

A Semi-supervised Speaker Identification Method for Audio Forensics Using Cochleagrams

Workshop on Engineering Applications

WEA 2017: Applied Computer Sciences in Engineering pp 55-64 | Cite as

- Steven Camacho (1) Email author (u3900217@unimilitar.edu.co)
- Diego Renza (1)
- Dora M. Ballesteros L. (1)

1. Universidad Militar Nueva Granada, Bogotá, Colombia

Conference paper

First Online: 29 August 2017

- [1 Readers](#)
- [290 Downloads](#)

Part of the [Communications in Computer and Information Science](#) book series (CCIS, volume 742)

Abstract

The general task in speaker identification for audio forensics is to identify the unknown speaker within an audio proof, who is suspected of a crime. Here, the voice of each person within a group of suspects is compared to the audio proof with the aim to determining which of them corresponds to the source. In this paper, a semi-supervised speaker identification method is proposed, which does not require a training stage. Also, the feature extraction is based on the use of cochleagrams for the previously selected words. The system can identify one or multiple suspects which have high similarity to the audio proof, or give a null response if none of the suspects satisfies a similarity threshold. The results of the proposed method are compared with the respective results of the same method but using spectrograms instead of cochleagrams. The performance of our system is measured through a confusion matrix (true and false positives, and true and false negatives) and global results are given in terms of overall accuracy and kappa index. According to several tests, our system has an overall accuracy higher than 0.97 and a kappa index around 0.78; this means a high confidence in the results of identification and rejection.

Keywords

Speaker identification Audio forensics Cochleagram Correct identification
Correct rejection

This is a preview of subscription content, [log in](#) to check access

Notes

Acknowledgment

This work is supported by the “Universidad Militar Nueva Granada-Vicerrectoría de Investigaciones” under the grant IMP-ING-2136 of 2016.

References

1. Ajmera, P.K., Jadhav, D.V., Holambe, R.S.: Text-independent speaker identification using radon and discrete cosine transforms based features from speech spectrogram. *Pattern Recogn.* **44**(10), 2749–2759 (2011)
[CrossRef](#) (<https://doi.org/10.1016/j.patcog.2011.04.009>)
[Google Scholar](#) (http://scholar.google.com/scholar_lookup?title=Text-independent%20speaker%20identification%20using%20radon%20and%20discrete%20cosine%20transforms%20based%20features%20from%20speech%20spectrogram&author=PK.%20Ajmera&author=DV.%20Jadhav&author=RS.%20Holambe&journal=Pattern%20Recogn.&volume=44&issue=10&pages=2749-2759&publication_year=2011)
2. Alegre, F.L.: Application of ANN and HMM to automatic speaker verification. *IEEE Lat. Am. Trans.* **5**(5), 329–337 (2007)
[CrossRef](#) (<https://doi.org/10.1109/TLA.2007.4378525>)
[Google Scholar](#) (http://scholar.google.com/scholar_lookup?title=Application%20of%20ANN%20and%20HMM%20to%20automatic%20speaker%20verification&author=FL.%20Alegre&journal=IEEE%20Lat.%20Am.%20Trans.&volume=5&issue=5&pages=329-337&publication_year=2007)
3. Almaadeed, N., Aggoun, A., Amira, A.: Speaker identification using multimodal neural networks and wavelet analysis. *IET Biometrics* **4**(1), 18–28 (2015)
[CrossRef](#) (<https://doi.org/10.1049/iet-bmt.2014.0011>)
[Google Scholar](#) (http://scholar.google.com/scholar_lookup?title=Speaker%20identification%20using%20multimodal%20neural%20networks%20and%20wavelet%20analysis&author=N.%20Almaadeed&author=A.%20Aggoun&author=A.%20Amira&journal=IET%20Biometrics&volume=4&issue=1&pages=18-28&publication_year=2015)
4. Ballesteros, L., Renza, D., Camacho, S.: An unconditionally secure speech scrambling scheme based on an imitation process to a Gaussian noise signal. *J. Inf. Hiding Multimedia Sig. Process* **7**(2), 233–242 (2016)
[Google Scholar](#) (<https://scholar.google.com/scholar?q=Ballesteros%2C%20L.%2C%20Renza%2C%20D.%2C%20Camacho%2C%20S.%20An%20unconditionally%20secure%20speech%20scrambling%20scheme%20based%20on%20an%20imitation%20process%20to%20a%20Gaussian%20n>

oise%20signal.%20J.%20Inf.%20Hiding%20Multimedia%20Sig.%20Process%207%282%29%2C%20233%E2%80%93242%20%282016%29)

