Austin Journal of Robotics & Automation

Austin Publishing Group

Open Access Full Text Article

Editorial

Shazam for People?

Vic Grout*

Vic Grout, Institute for Arts, Science and Technology, Glyndwr University, UK

***Corresponding author:** Vic Grout, Institute for Arts, Science and Technology, Glyndwr University, Plas Coch Campus Mold Road, Wrexham, North Wales, LL11 2AW, UK

Received: August 26, 2014; **Accepted:** October 10, 2014; **Published:** October 10, 2014

When Shazam first arrived on the scene, it was pretty amazing stuff; now, we rather take it for granted. But could the same idea soon work for people*?*

We know the scenario… You're in a bar or a shop or listening to the radio or TV… or… just about anything really… and you hear a song that you either like or think you recognise or both… but you don't know what it is. Frustrating, isn't it? At least it *was* until music identification services such as Shazam first appeared. After that, no worries; simply allow you're mobile to listen to the music for a few seconds, search the central database and, after a few more seconds, it reports back to you with full details of the name, artist and origin. It might even link you to a central library where you can find more of the same or possibly buy it.

Simple enough but, might the same principle one day work for *people*? It's really not that hard to imagine… Consider a future scenario: not too many years distant… You meet someone in the street. They introduce themselves as "John Green". As they're speaking, the headset you're wearing (maybe an implant) scans their face and analyses their voice and breath. It matches this against a global database and reports back to you…

"No, this isn't John Green. This is Paul White. He's 45 years old and lives in Sheffield; married with three children. He was arrested in 2003 for shoplifting and declared bankrupt in 2006. He works as a landscape gardener but his attendance record isn't very good. He smokes and has a chronic lung condition, which is making it difficult for him to get insurance. He votes Liberal Democrat …"

Then: *"Ah, but this is only the* free *stuff. If you're prepared to* pay*, I can tell you a lot more about him… "*

It sounds like a science fiction 'Big Brother for Everyone' nightmare scenario. But could it happen? If so, how soon?

We should probably start with an overview of how *Shazam*, and similar music identification services, work. It's clever and simple but perhaps a little *cleverer* and not *quite* as simple as it seems. Obviously, most such systems employ the microphone of whatever device is being used to gather a short sample of the music to be identified (for *Shazam*, it's about *10* seconds) and compare it with a central database (via a WiFi, 3G or 4G connection). But here's the thing…

We can't rely on a conventional 'one-size-fits-all' binary coding for the sample (or the database) because the same piece of music would have to be identified in *all* environments, through *all* media, in different formats, raw or compressed, at any volume and with various levels of quality or noise. Slight changes in the reproduction of a track will create a completely different binary code and it won't match.

What's needed instead is an 'impression' – an acoustic fingerprint of the music… Just as a human fingerprint has a unique combination of features (line breaks, connections, etc.) that remain the same even if the finger is stretched or squashed (or if you can only see a bit of it), so music should have some characteristics that can be recognised no matter what the underlying encoding is, or what part you're listening to. This isn't the same as simply identifying a tune by, say, humming it – there are indeed some apps that will do that; instead we're looking for a precise *instance* of a song here but accepting that there will be variations depending on the medium. For human fingerprints, we don't try to use bitmaps of the image because these would be all over the place and hard to compare; however, vector representation variants *will* identify the essential features. We need something similar for music.

Typical *Acoustic Fingerprints* will look for key features such as the music's tempo, prominent tones, and the rate at which the signal crosses the zero point, frequency/bandwidth or other 'spectral' characteristics. (Shazam identifies music fragments through spectrograms.) If this can be made accurate ('deterministic') enough – and quite obviously it *can* because it works, then the track fragment can be accurately matched against a database of millions of tunes to identify a particular recording, by a particular artist, at a particular time, in a particular place, etc. Still quite impressive really, isn't it, even though we're used to it?

So how might this work for *people*? What would it take to identify and profile someone you meet in the street? Well, let's start by sketching out the process…

1. Obtain a 'sample' of the person of interest to generate a *personal fingerprint*

2. Match this personal fingerprint against information in an online personal database

3. Return the match and associated information, possibly divided into 'free' and 'pay' categories

Two things are immediately obvious: 1 and 2 are interdependent and 3 are trivial. In other words, we need sufficient accuracy ('resolution') in 1 to allow for a unique match in 2 but this depends on the amount and quality of information available for both. However, if we can crack this then returning the result to the end user – and perhaps profiting by it – is straightforward (so we won't discuss it any further). So how close are we now and where are the gaps in existing technology to make this happen? Let's consider the key components…

Obtaining a Personal Sample/personal Fingerprint

So how much material can we 'get' on a stranger in the street? Well, we have to start by realising that it's not all about a *single type*

Citation: Grout V. Shazam for People?. Austin J Robot & Autom. 2014;1(1): 4.

of feature. Face recognition, for example, may be in its infancy and working on limited scales (although growing in both senses all the time) but we wouldn't be looking to use just that on its own. There are a variety of ways in which we could collect essential data from an individual and we'd combine it *all* to form the personal fingerprint – just like *Shazam* does for an acoustic fingerprint. It would depend a lot on the circumstances (what the person was *doing*, how close they were, etc.) as to what information we could get our hands on but the following are ways that we *might* look to identify people:

- Voice analysis
- Gait analysis
- Body size and shape
- Age estimation
- Breath analysis
- DNA
- biometrics
- Odour
- Clothes or uniform
- The car they drive
- Any identifiable technology on (or *in*) their person
- Location, occupation and known habits
- Association with colleagues, friends and family.

