Picture by John Tuffen

Future listening

AES London Convention speech looks at ways to enhance music understanding. Simon Duff reports

Editor's comment
Dave Robinson

AS *PSNE* GOES TO press, on this day, 5 August 2010, it seems we've finally got some of the info we've been after for, well, months and months.

If you're a channel 69 wireless user, you're going to receive compensation when you come to buy your new kit. If you're not on 69 now, you get nothing.

According to campaign group Save Our Sound UK, this represents 6% of the relevant frequencies, and is far from what the government is calling a "generous" funding scheme for reallocation. "To call this package 'generous' is an abuse of the English language," to quote SOS UK.

So can we find anything positive from this financial body-blow to the PMSE sector? Maybe. Today, 5 August, Ofcom revealed that "roughly 55%" of the cost to eligible ch69 users will be reimbursed. SOS UK hurriedly put a statement together at *PSNE*'s request, but didn't want to be drawn into making concrete quotes. And rightly so: they need time to examine the small print.

But let's look at it this way: if you received a substantial discount when you purchased your wireless systems, and the government are compensating at 55%, then you may get most of your cash back. Which is better than a kick in the teeth.

Here's something very important to bear in mind though: it would seem there is a certain element of a 'first come, first served' structure to the scheme, so don't dawdle on 23 September: get your application in smartish. Meanwhile, Ofcom have agreed to let *PSNE* readers know exactly what's involved: so make sure you check out the feature next month. ☙

Masataka Goto, leader of the Media Interaction Group at the National Institute of Advanced Industrial Science and Technology (AIST), Japan, gave this year's keynote speech at AES London in May.

The title of his talk was 'Music Listening in the Future'. Goto's mission is to help shape a future where music listening for a general public can be more active, more immersive, richer and deeper by using music-understanding technologies.

He explains: "Music-understanding technologies based on signal processing have the potential to enable new ways of listening to music for everyone. At the AIST we have developed nine different types of Active Music Listening interfaces, working with MP3 files running on PC, to demonstrate how end users can benefit from new technology. Active Music Listening aims at allowing the user to better understand the music they listen to and to actively influence the listening experience."

SmartMusicKIOSK is a computer programme that has a chorus-search function enabling users to access directly their favourite part of a song (and to skip others) while viewing a visual representation of its musical structure with the focus on chorus ('hook') sections. Technically the system automatically finds chorus sections by utilising a new signal processing method to analyse repetition of chord progression and melody line. With a patent approved on SmartMusicKIOSK in Japan, the US and Korea, Goto is in the process of seeking company partners in the audio world and believes that users could be computer consumers as well as those in the mobile devices marketplace.

Another area of development is Lyric-Synchroniser, which can automatically synchronise song lyrics to a recording and highlight the phrase being sung. A user can easily follow the current playback position and click on a word in the lyrics to listen to it. Goto has invented a method that allows users to extract vocals from existing stereo or mono files, and, with Hiromasa Fujihara and Hiroshi G Okuno, has developed a method that synchronises such extracted vocals with lyrics.

"Given polyphonic sound mixtures taken from available music recordings, our interfaces

Masataka Goto at AES London

thus enrich end-users' music listening experiences, and open up new ways of music listening in the future."

Music playback visualisation has also been realised. Animated synchronisation with music is made possible via a program called Cindy, in effect a virtual dancing tool driven by beat tracking. It can automatically synchronise motions and positions of computer-graphics dancers with musical beats. Users can enjoy different dance motion moods by changing them during music playback through pressing buttons.

Further software includes a music touch-up program called Drumix, developed with Kazuyoshi Yoshii and Okuno. This is an audio player with real-time drum-editing functions allowing users to change and edit the volume and timbre of drum sounds in existing music recordings and add new drum patterns during music playback. These functions can be achieved by an automatic drum sound detection method for polyphonic music recordings. VocaListener, developed with Tomoyasu Nakano, similarly aims at allowing users to change the timbre of existing vocal recordings. It is a singing synthesis system that automatically estimates pitch and dynamics parameters of commercial singing synthesisers such as Yamaha's Vocaloid so that the synthesised voice can mimic an existing vocal recording. Although the voice comes from singing synthesisers, it is human-like and very natural.
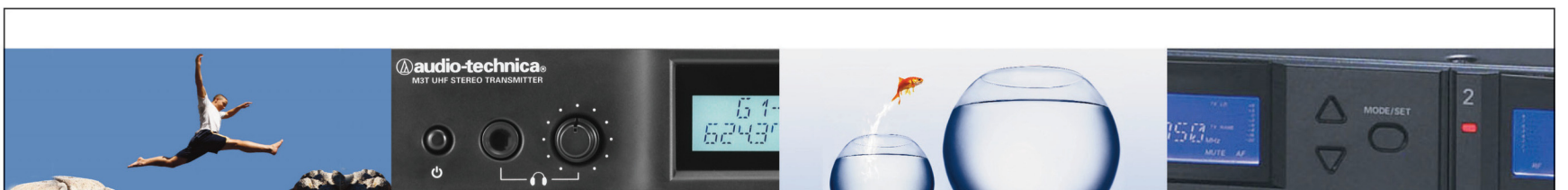
The last interface that Goto discussed during the lecture is Musicream, developed with Takayuki Goto. In essence, this is a discovery system developed with the intention of being connected to a flat rate, all you can hear music subscription service in the future. "The idea behind Musicream is to see if we can break free from stereotyped thinking of how music playback interfaces must be based on lists of song titles and artist names. Instead of listing up titles, disc icons representing music pieces flow one after another from virtual supply taps, and the user can select a disc and listen to it. By dragging a favourite disc in the flow, the user can easily pick out other similar pieces to create a playlist."

Located at Tsukuba Science City, just outside Tokyo, the AIST is the biggest national research institute in Japan. Goto leads researchers at the Media Interaction Group of AIST. At present most interfaces that have been developed are not commercially available but the group believes they can stimulate a new type of music technology industry and intends licensing its core technologies and patents to companies in the near future.

Goto is clearly determined and passionate about the possibilities that he is trying to open up. He concludes: "If listeners want to improve their music-understanding listening ability, methods to realise this have not yet been established. We aim to change this. Looking to the future I hope to develop new technologies around music-understanding ability training interfaces and singing information processing. It will involve areas that go further than ever before in combining signal processing and machine learning technologies and data from the internet. That is my next big research area. It will be amazing!" ☙

http://staff.aist.go.jp/m.goto/

✛ *AES Convention 129 is being held in San Francisco from 4-7 November.*