Picture this - automatic image categorisation

May 3 2005

Creating, storing and transmitting visual images has become increasingly easy. Yet the same problem always arises – how to categorise or classify visual images automatically without using external metadata or image thumbnails? There now may be an answer.
Researchers in the IST project LAVA have developed a method of automatically categorising the content of digital images, providing an effective means of storing and retrieving digital image content without having to rely on additional metadata. The developed techniques should be of real value for document and content management, making it much easier for users to search for images as well as text.

At the beginning of the project, the participants faced two major technological challenges: how to interpret and categorise efficiently large numbers of objects, scenes and events in real settings, and how to automatically acquire knowledge of categories for the convenient construction of further applications. To overcome them, they brought together two key groups of researchers.

“Our main idea was to bring together researchers from the machine-learning community with those from the computer-vision and cognitive science areas,” says LAVA’s Gabriela Csurka of Xerox Research Centre Europe (XRCE) in Grenoble.

“We began our approach by grouping together similar types of objects, such as bicycles or cars for example, and trying to find a way of categorising those that were common to a group,” says Csurka. “We applied machine-learning techniques to find the distinctions between

images by focusing on sections of images that were similar – sections that were common to other images with the same content." Researchers also needed to overcome the challenge of categorising image content despite views of objects from varying perspectives or from different distances.

A typical example of the challenges they faced was how to draw a distinction between an image of a car, and one of a stack of car tyres. Both picture types contain 'patches', or sections, of images that are the same. To overcome the problem, the team had to provide the system with the ability to examine key patches in other areas of the image. In this case, the software was programmed to check for other key content in the 'car' image, such as headlights or windows.

Real advance on earlier methods

At the close of the project at the end of April 2005, LAVA researchers have developed an integrated method of capturing visual images and identifying, automatically, the appropriate category for any captured objects or scenes, be they people, objects or simply landscapes. This confluence between machine-learning and vision interpretation has greatly enhanced the ability to build reliable vision-based detectors for everyday objects and events, they believe. Such systems can underpin novel applications of all kinds, including location identification and the description of meetings.

"We believe that we now have the state-of-the-art in image categorisation and event interpretation," Csurka says. "Our system does not rely on the whole shape of an image, but on local patches or parts of the image with similar geometric properties. So it is more versatile – we can cope with much larger intra-class variations and still correctly interpret the image. It is a real advance on what went before."

Underlining their achievement, the LAVA team, represented by the Gravir-INRIA laboratory and the University of Southampton, won 14 out of 18 competitions in detection, localisation and classification in the Visual Object Classes Challenge organised by the PASCAL network which emerged from LAVA's work in recognition of the importance of the field and the maturity of the existing research. The challenge itself aimed to compile a standardised collection of object recognition databases and provide a common set of tools for accessing and managing database annotations.

According to Chris Dance, project coordinator and head of XRCE laboratory, the techniques developed in LAVA will certainly be taken forward. "We will be working with Xerox business groups to integrate this new system into Xerox' document management offerings, making them pioneering products in this field and providing our customers with additional competitive advantage."

Potentially the developed technologies could be used for browsing images within documents, archiving images and managing photo for consumers, and searching for images on the Web. But they could also be applied to video surveillance, human-computer interaction, medical imaging and robotics.

Source: IST Results