

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 170](#)

-
[Number 2](#)

Year of Publication: 2017

Authors:

Ritesh A. Magre, Ajit S. Ghodke

10.5120/ijca2017914713

{bibtex}2017914713.bib{/bibtex}

Abstract

Recently lot of work has been done on audio visual speech recognition but less work has been done on visual speech and speaker recognition. This research belongs to human computer interaction (HCI) domain. HCI makes human computer interaction simple. This paper represents the creating of database of visual speech and speaker in English language and preprocessing of it to improve recognition accuracy. We have studied Tulipse1 database, AV Database and CUAVE Database on the basis of these different databases we have created our own database. This is useful for all researchers those are working HCI domain particularly Visual Speech and Speaker Recognition.

References

1. . Jana Trojanová, Marek Hruš z, Pavel Campr, Milosˇ Zˇ elezny “Design and Recording of Czech Audio-Visual Database with Impaired Conditions for Continuous Speech Recognition” Department of Cybernetics, Faculty of Applied Sciences, University OF West Bhoemia

Univerzitetni 22, 306 14, Plzen, Czech Republic.

2. . Paterson, E. K.: Audio Visual Speech Recognition for Difficult Environments. Ph.D. thesis, Clemson University(2002).
3. . Weber, K., Ikbali, S., Bengio, S., and Bourlard, H.: Robust Speech Recognition and Feature Extraction Using HMM2. *Computer Speech & Language*, 17 (2003) 2–3.
4. J.R. Movellan, “Visual speech recognition with stochastic networks,” in *Advances in Neural Information Processing Systems*, G. Tesauro, D. Touretzky, and T. Leen, Eds., vol. 7. MIT Press, Cambridge, 1995.
5. I. Matthews, Features for Audio-Visual Speech Recognition, Ph.D. thesis, School of Information Systems, University of East Anglia, UK, 1998.
6. C. C. Chibelushi, S. Gandon, J. S. D. Mason, F. Deravi, and R. D. Johnston, “Design issues for a digital audiovisual integrated database,” in *IEE Colloquium on Integrated Audio-Visual Processing for Recognition, Synthesis and Communication*, Savoy Place, London, Nov. 1996, number 1996/213, pp. 7/1–7/7.
7. A Adjoudani and C Benoit, On the integration of auditory and visual parameters in an HMM-based ASR, In *Stork and Henneke* [11], pages 461-471 .
8. C Bergler, H Hild , S Manke and A Waibel, Improving connected letter recognition by lipreading, In *Proc International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 557-560, Minneapolis, 1993.
9. M T Chan, Y Zhang and T S Huang, Real time lip tracking and bimodal continuous speech recognition, In *Proc IEEE 2nd workshop on multimedia signal processing*, pages 65-70, Redondo Beach, 1998.
10. C C Chibelushi, F Deravi and J S D Mason, Survey of audio-visual speech database, technical report, Department of electrical and electronic engineering, University of Wales, Swansea, 1996.
11. I Matthews, T Cootes, S Cox, R Harvey and J A Bangham, Lipreading using shape, shading and scale, In *Proceedings of Workshop on Audio visual speech processing* , pages 73-78, Terrigal 1998.
12. K Messar, J Matas, J Kittler, J Luetten and G Maitre, XM2VTS: The extended M2VTS database, In *Proc. 2nd International conference on audio and video based biometric person authentication (AVBPA)* page 72-76, Washington 1999.
13. J R Movellan and G Chadderdon, Channel seperability in audio visual integration of speech: A Bayesian approach, In *Stork and Hennecke* [11], pages 473-487.
14. E D Petjan, Automatic Lipreading to enhance speech recognition, In *Proc Global Telecommunication Conference (GLOBECOM)*, pages 265-272, Atlanta 1984.
15. P Teissier, J Robert-Ribes and J L Schwartz, Comparing models for audio visual fusion in noisy vowel recognition tasks, *IEEE transaction on speech and audio processing*, 7(6): 629-642, 1999.
16. CUAVE Database, Clemson university database for Audio visual experiments <http://www.ece.clemson.edu/speech/cuave.htm>.

Index Terms

Computer Science

Pattern Recognition

Keywords

Speech, Visual Speech, Lip Reading, Lip Database, Visual Speech Recognition, Speaker Recognition, Face detection, Lip Cropping .