
Bibliography

- Aitken, C. (1997). *Statistics and the Evaluation of Evidence for Forensic Scientists*. John Wiley & Sons, Chichester.
- Aitken, C. and Taroni, F. (2004). *Statistics and the Evaluation of Evidence for Forensic Scientists*. John Wiley & Sons, Chichester.
- Alexander, A., Botti, F., and Drygajlo, A. (2004). Handling Mismatch in Corpus-Based Forensic Speaker Recognition. In *Odyssey 2004, The Speaker and Language Recognition Workshop*, pages 69–74, Toledo, Spain.
- Alexander, A., Drygajlo, A., and Botti, F. (2005). NFI: Speaker recognition evaluation through a fake case. Case report, Swiss Federal Institute of Technology, Lausanne (EPFL), Lausanne, Switzerland.
- Arcienega, M., Alexander, A., Zimmerman, P., and Drygajlo, A. (2005). A bayesian network approach combining pitch and spectral envelope features to reduce channel mismatch in speaker verification and forensic speaker recognition. In *Inter-speech'2005 - Eurospeech - 9th European Conference on Speech Communication and Technology*, pages 2009–2012, Lisbon, Portugal.
- Arcienega, M. and Drygajlo, A. (2003). A Bayesian network approach for combining pitch and reliable spectral envelope features for robust speaker verification. In Kittler, J. and Nixon, M. S., editors, *Proc. 4th Int. Conf. on Audio- and Video-Based Biometric Person Authentication*, pages 78–85, Guildford, UK. Springer.
- Auckenthaler, R., Carey, M. J., and Lloyd-Thomas, H. (2000). Score normalisation for text-independent speaker verification system. *Digital Signal Processing (DSP), a review journal - Special issue on NIST 1999 Speaker Recognition Workshop*, 10 (1-3):42–54.
- Beck, S. D. (1998). *FBI Voice Database For Automated Speaker Recognition Systems, User's Manual for Training and Testing*. Federal Bureau of Investigation.

- Beck, S. D., Schwartz, R., and Nakasone, H. (2004). A bilingual multi-modal voice corpus for language and speaker recognition (LASR) services. In *Odyssey 2004, The Speaker and Language Recognition Workshop*, pages 265–270, Toledo, Spain.
- Ben, M., Blouet, R., and Bimbot, F. (2002). A Monte-Carlo method for score normalization in automatic speaker verification using Kullback-Leibler distances. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Orlando, Florida.
- Bimbot, F., Bonastre, J.-F., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., Merlin, T., Ortega-Garcia, J., Petrovska-Delacrétaz, D., and Reynolds, D. (2004). A tutorial on text-independent speaker verification. *EURASIP Journal on Applied Signal Processing*, 4:430–451.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press, Walten street, Oxford.
- Bolle, R. M., Ratha, N. K., and Pankanti, S. (2004). Error analysis of pattern recognition systems: the subsets bootstrap. *Comput. Vis. Image Underst.*, 93(1):1–33.
- Bolt, R. H., Cooper, F. S., David, E. E., Denes, P. B., Pickett, J. M., and Stevens, K. N. (1970). Speaker identification by speech spectrograms: A scientist’s view of its reliability for legal purposes. *J. Acoustic. Soc. Am.*, 47(2):597–603.
- Bolt, R. H., Cooper, F. S., David, E. E., Denes, P. B., Pickett, J. M., and Stevens, K. N. (1973). Speaker identification by speech spectrogram: some further observations. *J. Acoust. Soc. Amer.*, 54(2):531–537.
- Bonastre, J.-F., Bimbot, F., Boë, L.-J., Campbell, J. P., Reynolds, D. A., and Magrin-Chagnolleau, I. (2003). Auditory Instrumental Forensic Speaker Recognition. In *Proc. Eurospeech 2003*, pages 33–36, Geneva, Switzerland.
- Botti, F., Alexander, A., and Drygajlo, A. (2004a). An interpretation framework for the evaluation of evidence in forensic automatic speaker recognition with limited suspect data. In *Proceedings of 2004: A Speaker Odyssey*, pages 63–68, Toledo, Spain.
- Botti, F., Alexander, A., and Drygajlo, A. (2004b). On compensation of mismatched recording conditions in the bayesian approach for forensic automatic speaker recognition. *Forensic Science International*, 146(Supplement 1,2):S101–S106.