(Each of these techniques is a very real area of research in various stages of development – none are science fiction.) Oh, and *ordinary* fingerprints, of course. All this in addition to the face recognition we originally suggested. Naturally, some of these would have to be kept *very* up-to-date in the central database (see next section).

It's likely that, at least for the foreseeable future, none of these will provide a perfect solution in isolation but together they could well provide effective tools for individual identification. If we can just get a (sufficiently) deterministic fingerprint from a combination of these features then that would be enough. *"She's 5'7", approximately 35, estimated BMI of 27, she's wearing a chef's cap and she's in Luton, UK; she looks, smells, walks and talks like Penny Jones and she's with someone who matches (99.993067%) the personal fingerprint of her boss. She's carrying Penny Jones's mobile and credit card. We can say it's Penny Jones with 99.999994% accuracy."* And that's just with the 'uncertain' stuff – any 'real' biometric information will be even more precise and deterministic.

Of course, with several billion people in the world compared to several million songs, the requirements of a deterministic fingerprint will be far tougher for people than music. However, there's potentially a lot more to go on. Although some of the personal features suggested above are essentially 'fuzzy' in nature, a number are very precise and we can expect the resolution of all of them to improve over time. A crude calculation would suggest that we need the determinism of a personal fingerprint to be three or four orders of magnitude better than an acoustic one (although it then also depends on the quality of the database). This is a lot, certainly, but either a steady advance in the accuracy of *all* these measurements or a significant breakthrough in *one or two* would achieve it. Knowing, as we do how technology improves and advances, it would be unwise to bet against this.

So, given time, the *personal fingerprint* thing may be doable. What about the database to match it against?

Building up a Personal Database to Match Personal Fingerprints Against

This might seem straightforward on the surface but there are complications in the detail. We all know that the information 'on us' online is growing by the day. (Search for yourself on the Internet.) True, not all of it should be there and some of it's wrong and what can and can't be legally shared is something of a grey area; but we can't deny that there's plenty of information to be made available to anyone that wants it – or might be prepared to pay for it. This is already true of some conventional Internet 'services'; UK county court orders, for example, are publicly available and, for a price, the 'big data' corporate model can make an individual's available to you. Similarly, if you know someone has a criminal conviction – but you don't know what for, it only takes a few minutes intelligent searching to find out. (There are websites that collate this data – particularly for paedophiles and suchlike. Mind you, the information could be *wrong*, of course; Twitter, for example is easily the source of the most up-to-date *and* unreliable information on the Internet.) An automated process would achieve much more, much quicker. Naturally, there are huge legal issues here but the law has an abysmal record of staying relevant to emerging technology: there will always be loopholes – and it's hard to imagine *effective* legislation against automatic, intelligent web-searches reporting on what they find.

The *actual* problem is that most of the information that's available at present is the sort of material that would be supplied *back* to the user of the '*Shazam for People*'service. It's what would be *returned* in step 3 of the above process … the easy bit. What's needed is the information – and a fair *quantity* of it – that would allow for deterministic matching of the personal fingerprint. It's fair to say that, save for a few specialist applications for 'specialist' people (eg, criminal or terrorist databases) most of this isn't yet publicly available. There's no world-wide face or gait recognition system, no global index of shapes, sizes, locations, jobs, etc. Actually *identifying* people with accuracy and confidence isn't easy. Not only has that, but it all had to be kept up-to-date too, of course. At present, this looks like the biggest 'gap' in the process – the worst 'flaw' in the proposed system.

But realistically, how much longer is it likely to stay that way? There are numerous small systems already in existence in isolation. It's really just a question of linking then all together. Once again, there should be some laws to protect us here but history tells us that they *won't* when push comes to shove. The legalities of data are complex and have to change constantly to keep up with what technology allows us to do; generally they do this too late. The other point is that once personal data has been 'leaked' to the Web, it's almost impossible to get back. Prosecuting transgressors isn't much use to an individual whose personal data has been exploited and it doesn't even seem to be much of a deterrent in practice.

The other thing we know from experience is that digital

technologies don't have to be *perfect* to be released these days. We might never build a global system that was *100%* accurate but that won't stop trial versions appearing; once they do, improvements will follow quickly from real testing in the field. Of course, this again brings into play the issue that information returned could be *wrong* but we'll discuss that in the final section. It might even be worth considering the use of intelligent learning software to improve performance.

So, building up a database of personal fingerprints may be the biggest challenge – much of it legal – in this entire but, looking back at the way the global data field has evolved over the past decade; it's hard not to see it happening eventually.

So the Future is …?

