



Forensic speaker recognition: An evidence odyssey – Summary

Didier Meuwly – R&D, Informatics, The Forensic Science Service

Trident Court, 2920 Solihull Parkway, Birmingham Business Park

BIRMINGHAM B37 7YN, UK

Introduction

This summary describes firstly a general procedure for the evaluation of biometric technology and secondly the two forensic scenarios in which biometric technology can be involved. Then, an example of biometric evaluation in a simulated application is presented: 2 automatic speaker recognition systems are tested and compared in a forensic scenario.

Biometric Evaluation

Technology evaluation

The goal of a technology evaluation is to compare competing algorithms from a single technology. Testing of all algorithms is done on a standardized database collected by a “universal” sensor. Nonetheless, performance against this database will depend upon both the environment and the population in which it was collected. Because the database is fixed, results of technology tests are repeatable [1].

Scenarios evaluation

The goal of scenario testing is to determine the overall system performance in a prototype or simulated application. Testing is done on a complete system in an environment that models a “real-world” target application of interest. Each tested system will have its own acquisition sensor and so will receive slightly different data. Consequently, care will be required that data collection across all tested systems is in the same environment with the same population. Test results will be repeatable only to the extent that the modelled scenario can be carefully controlled [1].

Operational evaluation

The goal of operational testing is to determine the performance of a complete biometric system in a specific application environment with a specific target population. In general, operational test results will not be repeatable because of unknown and undocumented differences between operational environments [1].

Forensic scenarios

Biometric technology can be applied to infer the identity of the source of forensic trace material in two forensic scenarios defined in [2]: the forensic investigative scenario and the forensic evaluative scenario. These two scenarios are explained using automatic speaker recognition as an example.

Forensic investigative scenario: production of short lists

Automatic speaker recognition can be used to establish a short list of the most relevant sources of a questioned recording among a set of known potential speakers. The expression “known potential

speakers" enhances the fact that the hypotheses tested during a forensic investigation should be plausible but cannot be exhaustive. In this scenario the biometric technology is used to compute the likelihood of the questioned recording, given each known potential speaker.

Forensic evaluative scenario: production of likelihood ratios

Automatic speaker recognition can also be used to evaluate a questioned recording presented as a piece of evidence in front of a court of justice, when the prosecution and the defence support mutually exclusive hypotheses for the source of the questioned recording. If the prosecution supports the hypothesis that the defendant is the source of the questioned recording and the defence supports the hypothesis that another speaker is the source of the questioned recording, the biometric technology can be used to compute the likelihood ratio of the questioned recording, given the two competing hypotheses.

Comparison of 2 automatic speaker recognition systems tested in the forensic evaluative scenario

The Automatic SPeaker Individualisation Systems 1 and 2 (ASPIC 1 and 2) developed by the Swiss Institute of Technology (ITS-EPFL) and the University of Lausanne (IPS-UNIL) have been tested and compared in the forensic evaluative scenario. Likelihood ratios have been produced using a database of 3 sets chosen to model this scenario: 1) a set of recordings from speakers playing the role of the defendants, 2) a set of recordings from speakers modelling the potential population and 3) a set composed of questioned recordings simulating several forensic conditions: dialogues without or with noise, anonymous calls without or with disguise transmitted through the PSTN or the GSM network and recorded with a digital or an analogue recording system.

Conclusion

The definition of forensic scenarios as specific applications in which biometric technology can be involved enables the capability to set up evaluation frameworks to test and compare technologies such as automatic speaker recognition in forensic scenarios.

Bibliography

- [1] A. J. Mansfield and J. L. Wayman, "Report Best Practices in Testing and Reporting Performance of Biometric Devices For Biometrics Working Group," Centre for Mathematics and Scientific Computing, National Physical Laboratory, Teddington, Middlesex, UK Issue 2 draft 9, 08/02 2002.
- [2] G. Jackson, C. Champod, I. W. Evett, and S. Jones, "The forensic scientist - investigator or evaluator," personal communication, paper in preparation.