- Bouten, J. S. and van Leeuwen, D. A. (2004). Results of the 2003 NFI-TNO forensic speaker recognition evaluation. In *Odyssey 2004, The Speaker and Language Recognition Workshop*, pages 75–82, Toledo, Spain.
- Bregman, A. (1990). *Auditory Scene Analysis*. MIT Press, Cambridge Mass.
- Byrne, C. and Foulkes, P. (2004). The mobile phone effect on vowel formants. *The International Journal of Speech, Language and the Law*, 11:83–102.
- Cambier-Langeveld, T. (2005). Netherlands Forensic Institute (NFI), speaker recognition fake case evaluation. June 2-3, 2005, 8th Meeting of the ENFSI Expert Working Group for Forensic Speech and Audio Analysis (FSAAWG).
- Campbell, J. (1997). Speaker recognition: A tutorial. *Proceedings of IEEE*, 85:1437–1462.
- Campbell, W., Reynolds, D., Campbell, J., and Brady, K. (2005). Estimating and evaluating confidence for forensic speaker recognition. In *ICASSP*, volume 1, pages 717–720, Philadelphia, USA.
- Champod, C. and Evett, I. W. (2000). Commentary on: Broeders, A. P. A. (1999) 'Some observations on the use of probability scales in forensic identification', *Forensic Linguistics*, 6(2): 228-41. *Forensic Linguistics*, 7.2:238–243.
- Champod, C. and Meuwly, D. (2000). The inference of identity in forensic speaker recognition. *Speech Communication*, 31:193–203.
- Cook, R., Evett, I., Jackson, G., Jones, P. J., and Lambert, J. (1998). A model for case assessment and interpretation. *Science & Justice*, 38(3):151–156.
- Davis, S. and Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4):357–366.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society (B)*, 39:1–38.
- Dessimoz, D. (2004). Reconnaissance de locuteurs : Comparaison de performances entre la reconnaissance auditive par des profanes et de systemes automatiques. Project report, Institut de Police Scientifique, Ecole des Sciences Criminelles, University of Lausanne, Switzerland.
- Doddington, G. (1985). Speaker recognition - identifying people by their voices. *Proceedings of the IEEE*, 73(11):1651–1664.

- Doddington, G., Liggett, W., Martin, A., Przybocki, M., and Reynolds, D. A. (1998). Sheep, goats, lambs and wolves: a statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation. In *5th International Conference on Spoken Language Processing (ICSLP)*, page 0608, Sydney, Australia.
- Drygajlo, A., Meuwly, D., and Alexander, A. (2003). Statistical Methods and Bayesian Interpretation of Evidence in Forensic Automatic Speaker Recognition. In *Proc. Eurospeech 2003*, pages 689–692, Geneva, Switzerland.
- Dunn, R. B., Quatieri, T. F., Reynolds, D. A., and Campbell, J. (2001). Speaker recognition from coded speech in matched and mismatched conditions. In *2001: A Speaker Odyssey*, Crete, Greece.
- El-Maliki, M. (2000). *Speaker verification with missing features in noisy environments*. PhD thesis, Swiss Federal Institute of Technology, Lausanne, Switzerland.
- Evetts, I. (1998). Towards a uniform framework for reporting opinions in forensic science casework. *Science and Justice*, 38(3):198–202.
- Evetts, I. W. and Buckleton, J. S. (1996). Statistical analysis of STR data : Advances in forensic haemogenetics. *Springer-Verlag*, 6:79–86.
- Farrell, K. R., Mammone, R. J., and Assaleh, K. T. (1994). Speaker recognition using neural networks and conventional classifiers. *IEEE Transactions on speech and audio processing*, 2(1):194–205.
- Fraser, H. (2003). Issues in transcription: factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics*, 10(2):203–226.
- Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Academic Press, second edition.
- Furui, S. (1997). Recent advances in speaker recognition. *Pattern Recognition Letters*, 18(9):859–872.
- Gauvain, J. L. and Lee, C. H. (1994). Maximum a-posteriori estimation of multivariate Gaussian mixture observations. *IEEE Trans. Speech, Audio Processing*, 2:291–298.
- Gonzales-Rodriguez, J., Drygajlo, A., Ramos-Castro, D., Garcia-Gomar, M., and Ortega-Garcia, J. (2005). Robust estimation, interpretation and assessment of likelihood ratios in forensic speaker recognition. *Computer Speech and Language (to appear)*.

