Abstract

Recognition systems are commonly designed to authenticate users at the access control levels of a system. A number of voice recognition methods have been developed using a pitch estimation process which are very vulnerable in low Signal to Noise Ratio (SNR) environments thus, these programs fail to provide the desired level of accuracy and robustness. Also, most
Text-independent speaker recognition programs are incapable of coping with unauthorized attempts to gain access by tampering with the samples or reference database. The proposed text-independent voice recognition system makes use of multilevel cryptography to preserve data integrity while in transit or storage. Encryption and decryption follow a transform based approach layered with pseudorandom noise addition whereas for pitch detection, a modified version of the autocorrelation pitch extraction algorithm is used. The experimental results show that the proposed algorithm can decrypt the signal under test with exponentially reducing Mean Square Error over an increasing range of SNR. Further, it outperforms the conventional algorithms in actual identification tasks even in noisy environments. The recognition rate thus obtained using the proposed method is compared with other conventional methods used for speaker identification.

Reference

- A. K. Jain, A. Ross, S. Prabhakar, "An introduction to biometric recognition," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No.1, Jan. 2004, 4- 20.
- R. Clarke, “Human Identification in Information Systems: Management Challenges and Public Policy Issues,” Information Technology & People, Vol. 7, No. 4, 6–37, 1994.
- T. Kinnunen, H. Lib, “An Overview of Text-Independent Speaker Recognition: from Features to Supervectors,” 1-3.
- X. Huang, A. Acero, and H. W. Hon, Spoken Language Processing: a Guide to Theory, Algorithm, and System Development, Prentice-Hall, New Jersey, 2001.
- K. Theocharoulis, I. Papaefstathiou, C. Manifavas, “Implementing Rainbow Tables in High-End FPGAs for Super-Fast Password Cracking,” International Conference on Field Programmable Logic and Applications, 2010, 145-150.
- X. Huang, A. Acero, and H. Hon, Spoken Lang. Process.. Upper Saddle River, NJ: Prentice-Hall, 2001.
- A. Higgins, L. Bahler, and J. Porter, “Speaker verification using randomized phrase prompting,” Digital Signal Processing 1 (April 1991), 89–106.
- M. Sondhi, “New methods of pitch extraction,” IEEE Transactions on Audio and Electroacoustics, Vol. 16, No. 2, Jun 1968, 262-266.
- A. P. Varga and R. K. Moore, “Hidden Markov model decomposition of speech and noise,” in Proc. ICASSP’90, 1990, 845-848.
- T. Kinnunen, Licentiate’s Thesis, “Spectral Features for Automatic Text-Independent Speaker Recognition,” Department of Computer Science, University of Joensuu, December 2003, 5-11, 2-3.
- B. Goldburg, S. Sridharan, E. Dawson, “Design and cryptanalysis of transform-based analog speech scramblers,” IEEE Journal on Selected Areas in Communications, June 1993, 735-744.
- A. Menezes , P. Oorschot , S. Vanstone , R. Rivest, Handbook of Applied Cryptography, CRC Press, 1996, 169-179, 180-190.
- M. J. Ross, H. L. Shaffer, A. Cohen, R. Freudberg, and H. J. Manley, “Average magnitude difference function pitch extractor,” IEEE Transactions on Acoustics, Speech, Signal Processing, Vol. ASSP-22, Oct. 1974, 353-362.
- L. R. Rabiner, M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal, “A comparative performance study of several pitch detection algorithms,” IEEE Transactions on Audio, Signal,
and Speech Processing 24, 1976, 399-417.

- H. Kobayashi and T. Shimamura, "A weighted autocorrelation method for pitch extraction of noisy speech," IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, 2000, pp.1307-1310.

- L. Tan and M. Karnjanadecha, “Pitch Detection Algorithm: Autocorrelation and AMDF”, International Symposium on Communications and Information Technologies (ISCIT 2003), 1-4.

- L. R. Rabiner, B.H. Juang, Fundamentals of speech recognition, Prentice Hall, 1993.

- T. F. Quatieri, Discrete time speech signal processing, Prentice Hall, 2002.

- S. Jungpa, S. Hong, J. Gu, M. Kim, I. Baek, Y. Kwon, K. Lee, Sung-I Yang, "New speaker recognition feature using correlation dimension," Proceedings of the IEEE International Symposium on Industrial Electronics, 2001, ISIE 2001, Vol.1, 2001, pp.505-507.

- S. Kim, M. Ji, H. Kim, “Robust speaker recognition based on filtering in autocorrelation domain and sub-band feature recombination,” Pattern Recognition Letters, Elsevier Science, Vol. 31, No. 7, May 2010, 593-599.

- NOISEX-92 Database, Carnegie Mellon University, http://www.speech.cs.cmu.edu/comp.speech/Section1/Data/noisex.html, July 2011.

- L. R. Rabiner, “On the use of autocorrelation analysis for pitch detection,” IEEE Transactions on Acoustics, Speech, Signal Processing, Vol. ASSP-25, No. 1, 24–33, Feb. 1977.

- Y. J. Kim, and J. H. Chung, “Pitch synchronous cepstrum for robust speaker recognition over telephone channels,” Electronics letters, Vol. 40, No. 3, 207–209, 2004.

Index Terms

Computer Science  Pattern Recognition

Key words

Speaker Individuality  Text-independence  Pitch

Extraction

Voice Recognition

Autocorrelation
