

See what I see -- machines with mental muscle

October 1 2008



Showing the multi-dimensions of the recognition technology. MUSCLE logo

(PhysOrg.com) -- The way we use and interact with machines is undergoing a profound change as computers are programmed to learn from experience and see more how we see. European research into machine learning is pushing back the boundaries of computer capabilities.

Computers do not see things the way we do. Sure, they can manipulate recorded images, for example, but they currently understand very little about what is inside these pictures or videos. All the interpretation work must be done by humans, and that is expensive. But one European project is making computers more similar to us in their ability to interpret images and their surrounds.

Individuals from all walks of life, as well as sectors such as industry, services and education, stand to reap immense benefits from semi-autonomous, more intuitive machines that are able to do things which were, until now, either not possible, super expensive or the preserve of humans.

This has been made possible thanks to the developments in, and convergence of, methods for creating, obtaining and interpreting metadata - at its simplest level this is data about data, or facts about facts - in complex multimedia environments.

MUSCLE, an EU-funded super project which created a pan-European network of excellence involving more than 30 academic and research institutions from 14 countries, has come up not only with new paradigms but a range of practical applications.

Vast scale

The scale of the project was so vast, a special section to showcase its achievements has been set up in the 3D Second Life internet virtual world, which has millions of denizens.

The Virtual MUSCLE experience inside Second Life has been created as a one-stop information centre to ensure the continuation and sustainability of the project's achievements. Users are impersonated as avatars (computer representations of themselves) enabling them to experience multimedia content by literally walking through it. They are able to hold real-time conversations with other members of the community, exchange experiences, or just simply browse.

After an initial two years of collaborative research across the MUSCLE network, a series of showcases were established with several institutions working together on each one to produce practical applications.

Virtual tongue twister

One of these is an articulatory talking head, developed to help people who have difficulties in pronouncing words and learning vocabulary. This ‘insightful’ head models what is happening inside the human mouth, including where the tongue is positioned to make particular sounds, so the users can copy what they see on screen.

A second showcase functions as a support system for complex assembly tasks, employing a user-friendly multi-modal interface. By augmenting the written assembly instructions with audio and visual prompts much more in line with how humans communicate, the system allows users to easily assemble complex devices without having to continually refer to a written instruction manual.

In another showcase, researchers have developed multi-modal Audio-Visual Automatic Speech Recognition software which takes its cues from human speech patterns and facial structures to provide more reliable results than using audio or visual techniques in isolation.

Similarly, a showcase which has already attracted a lot of publicity, especially in the USA, is one that analyses human emotion using both audio and visual clues.

“It was trialled on US election candidates to see if their emotional states actually matched what they were saying and doing, and it was even tried out, visually only of course, on the enigmatic Mona Lisa,” says MUSCLE project coordinator Nozha Boujemaa.

Horse or strawberry?

Giving computers a better idea of what they are seeing or what the inputs

mean, another showcase developed a web-based, real-time object categorisation system able to perform searches based on image recognition - photos including horses, say, or strawberries! It can also automatically categorise and index images based on the objects they contain.

In an application with anti-piracy potential, one showcase came up with copy detection software. “This is an intelligent video method of detecting and preventing piracy. There is a lot of controversy at the moment about copyright film clips being posted on YouTube and other websites. This software is able to detect copies by spotting any variation from original recordings,” Boujemaa explains.

“Another application is for broadcasters to be able to detect if video from their archives is being used without royalties been paid or acknowledgement of the source being made.

Europe’s largest video archive, the French National Audiovisual Archive, has now been able to ascertain that broadcasters are only declaring 70% of the material they are using,” she tells ICT Results.

Other types of recognition software, effectively helping computers see what we see, can remotely monitor, detect and raise the alarm in a variety of scenarios from forest fires to old or sick people living alone falling over. The latter falls under the heading of “unusual behaviour” which also has applications in video security monitoring with “intelligent” cameras able to alert people in real time if they think somebody is suspicious.

“During the course of the project, we produced more than 600 papers for the scientific community, as well as having two books published, one on audiovisual learning techniques for multimedia and the other on the importance of using multimedia rather than just monomedia,” she says.

Although the massive project has now wound down, its legacy remains online, in print and most of all in a host of new applications that will affect the lives of people all over the world.

Provided by [ICT Results](#)

Citation: See what I see -- machines with mental muscle (2008, October 1) retrieved 25 April 2024 from <https://phys.org/news/2008-10-machines-mental-muscle.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.