Automated image analysis arises from handcraft and machine learning

May 24 2012

The amount of visual information increases with tremendous speed. The archives of television networks, image bank databases and social media in the web are all bursting with billions of pictures – and more is produced by the second. In order to organise these heaps of data and to find wanted information from it, the analysis of the images must be automatised.

In his recent doctoral dissertation for the Aalto University Department of Information and Computer Science, Ville Viitaniemi has studied methods for image analysis that are based on detection of visual categories.

"The content of images can be discerned and classified in countless ways. For a computer to know how to recognise and interpret images, it is useful to dissect them into prescribed categories," explains Viitaniemi.

The general task of automatic visual recognition and analysis has persisted throughout the existence of computers. Instead of presenting the computer an open question of what is in a picture, the computer is better off solving a bunch of small sub-tasks in which the images are dissected into categories. By choosing the right categories and combining them, the contents of images can be increasingly more accurately described.

"In my dissertation I look by experimentation for an efficient system for recognising visual categories."

Splice, recognise, fuse

The general mathematical model for recognising images is yet to be presented, and Viitaniemi says any such model would presently be computationally too heavy. The human brain on the other hand is not well enough known at the systemic level in order its mechanisms for visual recognition to be imitated.

"For now, the only method that works is 'an engineer's approach': to try to figure out which parts of the system, organised in which way, produce adequate results."

The three basic steps of the top-performing system of visual category detection are feature extraction, detection of the features, and the fusion of the results of the detection. In his research Viitaniemi strived to find the most efficient ways to execute these phases.

"First, the images under inspection are extracted of certain features such as colours, textures and shapes. Then the detection system is taught by methods of machine learning to detect the features from images. When a group of features have been detected, a fusion of the results follows," sums up Viitaniemi the process of visual analysis.

A bag of visual words into a support vector machine

For the extraction of features Viitaniemi wound up to prefer a method called Bag of Visual Words. A single image is broken down to 100–300 meaningful locations, after which the neighbourhood of each location is given a specific visual description.

"For each neighbourhood, a histogram is collected of the directions of its surrounding gradients. This way a useful feature is put together. A

feature characterising an entire image can then be created by looking into the statistics of the distribution of the local features."

The refined bags of visual words go into a support vector machine, which has been taught to recognise whether a feature belongs to certain category or not. Fed enough features, the machine will know whether it is a bird or an aeroplane on the sky of a picture.

"Different methods have to be experimented with, because a few successes in recognition tasks do not guarantee reliable performance. As long as we are not able to imitate the methods of image recognition of the brain, the best way is to experiment and experiment, through trial and error."

Provided by Aalto University

Citation: Automated image analysis arises from handcraft and machine learning (2012, May 24) retrieved 10 April 2024 from
https://phys.org/news/2012-05-automated-image-analysis-handcraft-machine.html