Why Feature Extraction is Hard

Christian Heipke

"An image speaks more than 1,000 words." This well-known Chinese saying describes much of the fascination of photogrammetry and remote sensing, especially of the images involved. But is it really true? Imagine two persons looking at the same image, say an aerial image of a town. What do they see? The first one might recognize a building, perhaps a house in a residential area with a red roof and a green yard. The other person may instead call this same building his home, the place where he was raised, and immediately recognize the door handle, because he used to bump his head on the handle when he was a kid. In point of fact, neither the door nor the handle may actually be visible in the aerial image.

Why are these two descriptions of the same image so different? The reason is that when we look at an image, we do not only see what is presented to us, instead we relate the features in the image to our background knowledge, our memories, and our experience about the scene in general (a house) and special objects in particular (the former home). Based on such aspects, it is of course difficult, if not impossible, to arrive at general conclusions about anything. Thus, we need to find a sort of compromise between our internal, subjective view of the depicted scene and the image and a more external, objective view, shared by others. Moreover, both views must be represented in a suitable, computer-readable way. The latter task has proven to be a formidable challenge, indeed, and is far from being solved even after a few decades of research and development in computer vision and digital photogrammetry.

In addition, humans often construct missing parts, for instance the door handle mentioned above. Consider too that a sketch of a cube on a piece of paper is generally seen as a three-dimensional cube.

In order to carry out feature extraction, single objects depicted in the scene must be recognized and described. This recognition assumes prior knowledge of objects as models, which first of all should be made available to the computer. The production of the object models is a major challenge in itself. Research has shown that both geometric and radiometric information on the various objects is necessary. For aerial imagery, the larger the scale of the images to be analyzed and the more details are required, the more important is geometric information, as one enters closer to the domain of human activities, which can be characterized by linear borders, symmetries, right angles, and other geometric aspects. For smaller resolutions, however, radiometric and spectral attributes dominate, which explains the good results of multispectral classification for satellite images of coarser resolution as well as the inferior results of the same technique for high resolution satellite and aerial images.

While image analysis is a very challenging and inspiring field of research—an example of a tree extraction result is shown in **Figure 1**—fully automatic systems do not seem realistic in foreseeable future. Semi-automatic procedures, which integrate the human operator into the entire evaluating process however, are being used successfully. A person remains responsible for tasks which require major decisions (*e.g.*, selection of algorithms and parameter control), quality control, and—where required—the correction of intermediate and final results. Thus, the best of both worlds can be put together for the benefit of exploiting images faster, more objectively, and sometimes also more accurately.

About the Author

Christian Heipke is a professor for photogrammetry and remote sensing at the University of Hannover, Germany. His main research interests lie in the area of automatic image analysis and geometric aspects of photogrammetry. He can be reached at *heipke@ipi.uni-hannover.de*.

Figure 1 *Two results of an automatic tree extraction from multi-spectral images and a digital surface model in several stages of resolution*