5. Ballesteros, L.D.M., Moreno, A.J.M.: A bit more on the ability of adaptation of speech signals. *Revista Facultad de Ingeniería Universidad de Antioquia* **66**, 82–90 (2013)
[Google Scholar](https://scholar.google.com/scholar?q=Ballesteros%2C%20L.D.M.%2C%20Moreno%2C%20A.J.M.%3A%20A%20bit%20more%20on%20the%20ability%20of%20adaptation%20of%20speech%20signals.%20Revista%20Facultad%20de%20Ingenier%C3%ADa%20Universidad%20de%20Antioquia%2066%2C%2082%E2%80%9390%20%282013%29) (<https://scholar.google.com/scholar?q=Ballesteros%2C%20L.D.M.%2C%20Moreno%2C%20A.J.M.%3A%20A%20bit%20more%20on%20the%20ability%20of%20adaptation%20of%20speech%20signals.%20Revista%20Facultad%20de%20Ingenier%C3%ADa%20Universidad%20de%20Antioquia%2066%2C%2082%E2%80%9390%20%282013%29>)
6. Campbell, J.P., Shen, W., Campbell, W.M., Schwartz, R., Bonastre, J.F., Matrouf, D.: Forensic speaker recognition. *IEEE Signal Process. Mag.* **26**(2), 95–103 (2009)
[Google Scholar](https://scholar.google.com/scholar?q=Campbell%2C%20J.P.%2C%20Shen%2C%20W.%2C%20Campbell%2C%20W.M.%2C%20Schwartz%2C%20R.%2C%20Bonastre%2C%20J.F.%2C%20Matrouf%2C%20D.%3A%20Forensic%20speaker%20recognition.%20IEEE%20Signal%20Process.%20Mag.%2026%282%29%2C%2095%E2%80%93103%20%282009%29) (<https://scholar.google.com/scholar?q=Campbell%2C%20J.P.%2C%20Shen%2C%20W.%2C%20Campbell%2C%20W.M.%2C%20Schwartz%2C%20R.%2C%20Bonastre%2C%20J.F.%2C%20Matrouf%2C%20D.%3A%20Forensic%20speaker%20recognition.%20IEEE%20Signal%20Process.%20Mag.%2026%282%29%2C%2095%E2%80%93103%20%282009%29>)
7. Daqrouq, K., Tutunji, T.A.: Speaker identification using vowels features through a combined method of formants, wavelets, and neural network classifiers. *Appl. Soft Comput.* **27**, 231–239 (2015)
[CrossRef](https://doi.org/10.1016/j.asoc.2014.11.016) (<https://doi.org/10.1016/j.asoc.2014.11.016>)
[Google Scholar](http://scholar.google.com/scholar_lookup?title=Speaker%20identification%20using%20vowels%20features%20through%20a%20combined%20method%20of%20formants%2C%20wavelets%2C%20and%20neural%20network%20classifiers&author=K.%20Daqrouq&author=TA.%20Tutunji&journal=Appl.%20Soft%20Comput.&volume=27&pages=231-239&publication_year=2015) (http://scholar.google.com/scholar_lookup?title=Speaker%20identification%20using%20vowels%20features%20through%20a%20combined%20method%20of%20formants%2C%20wavelets%2C%20and%20neural%20network%20classifiers&author=K.%20Daqrouq&author=TA.%20Tutunji&journal=Appl.%20Soft%20Comput.&volume=27&pages=231-239&publication_year=2015)
8. Day, P., Nandi, A.K.: Robust text-independent speaker verification using genetic programming. *IEEE Trans. Audio Speech Lang. Process.* **15**(1), 285–295 (2007)
[CrossRef](https://doi.org/10.1109/TASL.2006.876765) (<https://doi.org/10.1109/TASL.2006.876765>)
[Google Scholar](http://scholar.google.com/scholar_lookup?title=Robust%20text-independent%20speaker%20verification%20using%20genetic%20programming&author=P.%20Day&author=AK.%20Nandi&journal=IEEE%20Trans.%20Audio%20Speech%20Lang.%20Process.&volume=15&issue=1&pages=285-295&publication_year=2007) (http://scholar.google.com/scholar_lookup?title=Robust%20text-independent%20speaker%20verification%20using%20genetic%20programming&author=P.%20Day&author=AK.%20Nandi&journal=IEEE%20Trans.%20Audio%20Speech%20Lang.%20Process.&volume=15&issue=1&pages=285-295&publication_year=2007)
9. Devika, A., Sumithra, M., Deepika, A.: A fuzzy-GMM classifier for multilingual speaker identification. In: *2014 International Conference on Communications and Signal Processing (ICCSP 2014)*, pp. 1514–1518. IEEE (2014)
[Google Scholar](https://scholar.google.com/scholar?q=Devika%2C%20A.%2C%20Sumithra%2C%20M.%2C%20Deepika%2C%20A.%3A%20A%20fuzzy-GMM%20classifier%20for%20multilingual%20speaker%20identification.%20In%3A%202014%20International%20Conference%20on%20Communications%20and%20Signal%20Processing%20%28ICCSP%202014%29%2C%20pp.%201514%E2%80%931518.%20IEEE%20%282014%29) (<https://scholar.google.com/scholar?q=Devika%2C%20A.%2C%20Sumithra%2C%20M.%2C%20Deepika%2C%20A.%3A%20A%20fuzzy-GMM%20classifier%20for%20multilingual%20speaker%20identification.%20In%3A%202014%20International%20Conference%20on%20Communications%20and%20Signal%20Processing%20%28ICCSP%202014%29%2C%20pp.%201514%E2%80%931518.%20IEEE%20%282014%29>)

