

Emulating DNA: Rigorous Quantification of Evidential Weight in Transparent and Testable Forensic Speaker Recognition

Joaquin Gonzalez-Rodriguez, *Member, IEEE*, Phil Rose, Daniel Ramos, *Student Member, IEEE*, Doroteo T. Toledano, *Member, IEEE*, and Javier Ortega-Garcia, *Member, IEEE*

Abstract—Forensic DNA profiling is acknowledged as the model for a scientifically defensible approach in forensic identification science, as it meets the most stringent court admissibility requirements demanding transparency in scientific evaluation of evidence and testability of systems and protocols. In this paper, we propose a unified approach to forensic speaker recognition (FSR) oriented to fulfil these admissibility requirements within a framework which is transparent, testable, and understandable, both for scientists and fact-finders. We show how the evaluation of DNA evidence, which is based on a probabilistic similarity-typicality metric in the form of likelihood ratios (LR), can also be generalized to continuous LR estimation, thus providing a common framework for phonetic-linguistic methods and automatic systems. We highlight the importance of calibration, and we exemplify with LRs from diphthongal F-pattern, and LRs in NIST-SRE06 tasks. The application of the proposed approach in daily casework remains a sensitive issue, and special caution is enjoined. Our objective is to show how traditional and automatic FSR methodologies can be transparent and testable, but simultaneously remain conscious of the present limitations. We conclude with a discussion on the combined use of traditional and automatic approaches and current challenges for the admissibility of speech evidence.

Index Terms—Admissibility of speech evidence, calibration, Daubert, deoxyribonucleic acid (DNA), forensic speaker recognition (FSR), likelihood ratio (LR).

expressed in the form of individualisation (hard *match*), as categorical opinion of identity of sources, exclusion (*nonmatch*) statements, or making use of verbal scales of probability of hypothesis, given evidence. The process leading from evidence to conclusion is often opaque, either because it lacks scientific rigor and is inherently unfalsifiable, or because the approach is inadequately tested, and thus cannot quote random match probabilities or estimate the chance of error. Not surprisingly, this has often resulted in legal discussion about the acceptance of expert testimony. Contrasting with this, DNA profiling [1], [5], [65] has solid and well-known scientific foundations, and is probabilistic [18], [60]. Avoiding individualization or exclusion statements for the determination of the source of the evidence, DNA evidence is often presented using frequencies, match probabilities, and inclusion or exclusion probabilities [24], but many influential forensic scientists [1], [16], [27], [26] advocate assessing the weight of the evidence with *likelihood ratios* (LRs) [24], [29]. This LR is a quotient of a *similarity* factor, which supports the hypothesis that a questioned sample was left by a given suspect, and a *typicality* factor, which quantifies support for the hypothesis that the questioned sample was left by someone else in a relevant population. In DNA typing, this likelihood-ratio approach for evidence analysis [1], [28] has been held up as a model of an explicit and probabilistic framework which allows quantification and inspection by fact finders