- Gonzalez-Rodriguez, J., Ramos-Castro, D., Garcia-Gomar, M., and Ortega-Garcia, J. (2004). On robust estimation of likelihood ratios: the ATVS-UPM system at 2003 NFI/TNO forensic evaluation. In *Odyssey 2004, The Speaker and Language Recognition Workshop*, Toledo, Spain.
- Gray, R. M. (1984). Vector quantization. *IEEE ASSP Magazine*, 1:4–29.
- Gruber, J. S., Poza, F., and Pellicano, A. J. (1993). *Audio Recordings: Evidence, Experts and Technology*, volume 48 of *Am Jur Trials*. Lawyers Cooperative Publishing.
- Hennebert, J., Melin, H., Petrovska, D., and Genoud, D. (2000). POLYCOST: A telephone-speech database for speaker recognition. *Speech Communication*, 31:265–270.
- Hermansky, H. (1990). Perceptual linear predictive (PLP) analysis of speech. *J. Acoust. Soc. Amer.*, 87(4):1738–1752.
- Hermansky, H. (1994). RASTA Processing of Speech. *IEEE Trans. on Speech, and Audio Proc*, 2(4):578–589.
- Hollien, H. (1990). *The Acoustics of Crime, The New Science of Forensic Phonetics*. Number ISBN 0-306-43467-9. New York: Plenum Publishing Corporation.
- Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. Springer.
- Kerstholt, J., Jansen, E., Amelsvoort, A. v., and Broeders, A. (2003). Earwitness line-ups: effects of speech duration, retention interval and acoustic environment on identification accuracy. In *Proc. Eurospeech 2003*, pages 709–712, Geneva, Switzerland.
- Künzel, H. (1997). Some general phonetic and forensic aspects of speaking tempo. *Forensic Linguistics*, 4/1:48–83.
- Künzel, H. J. (1998). Forensic speaker identification: A view from the crime lab. In *Proceedings of the COST Workshop on Speaker Recognition by Man and Machine*, pages 4–8, Technical University of Ankara, Ankara, Turkey.
- Koenig, B. (1990). Authentication of forensic audio recordings. *Journal of the Audio Engineering Society*, 38(1/2):3–33.
- Koolwaaij, J. and Boves, L. (1999). On decision making in forensic casework. *Forensic Linguistics, the International Journal of Speech, Language and the Law*, 6(2):242–264.

- Kreyszig, E. (1999). *Advanced engineering mathematics*. John Wiley and Sons, Singapore, 8th edition.
- Li, K. P. and Porter, J. (1988). Normalizations and selection of speech segments for speaker recognition scoring. In *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-88.*, volume 1, pages 595–598, New York, USA.
- Loevinger, L. (1995). Science as evidence. *Jurimetrics*, 35(2):153–190.
- Majewski, W. and Basztura, C. (1996). Integrated approach to speaker recognition in forensic applications. *Forensic Linguistics*, 3(1):50–64.
- Makhoul, J. (1975). Linear prediction: A tutorial review. *Proc. IEEE*, 63:561–580.
- Martin, A., Doddington, G., Kamm, T., Ordowski, M., and Przybocki, M. (1997). The DET Curve in Assessment of Detection Task Performance. In *Proc. Eurospeech '97*, pages 1895–1898, Rhodes, Greece.
- Martin, R. (1994). Spectral subtraction based on minimum statistics. In *EUSIPCO-94*, pages 1182–1185.
- Meuwly, D. (2000). *Voice Analysis, in : Encyclopedia of Forensic Science*, pages 1413 – 1420. London: Academic Press Ltd.
- Meuwly, D. (2001). *Reconnaissance de locuteurs en sciences forensiques: l'apport d'une approche automatique*. PhD thesis, IPSC, University of Lausanne.
- Meuwly, D., Alexander, A., Drygajlo, A., and Botti, F. (2003). Polyphone-IPSC: A shared speakers database for evaluation of forensic-automatic speaker recognition systems. In *Forensic Science International*, volume 136, page 367, Istanbul, Turkey. Elsevier.
- Meuwly, D. and Drygajlo, A. (2001). Forensic Speaker Recognition Based on a Bayesian Framework and Gaussian Mixture Modeling (GMM). In *2001, A Speaker Odyssey: The Speaker Recognition Workshop*, pages 145–150.
- Meuwly, D., El-Maliki, M., and Drygajlo, A. (1998). Forensic speaker recognition using Gaussian mixture models and a bayesian framework. In *8th COST 250 workshop: Speaker Identification by Man and by Machine: Directions for Forensic Applications*, pages 52–55.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits in our Capacity for Processing Information. *The Psychological Review*, 63:81–97.

