New software advances photo search and management in online systems

October 15 2007

Searching for digital photographs could become easier with a Penn State-developed software system that not only automatically tags images as they are uploaded, but also improves those tags by "learning" from users' interactions with the system.

"Tagging itself is challenging as it involves converting an image's pixels to descriptive words," said James Wang, lead researcher and associate professor of information sciences and technology (IST). "But what is novel with the 'Tagging over Time' or T/T technology is that the system adapts as people's preferences for images and words change."

In other words, the system can accommodate evolving vocabulary and interpretations to images that people have uploaded and are uploading to systems such as Yahoo's Flickr. This allows the T/T system's vocabulary to grow, replacing old tags with more relevant and more specific new tags, Wang said.

In tests, the T/T technology correctly annotated four out of every 10 images, a significant improvement over the researchers' earlier annotation system, ALIPR or Automatic Linguistic Indexing of Pictures-Real Time. That system offered users a list of 15 possible annotations or words for an image-one of which was correct for 98 percent of images tested.

"The bottom line is that the system makes it easier to find photographs and is able to improve its performance by itself as time passes," said

Ritendra Datta, a graduate student in computer science working with Wang. "The advancement means time savings for consumers as well as improved searching and referral capabilities."

The system was described in a paper, "Tagging Over Time: Real-world Image Annotation by Lightweight Meta-learning," presented at the recent ACM Multimedia 2007 conference in Augsburg, Germany. The authors were Datta; Dhiraj Joshi, a former graduate student in computer science; Jia Li, associate professor, Department of Statistics; and Wang. Penn State has filed a provisional patent application on this invention.

In the researchers' previous system, pixel content of images was analyzed to suggest annotations. In the new software, researchers have added a machine-learning component that enables the computer to learn from the user's interactions with photo-sharing systems.

Images of the World Trade Center, for instance, once would have been tagged with "financial center," "business" or "market success." Users today, however, have different associations or tags for the Twin Towers. The researchers' new system has the capability to learn from such changes and reflect them automatically in refining the old tags and generating future tags.

As the system adapts to such changes, it also enhances tagging performance. With an initial accuracy of 40 percent, the system's precision improves over time and can reach a level of up to 60 percent correct, Datta said.

In a companion paper, "Learning the Consensus on Visual Quality for Next-Generation Image Management," which also was presented at the ACM conference, the researchers described a new system which can automatically select "aesthetically pleasing" images and isolate out those images of low or poor quality. To do this, the system uses visual features

such as contrast, depth-of-field indicators, brightness and region composition from publicly rated photographs to learn the statistical models for high- and low-quality images.

"With this system, users can more easily identify the best photographs in their collections," Datta said. "The system also suggests images which should be deleted from the digital cameras to make storage space for new photographs, for example."

The system can also improve image search engines by prioritizing visually pleasing images among the search results, Wang added. The National Science Foundation supported research on both systems.

Source: Penn State