10. Hansen, J.H., Hasan, T.: Speaker recognition by machines and humans: a tutorial review. *IEEE Signal Process. Mag.* **32**(6), 74–99 (2015)
CrossRef (<https://doi.org/10.1109/MSP.2015.2462851>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Speaker%20recognition%20by%20machines%20and%20humans%3A%20a%20tutorial%20review&author=JH.%20Hansen&author=T.%20Hasan&journal=IEEE%20Signal%20Process.%20Mag.&volume=32&issue=6&pages=74-99&publication_year=2015)
11. Hu, Y., Wu, D., Nucci, A.: Fuzzy-clustering-based decision tree approach for large population speaker identification. *IEEE Trans. Audio Speech Lang. Process.* **21**(4), 762–774 (2013)
CrossRef (<https://doi.org/10.1109/TASL.2012.2234113>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Fuzzy-clustering-based%20decision%20tree%20approach%20for%20large%20population%20speaker%20identification&author=Y.%20Hu&author=D.%20Wu&author=A.%20Nucci&journal=IEEE%20Trans.%20Audio%20Speech%20Lang.%20Process.&volume=21&issue=4&pages=762-774&publication_year=2013)
12. Maher, R.C.: Audio forensic examination. *IEEE Signal Process. Mag.* **26**(2), 84–94 (2009)
CrossRef (<https://doi.org/10.1109/MSP.2008.931080>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Audio%20forensic%20examination&author=RC.%20Maher&journal=IEEE%20Signal%20Process.%20Mag.&volume=26&issue=2&pages=84-94&publication_year=2009)
13. Morrison, G.S., Sahito, F.H., Jardine, G., Djokic, D., Clavet, S., Berghs, S., Dorny, C.G.: Interpol survey of the use of speaker identification by law enforcement agencies. *Forensic Sci. Int.* **263**, 92–100 (2016)
CrossRef (<https://doi.org/10.1016/j.forsciint.2016.03.044>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Interpol%20survey%20of%20the%20use%20of%20speaker%20identification%20by%20law%20enforcement%20agencies&author=GS.%20Morrison&author=.%20FH.%20Sahito&author=G.%20Jardine&author=D.%20Djokic&author=S.%20Clavet&author=S.%20Berghs&author=CG.%20Dorny&journal=Forensic%20Sci.%20Int.&volume=263&pages=92-100&publication_year=2016)
14. Nemati, S., Basiri, M.E.: Text-independent speaker verification using ant colony optimization-based selected features. *Expert Syst. Appl.* **38**(1), 620–630 (2011)
CrossRef (<https://doi.org/10.1016/j.eswa.2010.07.011>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Text-independent%20speaker%20verification%20using%20ant%20colony%20optimization-based%20selected%20features&author=S.%20Nemati&author=ME.%20Basiri&journal=Expert%20Syst.%20Appl.&volume=38&issue=1&pages=620-630&publication_year=2011)
15. Univaso, P., Ale, J.M., Gurlekian, J.A.: Data mining applied to forensic speaker identification. *IEEE Lat. Am. Trans.* **13**(4), 1098–1111 (2015)
CrossRef (<https://doi.org/10.1109/TLA.2015.7106363>)

