



Mobile phones are no longer simple communication tools, but sophisticated portals granting entry to the digital world. Consumers expect these pocket-sized units to offer multiple, high-quality functions, but performance is often inhibited by processing power. To improve their effectiveness, scientists are developing techniques to deliver media with greater fidelity, supporting the development of next generation electronics.

# Expanding the audio and video capabilities of handheld mobile devices

**Portable telecommunications** have undeniably transformed 21st century cultures, and global interactions within them. Indeed, according to the International Telecommunications Union, they may soon outnumber people – with a potential 7.3 billion mobiles in circulation by 2014. Such a vast market creates incredible opportunities for savvy science, with users clamouring for novel features, improved speed and errorless multimedia content delivery.

To meet these demands, a recent EU funded project, AVID-MODE, has fostered synergies between business knowhow and academic expertise. Inaugurated in 2009, the four year initiative has undertaken extensive research into mobile technology, which anticipates the sector's future evolution. Clearly, a phone's voice communication function is now merely one of myriad, expanding arrays of utilities. Customised via downloads and upgrades, a

phone can be easily transformed into a bespoke, all-in-one platform, delivering GPS services, innumerable apps, music and streamed video content. "Mobile devices play an ever greater part in our everyday lives, so consumers have higher expectations of their capabilities" observes the Project Coordinator, Dr. Athanasios Mouchtaris, an Assistant Professor in the Computer Science Department at the University of Crete and

allowed us to explore different, yet interconnected research themes" details Mouchtaris. "One is audio signal processing. We aim to create highly detailed, immersive rendering, and optimise sound reproduction" Conducted in Crete, experiments into this subject required important foundational work into sound source separation and localisation – which help identify and contextualise the origin of the human voice.

## "We've developed a method for using mobile devices to transmit a convincing, 3D sonic environment"

Affiliated Researcher at the Institute of Computer Science (ICS) of the Foundation for Research and Technology-Hellas (FORTH). "People want to see, for example, high quality TV relayed by these devices."

During the 1.3 million Euro venture, four organisations – Greece's FORTH and the University of Thessaly, Italian electronics

firm T.W.S. Electronics and Chinese multimedia specialists Cidana –

undertook an extensive collaboration. "Academics

approach problems in a certain way" perceives Mouchtaris. "They often tend to be rather inclined towards theory and research whereas, industrially, pragmatic solutions are favoured. Because of the different interests involved in AVID-MODE, the project's outcomes attained a tangible, practical dimension, which is more relevant than a purely hypothetical approach."

"Engaging with a number of partners, including manufacturers and software engineers, has

"We've developed a method for using mobile devices to transmit a convincing, 3D sonic environment" reports Mouchtaris. "This technique uses an array of microphones. Employed, for example, in the context of a meeting, it allows individual speakers to be clearly discerned, and their physical positions audibly recreated." The information is encoded in a low bit rate signal, and transmitted to the listener. In their earpiece, they receive a detailed sound reproduction that emanates from different directions, emulating the space in which the speakers are situated. "Of course, you can already recreate a 3D environment using headphones or loudspeakers," notes the Greek researcher, "but not only did we create an algorithm to realise this effect, but also low-cost hardware to ensure it can be made affordable on a mobile."

"We encountered several fundamental challenges during the project" he relates. "Significantly, mobile devices lack the capabilities of a regular PC. Processing power is reduced – so we needed to design algorithms that could operate in real time within this environment, which was testing.





Similarly, we faced limitations on the rendering of sound and video, and also on energy requirements. Batteries must be used frugally, and not monopolised by a single task. It's critical that the unit should continue to operate as a phone, whilst other functions run simultaneously."

At the University of Thessaly, identical constraints applied to another team, whose efforts were concentrated in the visual domain. "Our starting point was the variety of codecs used primarily by video applications" reveals Dr. Ioannis Katsavounidis, an Associate Professor at the institution. "These are essential to video transmission. Without them, we'd need vast bit rates, which are unsustainable via broadcast or streaming methods. Common codecs include older variants like MPEG2, and newer types like MPEG4 and, also, AVS." Although they continue to improve, there's much debate about their respective merits.