So, putting all this together – certainly in the long term, there doesn't appear to be much to stop this. A reasonable estimate for something *like* '*Shazam for People*'appearing in the public domain might be *five years*? (Of course, governments, security services and the military may well be even more advanced in this at this very moment in time – they're always ahead of the rest.) But if you're not convinced it's this close – that this seems a little fanciful, try playing the '*man (or woman) on a train*' game…

Imagine you're sitting across from someone on a train. They're reading or working on something. You have a mobile and an Internet connection. How close can you get to finding out who they are and everything about them? Well, as we know, the 'finding stuff on the Internet and reporting back' stage (3) is the easy bit. What's needed in the first instance is that little bit of critical information that links the person sitting in front of you to the Internet data (1 & 2). What would make it dead easy, of course, would be a sight of a business card or name badge with an email or web address, etc. but we can probably achieve much the same in other ways. Just a few examples:

- *They could be an academic reading a paper on a certain subject; you might know where they got on the train or where they're getting off. A quick look at the photos on the 'Department of X' webpage for the 'University of Y' and you're off …*
- *You might be listening to a conversation between two people; they might be discussing other people or places or events or interests or problems or … Remember how easily Google can link all of that stuff together?*
- *They might be identifiable as working for (clothing) or belonging to (badge) a particular institution. With one other key piece of information (any of the above), that might be enough to identify them …*
- *"I've just posted ABC on XYZ", you hear. That should be easy enough to find …*
- *If you're still missing that key piece of information, you could try* talking *to them and steering the conversation in that direction …*

None of these techniques have guaranteed accuracy, of course, but they might get you close and there's no doubt they'd work some of the time. As time goes by, 'some' of the time is likely to become 'most' of the time and so on… And most of this, of course, is entirely legal. Pretty sneaky of course – highly questionable, ethically – but generally within the law. What wouldn't necessarily be legal is systematically collating this information and making it available as a service – free or for a fee. However, don't jump to the comfortable conclusion that this would automatically be *against* the law. Some of it would be but much wouldn't. (If it can already be argued to be in the public domain it may well be OK.) Remember, it's likely to be the linking of general material to the personal fingerprint that tests the lawyers, not the operation and results of an open web-search. The business brains usually get in ahead of the legal brains and that's all it generally takes to do the damage. Let's see what happens, shall we?

An easily overlooked obstacle to some of the above suggestions, though, is that they rely on *human* intelligence to find ways of matching the individual to their data – and *we* can still be pretty sneaky in ways that machines possibly can't. The equivalent 'semantic' intelligence of the Web still hasn't quite delivered on this yet but – once more – it's improving all the time and don't forget that this is the essential purpose of the 'personal fingerprint' and the database to match it against. Tough, yes, but coming without much doubt. The '*man on the train*' game is really to help us see where the gaps are *currently* in the proposed system; when we look carefully, there may not actually be that many.

Accuracy is another big issue here. Both the matching of the individual and the information stored about them are subject to error. If the resolution of the sample isn't sufficient, we won't be able to find the right person; if the information's been pulled together from open contributions, it will be hugely inaccurate – at least to begin with. Will this prove to be a deal-breaker for '*Shazam for People*'? Unlikely really: the Internet has always been like that. Google, for example, doesn't always understand what you're looking for and the information it returns isn't always reliable. Wikipedia is infamously poor but universally popular. '*Shazam for People*'will probably lack both focus and accuracy in its early stages but it will improve on both counts and, ultimately, people will use (and possibly pay for) the service that works best. Like all developmental early releases, it will be rubbish at first and get gradually better.

Of course, there are actually *much bigger questions* than 'will it happen?' – That's probably a given really, considering all the above. But what of morality and ethics? Technology's never been particularly good at dealing with those and it probably never will be. Moral objections (and, yes indeed, it *does* seem right to have *massive* objections to such a system) rarely seem to actually *impede* technological advance – for good or bad. We're usually left to *dea lwith* the fallout from technology rather than having much say in whether we *want* it. But of course, we *do* want it; at least we want it about *other people*; we're just not comfortable when the tables are turned on *us*. ('*Smart knows what the neighbours paid*'.) As long as there are willing customers, there will be service providers.

Ultimately, we're probably just going to have to get used to having fewer secrets. If we each have a 'bubble' around us with our personal stuff inside of it and the public bits outside, then what's going to happen? It's possible that the bubble won't actually burst but it's *very* likely to shrink… *a lot*. We may, in the future, have to look to protecting our closest secrets with redoubled effort and watch the less important things float away into the public cloud. But even then it might be that the dreaded spectre of market forces comes into play here; it could even split the payers and non-payers (the 'haves' and the

'have nots') in two. People with access to the system might actually know more about a person than the person themselves. *"He thinks he has two kids but one of them isn't his."* We know it's going to be hard to stop this so we may just have to adapt to it.

One possible piece of social advice for the next decade might well be: *If you've got any skeletons in the cupboard and the door's loose, it might be worth getting them voluntarily out into the light of day now?*

Austin J Robot & Autom - Volume 1 Issue 1 - 2014
ISSN : 2471-0407 | www.austinpublishinggroup.com

Citation: Grout V. Shazam for People?. Austin J Robot & Autom. 2014;1(1): 4.