- Google Scholar (http://scholar.google.com/scholar_lookup?title=Data%20mining%20applied%20to%20forensic%20speaker%20identification&author=P.%20Univaso&author=JM.%20Ale&author=JA.%20Gurlekian&journal=IEEE%20Lat.%20Am.%20Trans.&volume=13&issue=4&pages=1098-1111&publication_year=2015)
16. Wu, J.D., Tsai, Y.J.: Speaker identification system using empirical mode decomposition and an artificial neural network. *Expert Syst. Appl.* **38**(5), 6112–6117 (2011)
CrossRef (<https://doi.org/10.1016/j.eswa.2010.11.013>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Speaker%20identification%20system%20using%20empirical%20mode%20decomposition%20and%20an%20artificial%20neural%20network&author=JD.%20Wu&author=YJ.%20Tsai&journal=Expert%20Syst.%20Appl.&volume=38&issue=5&pages=6112-6117&publication_year=2011)
17. Wu, Z., Evans, N., Kinnunen, T., Yamagishi, J., Alegre, F., Li, H.: Spoofing and countermeasures for speaker verification: a survey. *Speech Commun.* **66**, 130–153 (2015)
CrossRef (<https://doi.org/10.1016/j.specom.2014.10.005>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Spoofing%20and%20countermeasures%20for%20speaker%20verification%20a%20survey&author=Z.%20Wu&author=N.%20Evans&author=T.%20Kinnunen&author=J.%20Yamagishi&author=F.%20Alegre&author=H.%20Li&journal=Speech%20Commun.&volume=66&pages=130-153&publication_year=2015)
18. Xing, Y., Li, H., Tan, P.: Hierarchical fuzzy speaker identification based on FCM and FSVM. In: 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 311–315. IEEE (2012)
Google Scholar (<https://scholar.google.com/scholar?q=Xing%2C%20Y.%2C%20Li%2C%20H.%2C%20Tan%2C%20P.%3A%20Hierarchical%20fuzzy%20speaker%20identification%20based%20on%20FCM%20and%20FSVM.%20In%3A%202012%209th%20International%20Conference%20on%20Fuzzy%20Systems%20and%20Knowledge%20Discovery%20%28FSKD%29%2C%20pp.%20311%E2%80%93315.%20IEEE%20%282012%29>)
19. Zhao, X., Shao, Y., Wang, D.: Casa-based robust speaker identification. *IEEE Trans. Audio Speech Lang. Process.* **20**(5), 1608–1616 (2012)
CrossRef (<https://doi.org/10.1109/TASL.2012.2186803>)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Casa-based%20robust%20speaker%20identification&author=X.%20Zhao&author=Y.%20Shao&author=D.%20Wang&journal=IEEE%20Trans.%20Audio%20Speech%20Lang.%20Process.&volume=20&issue=5&pages=1608-1616&publication_year=2012)

Copyright information

© Springer International Publishing AG 2017

About this paper

Cite this paper as:

Camacho S., Renza D., Ballesteros L. D.M. (2017) A Semi-supervised Speaker Identification Method for Audio Forensics Using Cochleagrams. In: Figueroa-García J., López-Santana E., Villa-Ramírez J., Ferro-Escobar R. (eds) Applied Computer Sciences in Engineering. WEA 2017. Communications in Computer and Information Science, vol 742. Springer, Cham

- DOI (Digital Object Identifier) https://doi.org/10.1007/978-3-319-66963-2_6
- Publisher Name Springer, Cham
- Print ISBN 978-3-319-66962-5
- Online ISBN 978-3-319-66963-2
- eBook Packages [Computer Science](#)

- [About this book](#)
- [Reprints and Permissions](#)

Personalised recommendations

SPRINGER NATURE

© 2017 Springer International Publishing AG. Part of [Springer Nature](#).

Not logged in Not affiliated 77.111.246.8