Katsavounidis and his associates undertook a meticulous analysis to compare the performance of MPEG2, H.264 and AVS codecs, conducting months' worth of computer-based trials, using a variety of content, including film, animation and news programming. "The results will surprise both academia and industry, at least in terms of comparison between H.264 and AVS" he says. "The Chinese AVS codec presents a very good case – it exhibits compression efficiency within 10% of H.264, yet is quite simpler. Consequently, it seems generally more apt for mobile devices." Also, he adds, usage of AVS, unlike some of its better known comparators, does not incur royalty fees or

license payments. MPEG2 – the oldest of the trio being assessed – often fell into third place but, intriguingly, triumphed in at least one category, displaying less vulnerability to errors than its peers. The next radical progression in codecs may well arrive when H.265 – only published at the start of 2013 – takes off, believes Katsavounidis. "It's anticipated to make a big splash – potentially as large as the arrival of Blu-Ray" he muses.

Hand gesture recognition is another, hitherto underutilised technology that the group believes is primed for greater exploitation in mobile systems. The man-machine interface, they assert, should be far more dynamic than the humble touch-screen. "We decided to concentrate on devices with only one camera" confirms Dr. Katsavounidis. "It's possible to create a two camera solution using stereoscopic vision, which allows you to access depth information – and thus additional parameters and accuracy. Using one low quality, SD camera makes the challenge more difficult – but helped us to generate a solution which is universally applicable, and can operate in minimal lighting."

To be deemed truly effective, the device must achieve a high rate of recognition (around 95% is, at present, state-of-the art for the industry). Impressively, the group has surpassed this benchmark – devising a system with over a 99% recognition rate, which is soon to be presented at the ICASSP conference in Canada. "I'm really proud that we've pushed the envelope so far" says Katsavounidis. "Many other systems are very complex, or require algorithms that cannot even be computed on a mobile device. Our approach, however, is very computationally efficient, and achieves a pragmatic balance between these considerations."

Within twelve months, the AVID-MODE collective anticipate that the next generation systems portended by their trailblazing work will start to penetrate the mainstream. If industrialising their own discoveries may be a little further off, Katsavounidis is optimistic that their progress could yet herald a local renaissance. "Using this work as a foundation, I'm hoping some of my students can start their own companies" he confides. "Greece is in a dire financial situation, so sparks of progress like this are very exciting, and may not only help change the way we ultimately communicate, but offer new ways to reinvigorate the economy". ★



## AT A GLANCE

### Project Information

#### Project Title:

AVID-MODE: Efficient Location-Aware Audio-Visual Delivery of High-Quality Content to Mobile Devices

#### Project Objective:

The objective of AVID-MODE is the enhancement of high-quality video, sound, and television delivery capabilities of mobile devices that will enable them to provide interactive and immersive media applications. The project addresses limitations to current audio and video coding methods which restrict the use of mobile devices to offering only a low-quality streaming media presentation to the user.

#### Project Duration and Timing:

48 months, June 2009 to May 2013

#### Project Funding:

European Commission's Seventh Framework Programme (FP7), Marie Curie Industry-Academia Partnerships and Pathways (IAPP), total funding 1,323,240 Euro

#### Project Partners:

- Coordinator: Foundation for Research and Technology – Hellas (FORTH) (Greece)
- University of Thessaly (Greece)
- TWS Electronics (Italy)
- Cidana (China)

## MAIN CONTACT



### Prof. Athanasios Mouchtaris

Prof. Athanasios Mouchtaris received the Ph.D. degree from the University of Southern California (USC), Los Angeles, in 2003. From 2003 to 2004 he was a Postdoctoral Researcher in the Electrical and Systems Engineering Department of the University of Pennsylvania, Philadelphia. Since 2007 he has been an Assistant Professor in the University of Crete and an Affiliated Researcher in FORTH-ICS.

#### Contact:

Tel: +30-2810-391753

Email: mouchtar@ics.forth.gr

Web: [www.ics.forth.gr/~mouchtar](http://www.ics.forth.gr/~mouchtar)  
[www.avid-mode.eu](http://www.avid-mode.eu)

