Dublin Institute of Technology
ARROW@DIT

Conference papers

Digital Media Centre

2006-10-01

The Use of Task Based Mood-Induction Procedures to Generate High Quality Emotional Assets

Brian Vaughan
Dublin Institute of Technology, brian.vaughan@dit.ie

Charlie Cullen
Dublin Institute of Technology, charlie.cullen@dit.ie

Spyros Kousidis
Dublin Institute of Technology, spyros.kousidis@dit.ie

Yi Wang
Dublin Institute of Technology, yi.wang@dit.ie

Ciaran McDonnell
Dublin Institute of Technology, ciaran.mcdonnell@dit.ie

See next page for additional authors

Follow this and additional works at: http://arrow.dit.ie/dmccon

Part of the Computer Sciences Commons

Recommended Citation

Vaughan, B. et al (2006) The Use of Task Based Mood-Induction Procedures to Generate High Quality Emotional Assets. *IT&T 2006 Conference,* Carlow, Ireland. October 25 - 26.

This Conference Paper is brought to you for free and open access by the Digital Media Centre at ARROW@DIT. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@DIT. For more information, please contact yvonne.desmond@dit.ie, arrow.admin@dit.ie.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License

Authors
Brian Vaughan, Charlie Cullen, Spyros Kousidis, Yi Wang, Ciaran McDonnell, and Dermot Campbell

This conference paper is available at ARROW@DIT: http://arrow.dit.ie/dmccon/36

The Use of Task Based Mood-Induction Procedures to Generate High Quality Emotional Assets.

B. Vaughan , C. Cullen, , S. Kousidis , Wang Yi , C. McDonnell and D. Campbell
DMC, DIT, Aungier Street, Dublin 2
brian.vaughan@dit.ie

Abstract

Detecting emotion in speech is important in advancing human-computer interaction, especially in the area of speech synthesis [1]. This poster details experimental procedures based on Mood Induction Procedure 4 [2], *using performance related tasks to engender natural emotional responses in participants. These tasks are aided or hindered by the researcher to illicit the desired emotional response. These responses will then be recorded and their emotional content graded to form the basis of an emotional speech corpus. This corpus will then be used to develop a rule-set for basic emotional dimensions in speech.*

1. Introduction

Research has shown that there is potential for grading the basic emotional content of natural speech using the dimensions of activation and evaluation [3] . There is considerable debate concerning the methods used for obtaining emotional content, and distinction is made between natural [4] and acted speech [5], mainly as regards to their level of authenticity. In this research, Mood Induction Procedures (MIP's) [2] are used to stimulate emotions in controlled situations. In these experiments, participants sit in isolation booths and must co-operate to complete a timed task [6].

2. Methods

In order to illicit natural emotional responses, the researcher aids or hinders the participants in their co-operative task. This is achieved by varying the instructions given or the materials or time available. The use of sound isolation booths serves to prevent outside distractions from affecting the participants as well as allowing the dialogue to be recorded as two separate high quality audio streams, specifically at a 24bit/192Khz [7] professional resolution. These high quality audio assets obtained in a properly isolated acoustic environment will allow a robust and effective analysis to be preformed with reliable acoustic results. The assets obtained will subsequently be graded according to their emotional content, and then assessed so that they can be segregated into relevant emotional dimensions.

3. Results

There are a number of methods [3] that can be used to evaluate the emotional dimensions of the recorded audio, but listening tests are considered to be the most statistically robust. The listening tests will be carried out

by participants who did not take part in the co-operative task based experiment. The assets will be graded using the FeelTrace [8] tool in order to specify the emotional dimension of each clip.

4. Conclusion

Using the FeelTrace tool to analyse and grade the emotional content of the audio clips allows for a robust and justifiable set of results that can then be used to specify the emotional dimension of the clips in the corpus. These emotionally graded assets can then be analysed further in order to create a rule set for basic emotional dimensions in speech.

5. References

1. Schroeder, M., *Speech and Emotion Research. An Overview of Research Frameworks and a Dimensional Approach to Emotional Speech Synthesis*, in *Faculty of Philosophy*. 2004, Universit¨at des Saarlandes. p. 288.
2. Gerrards-Hesse, A., K. Spies, and F.W. Hesse, *Experimental inductions of emotional states and their effectiveness: A review.* British Journal of Psychology, 1994. **85**(1): p. 55-78.
3. Cowie, R., et al., *Emotion recognition in human-computer interaction.* IEEE Signal Processing Magazine, 2001. **18**(1): p. 32-80.
4. Campbell, N. *Databases of emotional speech*. in *ISCA Workshop on Speech and Emotion*. 2000. Northern Ireland.
5. Roach, P., et al., *Transcription of Prosodic and Paralinguistic Features of Emotional Speech.* Journal of the International Phonetic Association, 1998(28): p. 83-94.
6. Kehrein, R. *The prosody of authentic emotions.* in *Speech Prosody*. 2002. Aix-en-Provence, France.
7. Watkinson, J., *Art of Digital Audio*, ed. M. Burlington. Vol. Third Edition. 2000: Focal Press.
8. Cowie, R., et al. *'FEELTRACE': An instrument for recording perceived emotion in real time*. in *ISCA Workshop on Speech and Emotion*. 2000. Northern Ireland.