- Nakasone, H. and Beck, S. D. (2001). Forensic Automatic Speaker Recognition. In *2001: A Speaker Odyssey*, pages 139–144, Crete, Greece.
- Navratil, J. and Ramaswamy, G. N. (2003). The awe and mystery of T-norm. In *Proc. Eurospeech 2003*, pages 2009–2012, Geneva, Switzerland.
- Niemi-Laitinen, T., Saastamoinen, J., Kinnunen, T., and Fränti, P. (2005). Applying MFCC-based automatic speaker recognition to GSM and forensic data. In *2nd Baltic Conf. on Human Language Technologies (HLT'05)*, pages 317–322, Tallinn, Estonia.
- Oglesby, J. Mason, J. (1989). Speaker recognition with a neural classifier. In *Proceedings First IEE International Conference on artificial Neural Networks*, volume 313, pages 306–309.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Fransisco: Morgan Kaufmann Publishers.
- Pfister, B. and Beutler, R. (2003). Estimating the weight of evidence in forensic speaker verification. In *Proc. Eurospeech 2003*, pages 701–704, Geneva, Switzerland.
- Reynolds, D. A. (1992). *A Gaussian mixture modeling approach to text-independent speaker identification*. PhD thesis, Georgia Institute of Technology, Atlanta, Georgia.
- Reynolds, D. A. (1994). Experimental evaluation of features for robust speaker identification. *IEEE Transactions on Speech and Audio Processing*, 2(4):639–643.
- Reynolds, D. A. (1995). Speaker identification and verification using Gaussian mixture speaker models. *Speech Commun.*, 17(1-2):91–108.
- Reynolds, D. A. (2003). Channel robust speaker verification via feature mapping. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, volume 2, pages 53–56.
- Reynolds, D. A., Doddington, G. R., Przybocki, M. A., and Martin, A. F. (2000). The NIST speaker recognition evaluation - overview methodology, systems, results, perspective. *Speech Communication*, 31(2-3):225–254.
- Reynolds, D. A. and Rose, R. C. (1995). Robust text-independent speaker identification using Gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3(1):72–83.

- Richiardi, J., Prodanov, P., and Drygajlo, A. (2005). A probabilistic measure of modality reliability in speaker verification. In *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing 2005*, pages 709–712, Philadelphia, USA.
- Robertson, B. and Vignaux, G. A. (1995). *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*. John Wiley and Sons, Chichester.
- Rose, P. (2002). *Forensic Speaker Identification*. Forensic Science. London & New York: Taylor and Francis.
- Rose, P. (2003). *The technical comparison of forensic voice samples*. Freckelton I. and Selby H, series editors, Issue 99, Expert evidence. Sydney: Thomson Lawbook Company.
- Rose, P. (2005). Technical forensic speaker recognition: evaluation, types and testing of evidence. *Computer Speech and Language (to appear)*.
- Rosenberg, A. (1976). Automatic speaker verification: A review. *Proceedings of the IEEE*, 64(4):475–487.
- Rossy, Q. (2003). Simulation de cas reels de reconnaissance de locuteurs au moyen du logiciel ASPIC. Project report, Institut de Police Scientifique, Ecole des Sciences Criminelles, University of Lausanne, Switzerland.
- Schmidt-Nielsen, A. and Crystal, T. H. (2000). Speaker verification by human listeners: Experiments comparing human and machine performance using the NIST 1998 speaker evaluation data. *Digital Signal Processing*, 10:249–266.
- Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- Sohn, J., Kim, N., and Sung, W. (1999). A statistical model-based voice activity detector. *IEEE Signal Process. Lett.*, 6(1):1–3.
- Soong, F., Rosenberg, A., Rabiner, L., and Juang, B. (1985). A vector quantization approach to speaker recognition. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 10, pages 387–390.
- Srinivasan, K. and Gersho, A. (1993). Voice activity detection for cellular networks. In *Proc. IEEE Speech Coding Workshop*, number 85-86.
- Tippett, C. F., Emerson, V., Fereday, M., Lawton, F., and Lampert, S. (1968). The evidential value of the comparison of paint flakes from sources other than vehicles. *Journal of Forensic Sci. Soc.*, 8:61–65.

-
- Voiers, W. D. (1964). Perceptual bases of speaker identity. *Journal of the Acoustic Society of America*, 36(6):1065–1073.
- Wolf, J. (1972). Efficient acoustic parameters for speaker recognition. *Journal of the Acoustical Society of America*, 51:2044–2056.
- Yarmey, A. D. (1995). Earwitness Speaker Identification. *Psychology, Public Policy, and Law*, 1(4):792–816.
- Zimmermann, P. (2005). Analyse de l'influence des conditions d'enregistrement dans la reconnaissance automatique de locuteurs en sciences forensiques. Project report, Institut de Police Scientifique, Ecole des Sciences Criminelles, University of Lausanne, Switzerland.