibition in Hannover. These examples show the rapid progress in the technology of mobile phone cameras.

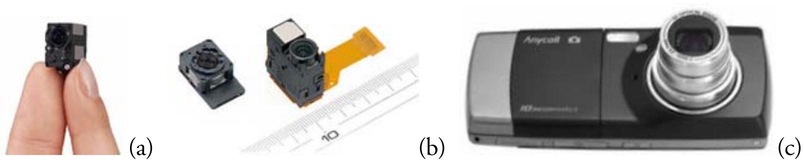

Figure 1. (a) Sharp's 2 Mpixel CCD camera module LZP0P3738, (b) Sharp's 3 Mpixel CCD camera modules LZ0P3751 and LZ0P3758, (c) Samsung's 10 Mpixel camera integrated mobile SCH-B600.

Due to the very limited size and restricted material and equipment costs, the production of mobile phone cameras is a challenge. The impact of their production specifications on the stability of interior orientation and 3D object reconstruction capabilities has not adequately been studied. This work investigates the accuracy potential of two recent mobile phone cameras and compares them with respect to two off-the-shelf digital still video cameras.

Four cameras are used (Figure 2). Two of them are mobile phone cameras (Sony Ericsson K750i and Nokia N93) and two of them are off-the-shelf digital still video cameras (Sony DSC W100 and Sony DSC F828). The mobile phone cameras have CMOS sensors of smaller size than the CCD chips in the off-the-shelf cameras and partly much smaller lenses.

Figure 2. Cameras used in our tests: (a) Sony Ericsson K750i, (b) Nokia N93, (c) Sony DSC W100, (d) Sony DSC F828.

The photogrammetric calibration field at the Institute of Geodesy and Photogrammetry (HIL C57.3, ETH Zurich) was used. It is 3.4 x 2.0 x 1.0 m3 in size. The 3D coordinates of 87 well distributed control points (GCP) were measured using a Leica Axyz system. The average precision values of the GCPs are ±0.03, ±0.05 and ±0.03 mm for X, Y and Z axes, respectively.

For the calibration of the K750i, eighteen images from three locations (each of which has three stations, i.e., down, middle and up) were taken in a convergent geometry mode. Nine images are taken in normal mode and the rest of them are rotated. For the calibration of the N93, W100 and F828 cameras an image acquisition geometry with thirteen images was used. First nine images are in normal mode and the rest are the rotated ones. The image measurements were performed with the Least Squares template matching (Gruen, 1985) using the in-house developed software BAAP. Another in-house developed software SGAP (Beyer, 1992) was used for the bundle block adjustment with self-calibration. The K750i, N93 and W100 have only the JPEG output option. Their image measurements were carried out on their original JPEG images. For the F828, TIFF output images were used for the image measurements.

In spite of giving the worst results in the test, the K750i still can offer sub-millimeters accuracy in object space. Both block-invariant 10 and 44 additional parameter sets cannot compensate the systematic errors fully. The first three cameras of the test K750i, N93 and W100 give identical standard deviation values for the image observations (between 1/4 - 1/5 pixel). They all apply a chip level image enhancement for sharpening the images. This low level image enhancement, while improving the visual quality, is probably reducing the geometric quality of the cameras. They show noticeably block-variant systematic errors after the self-calibrating bundle adjustment with block-invariant additional parameters. The N93 and W100 have same lens systems (Zeiss, Vario-Tessar). The W100 has a CCD sensor of larger size with 8 Mpixels. It is 2.5 times larger than the CMOS sensor of N93. According to theoretical expectations, the N93 should give an accuracy of factor 1.6 (sqrt(2.5)) worse compared to the W100. The N93 almost strictly meets this expectation by giving 1.7-1.9 times worse numbers than the W100. On the other hand, there is a large difference between those two cameras, considering the size of the imaging system and the cost of the materials used in the construction. In this respect, the accuracy performance of the N93, as compared to the W100, is noteworthy. Although the W100 and F828 have the same image format with 8 Mpixels, the expectation of equal accuracy does not hold here. The W100 gives substantially worse accuracy numbers (almost 3 times) than the F828. This is mainly due to a better lens system of the F828 and (possibly) the degrading chip level image enhancement operation of the W100.

The use of JPEG images for the image measurements of K750i, N93 and W100 raised the question whether the JPEG compression has an effect on the results. The original TIFF images of the F828 were converted to quality level 100 (maximum quality) and 70 JPEG images, using the free software IrfanView (version 3.98, http://www.irfanview.com/). Even when going up to a factor of 42 compression rate we did get only a small reduction in accuracy (9% in depth direction). This can be considered harmless.

We believe that with a proper calibration and data processing software performance these devices can be used for many photogrammetric tasks which require an accuracy of around 1:10 000. The integration of GPS receivers and motion sensors will further broaden their applicability. Also, it is to be expected that the quality and performance of the integrated cameras will further improve, together with the on-board processing functions. This may allow one day such a device to be used as a stand-alone photogrammetric data acquisition and processing tool, at least for smaller projects.

6th ARIDA Workshop on "Innovations in 3D Measurement, Modeling and Visualization"
Centre for Scientific & Technological Research - B. Kessler Foundation, Povo (Trento), Italy, February 25-26, 2008