A new approach to computing likelihood ratios based on principal component analysis

Balamurali B.T.Nair[1], *Esam A.S. Alzqhoul*[1], *Bernard J. Guillemin*[1*]

[1]Department of Electrical & Computer Engineering, The University of Auckland, New Zealand

*bj.guillemin@auckland.ac.nz

A number of probabilistic procedures have been proposed for computing the likelihood ratio (LR) in a forensic voice comparison (FVC), such as Univariate Kernel Density (UKD), Multivariate Kernel Density (MVKD), and Gaussian Mixture Model – Universal Background Model (GMM-UBM) analyses [1]. Among these, MVKD has grown in popularity, a feature being its ability to handle correlations between multiple input parameters from the same speech segment. An important point to note in respect to MVKD, and one that is sometimes overlooked by researchers, is that it was designed to handle small numbers of parameters in the region of 3 or 4 only [2]. Some researchers have started using it for larger numbers of parameters in the region of 14 [3]. Though they have presented promising results, in our experience MVKD will sometimes produce nonsense results in such situations. The underlying cause of this fragility is linked to the inversion of illconditioned matrices, this becoming problematic as matrix size increases.

In this paper we present a new approach to computing LRs based on principal component analysis (PCA). With this approach, termed Principal Component Analysis Kernel Likelihood Ratio (PCAKLR), multiple input parameters are first transformed into sets of uncorrelated parameters. A LR is then computed individually for each of these sets using UKD and then an overall LR determined by combining these individual LRs through simple multiplication according to the naive Baye's approach. From a functional standpoint, PCAKLR is equivalent to MVKD in that suspect data is modelled using a normal distribution, background data using a kernel density distribution, and correlations are taken into account in the comparison process. The key difference between them, however, is that PCAKLR is mathematically robust irrespective of the number of input parameters.

In order to compare the performance of MVKD with PCAKLR, we have used a speech database comprising 297 male speakers compiled by the Japanese National Research Institute of Police Science (NRIPS) for forensic speaker recognition tests. Tokens of the five Japanese vowels /i e a o u/ have been extracted from this database and 14 linear prediction cepstral coefficients (LPCC) computed for each token. The database has been divided into three groups of speakers, 99 for the Development group (used for calibrating and fusing results for individual vowels), 99 for the Testing group and 99 for the Backround group. An interesting feature of PCAKLR is the manner in which results for multiple segments (e.g., vowels) can be combined. In respect to MVKD, this is traditionally done using logistic regression fusion. Though the same approach can be used for PCAKLR, an alternative strategy is to bring together results for individual vowels into one large data matrix and leave it to the PCA stage of PCAKLR to sort this data out, removing all correlations. Fig. 1 shows a Tippett plot comparing the performance of MVKD and PCAKLR for 4 input parameters. In the case of MVKD, fusion has been used to combine results for all 5 vowels, but for PCAKLR the above strategy has been implemented (termed $PCAKLR_{NF}$ for No Fusion). Using log-likelihood-ratio cost (C_{llr}) for assessing the accuracy of a FVC, for MVKD $C_{llr} = 0.3739$ and for $PCAKLR_{NF}$ $C_{llr} = 0.3592$. As expected, these C_{llr} values, together with the Tippett plot results, confirm that MVKD and PCAKLR have performed similarly for small numbers of parameters. But the real advantage of PCAKLR is that it can reliably handle large numbers of parameters, something which, in our experience, MVKD cannot do.

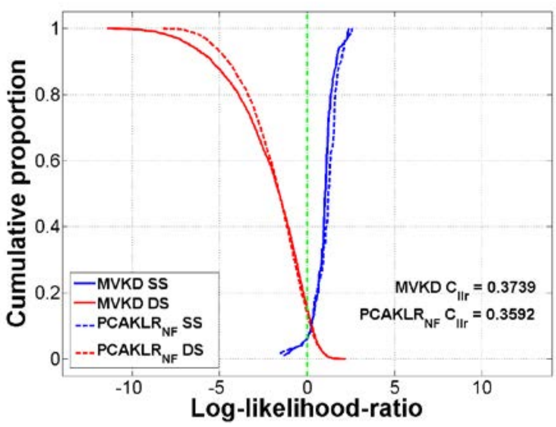

Fig. 1 Tippett plot of combined results for five vowels. MVKD – fusion; PCAKLR – no fusion

References

[1] Morrison, G.S., A comparison of procedures for the calculation of forensic likelihood ratios from acoustic-phonetic data: Multvariate kernel density (MVKD) versus Gaussian mixture model-universal background model (GMM-UBM). Speech Communication, 2010.

[2] Aitken, C.G.G. and D. Lucy, *Evaluation of trace evidence in the form of multivariate data.* Applied Statistics, 2004: p. 109-122.

[3] Rose, P., Forensic Voice Comparison With Japanese Vowel Acoustics - A Likelihood Ratio-Based Approach Using Segmental Cepstra. The 17th International Congress of Phonetic Sciences (ICPHS XVII), 2011.