

Putting pictures into words

September 16 2008

(PhysOrg.com) -- Visual images can contain a wealth of information, but they are difficult to catalogue in a searchable way. European researchers are generating and combining scraps of information to create a searchable picture.

Digital images can open our eyes to the most extraordinary detail and beauty. But there is one major drawback. The information in an image is purely visual. It tells us nothing about when or where the image was taken. It tells us nothing about the people in the image. We do all that interpretation ourselves. That makes cataloguing and searching for particular images difficult – whether you work for an art gallery or you are updating your family album.

Scientists on a major European research project called aceMedia have taken important steps towards a solution to this problem. They are building an information layer into digital image files. Their vision is that image files will come with content information, metadata (background information for use on the internet) and an intelligence layer that automatically generates word-searchable data about the image.

An extra ‘information layer’, that adds both automatically generated and manually generated information to images, would revolutionise image searching on the internet as well as on your home computer or mobile phone. The technologies developed in the EU-funded aceMedia project have sparked interest from a range of commercial companies, looking to exploit the ideas in a host of directions.

Building a picture puzzle

The project re-used, developed and combined a series of technologies that provide greatly enriched content information about an image.

One of the technologies exploited by an aceMedia team uses software that can identify low-level visual descriptors, such as consistent areas of colour that may be sky, sea, sand or possibly snow, and information about texture, edge and shape.

Combining the low-level descriptors with sets of contextual rules held in domain ontologies (such as the fact that consistent areas of blue at the top of an image are likely to be sky, or that beach and snow are unlikely to appear in the same picture) turns data into a rich information source.

“Turning low-level descriptors into useful information is a very difficult step,” according to Yiannis Kompatsiaris, Head of the Multimedia Knowledge Laboratory at the Informatics and Telematics Institute in Thessaloniki, Greece and one of the lead researchers on aceMedia. His team was involved in structuring knowledge and adding it to the domain ontologies that classified and identified the information provided by the low-level descriptors.

Data from low-level descriptors was also combined with the results from specific detectors, such as the kinds of face detectors that are commercially available on some cameras today. All add further clues or searchable data for image users.

Another layer of information can be added by the individual user. They can add rules defining their personal preferences, profiles and policies to create a personalised filing system. ‘Inferencing’ techniques, filtering and similarity algorithms were used to make that personal filing simpler.

To enable easier searching, some of the aceMedia researchers incorporated natural language processing techniques into the mix, which mean you can use everyday language when searching for an image.

The ace in your hand

AceMedia researchers drew together their full range of technologies in an Autonomous Content Entity (ACE) Framework. The ACE Framework defines APIs to support networking, database management, scalable coding, content pre-processing, content visualisation, knowledge-assisted content analysis, as well as context analysis and modelling modules.

Using the framework, ACEs can be created that contain all of the rules, metadata and content information. They become a part of the image file.

For video, aceMedia researchers developed a scalable video codec, the aceSVC. Pical scalable video coding chain consists of three main modules – an encoder, extractor and decoder. The aceSVC enables video playing, reviewing and video analysis in the compressed domain.

As part of the project, aceMedia researchers demonstrated the benefits of automated content sharing and easier content management that ACEs could provide on a series of home network devices, including PCs, mobile phones and set-top boxes.

While the vision of the aceMedia project was to combine technologies, each delivering a piece to the overall information puzzle, they are not interdependent according to Kompatsiaris.

“The tools we developed in aceMedia are scalable to many concepts and many environments,” he confirms.

Adding time and location

“In five years time, a good number of these technologies will be in common use – combined with a number of technologies that have grown in popularity since the aceMedia project started, such as geo-tagging using GPS receivers. I think cameras in the future will know their position and be able to combine that information with content analysis to give much better results than we are capable of at the moment. For example, if the camera knows it is in a mountainous environment then it can analyse the content of the image much more efficiently.

Provided by [ICT Results](#)

Citation: Putting pictures into words (2008, September 16) retrieved 2 May 2024 from <https://phys.org/news/2008-09-pictures-words.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.