

# Robotic perception, on purpose

October 26 2009

---



(PhysOrg.com) -- European researchers developed technology that enables a robot to combine data from both sound and vision to create combined, purposeful perception. In the process, they have taken the field to a new level.

Currently, computer vision is good at recognising objects in images and videos and has been successfully employed in several specialised industrial applications, such as quality control during microchip fabrication.

But robotic perception is much weaker in less defined situations, like understanding and responding to human behaviour and even conversations. Yet, it is precisely this sort of interaction which promises the most compelling applications for future humanoid technology, where people-like robots can act as guides, or mix with people, or use

perception to infer appropriate actions.

More importantly, these broad robotic applications will deliver insights into other disciplines, like cognition and neuroscience.

A truly perceptive robot, capable of acting independently and appropriately in complex situations remains a distant goal, but European researchers brought it much closer with their Perception-on-Purpose (POP) project.

## **Original, by design**

“The originality of our project was our attempt to integrate two different sensory modalities, namely sound and vision,” explains Radu Horaud, POP’s coordinator.

“This was very difficult to do, because you are integrating two completely different physical phenomena,” he adds.

Vision works from the reflection of light waves from an object, and it allows the observer to infer certain properties, like size, shape, density and texture. But with sound you are interested in locating the direction of the source, and trying to identify the type of sound it is.

## **Tricky issue**

On its own, sound is difficult to pinpoint, because it needs to be located in a 3D space. Then there is the problem of background noise, such as an open window letting in sounds from next door.

But it turned out that integrating two different senses helped the researchers in their bid to locate and tune into relevant sounds.

“It is not that easy to decide what is foreground and what is background using sound alone, but by combining the two modalities - sound and vision - it becomes much easier,” reveals Horaud.

“If you are able to locate ten sound sources in ten different directions, but if in one of these directions you see a face, then you can much more easily concentrate on that sound and throw out the other ones.”

## **Integrated technology**

This was one approach that the team took and, with the algorithms they developed, their robot, called Popeye, was able to identify the speaker with a fair degree of reliability.

“There is more work to be done on that aspect of the work, it is not completely robust yet,” warns Horaud.

Still, it was a very strong result, and what makes it even more impressive is that the team managed to integrate all the technology into a neat and compact robotic platform.

“Most often, sound research is conducted in specialised labs, with arrays of microphones and a very controlled acoustic environment. But we integrated our two microphones and two cameras onto the head of our Popeye. The idea is to have an agent-centred [cognitive](#) system,” Horaud stresses.

## **Powerful technology**

The Popeye packs a lot of powerful technology into a small space and offers purposeful robotic perception. This is important because Horaud

argues persuasively that, in evolutionary terms, multi-sensory perception and cognition are linked.

By perceiving a hand-held object with their two eyes, for example, monkeys - and the first hominids after them - developed stereo vision and hence were able to learn many properties of an object from combined tactile and visual data. Over time, they developed new skills, including building tools, from this information.

Horaud feels, too, that some modern uses of artificial intelligence (AI), like chess applications, are limited because they do not learn from their environment. They are programmed with abstract data - say, chess moves - and they process that.

“They cannot infer predicates from natural images; they cannot draw abstract information from physical observations,” he stresses.

For now, POP has achieved many of its aims and developed very promising approaches. Commercial applications for this type of technology are not out of the question, and the researchers also hope to continue their work in a further project.

That project would look at extending some of POP’s results into a functioning humanoid robot. In the meantime, POP’s work means that the purposefully perceptive [robot](#) has become a not-so-distant future technology.

More information: [perception.inrialpes.fr/POP/](http://perception.inrialpes.fr/POP/)

Provided by ICT Results ([news](#) : [web](#))

Citation: Robotic perception, on purpose (2009, October 26) retrieved 17 April 2024 from <https://phys.org/news/2009-10-robotic-perception-purpose.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.