





EMOVO Corpus: an Italian Emotional Speech Database

GIOVANNI COSTANTINI^{1,2}, IACOPO IADAROLA⁴, ANDREA PAOLONI³, MASSIMILIANO TODISCO^{1,3}

¹ Dipartimento di Ingegneria Elettronica, Università "Tor Vergata", Roma ² Istituto di Acustica e Sensoristica "O. M. Corbino", Roma ³ Fondazione "Ugo Bordoni", Roma ⁴ Dipartimento di Linguistica, Università "La Sapienza", Roma

- A corpus of emotional voice is essential to build systems of emotion recognition and systems of text-to-speech synthesis with emotion rendering.
- The introduction of emotive voice in man-machine communication finds its most important application in building user interfaces for many services, in particular call centers.

EMOVO Corpus

Applications

- Lie detector (X13-VSA PRO): a system capable to assess whether a user is lying.
- E-learning.
- Automotive industry, in particular to enhance safety of the driver
- Video games
- Man-to-robot communication
- Investigation on human brain malfunctions
- Interaction with audio and visual entertainment systems

Validation

- The listeners were organized in two groups of 12 persons each, hosted in two different labs
- Subjects were asked to guess between 2 emotions
- From each actor's voice 2 nonsense sentences were proposed: this prevented semantic content from biasing the guess among 7 outcomes
- Each subject listened to the signals from a male and a female actor for a total of 84 tests
- The results of the two labs were statistically compatible. Hence, one table only is presented in one table (see Table I)

6 Emotions disgust, joy, fear, anger, surprise, sadness + neutral	6 Actors 3 males 3 females	14 Phrases
Nonsense sentences 'The strong house wants with bread' 'The force is up and red garlic' 'The cat is flowing in pear'	ntically neutral sentences One forty-three ten th	mbers, of foods) housand fifty-seven twenty' g of lamb limoncello'
Short sentences 'Workers get up early' 'Firefighters are equipped with a gun' 'The waterfall makes a lot of noise'	Long sentences 'A moment later he hath walked and stumbled' "Now I take the sweatshirt and go for a walk' The next fall Tony will leave for Spain in the first half of October' 'I would like the telephone number of Mr. Piatti'	Interrogative sentences 'Saturday night, what will?' 'Bring with you that thing?'

ole I			RECOGNIZED EMOTION							
Table I	NEUTRAL	DISGUST	JOY	FEAR	ANGER	SURPRISE	SADNESS			
NEUTRAL	93%	1%	0%	0%	4%	0%	2%			
DISGUST	3%	67%	2%	6%	10%	6%	6%			
JOY	2%	4%	65%	7%	7%	10%	4%			
FEAR	2%	7%	2%	74%	3%	3%	9%			
ANGER	1%	1%	1%	3%	92%	1%	1%			
SURPRISE	1%	3%	4%	1%	1%	81%	9%			
SADNESS	2%	2%	1%	3%	0%	0%	92%			
	JOY FEAR ANGER URPRISE	JOY 2% FEAR 2% URPRISE 1%	JOY 2% 4% FEAR 2% 7% ANGER 1% 1% URPRISE 1% 3%	DISGUST 3% 67% 2% JOY 2% 4% 65% FEAR 2% 7% 2% ANGER 1% 1% 1% URPRISE 1% 3% 4%	DISGUST 3% 67% 2% 6% JOY 2% 4% 65% 7% FEAR 2% 7% 2% 74% ANGER 1% 1% 1% 3% URPRISE 1% 3% 4% 1%	DISGUST 3% 67% 2% 6% 10% JOY 2% 4% 65% 7% 7% FEAR 2% 7% 2% 74% 3% ANGER 1% 1% 1% 3% 92% URPRISE 1% 3% 4% 1% 1%	DISGUST 3% 67% 2% 6% 10% 6% JOY 2% 4% 65% 7% 7% 10% FEAR 2% 7% 2% 74% 3% 3% ANGER 1% 1% 1% 3% 92% 1% URPRISE 1% 3% 4% 1% 1% 81%			

Recording data

- Data were recorded in the laboratories of the Fondazione Ugo Bordoni
- Recording was captured at 48 kHz, 16 bit, PCM stereo way format
- 588 records (14 phrases x 6 actors x (6 + 1) emotional states)
- Total duration of recorded speech: around 60 minutes

Conclusion

- EMOVO is the first corpus of emotional voice for Italian
- Speech emotion has been simulated by actors
- The EMOVO corpus can be downloaded at:

http://voice.fub.it/EMOVO

References

- 1. Cowie, R., Cornelius, R.R. (2003). Describing the Emotional States that Are Expressed in Speech. Speech Communication, 40(1,2), 2-32.
- 2. Davitz, J.R., 1964. A review of research concerned with facial and vocal expressions of emotion. In Davitz, J.R. (Ed.), The Communication of Emotional Meaning. McGraw-Hill, NY, pp. 13–29.
- 3. Douglas-Cowie E., Campbell N., Cowie R. & Roach P. (2003), Emotional speech: towards a new generation of databases, Speech communication, 40, pp. 33-60.
- Giovannella C., Floris D., Paoloni A., An exploration on possible correlations among perception and physical characteristics of of EMOVO's emotional portrayals. IxD&A Journal, N. 15, 2012, pp. 102-111.
- 5. Giovannella C., Floris D., Paoloni A., Transmission of vocal emotion: Do we have to care about the listener? The case of the Italian speech corpus EMOVO, 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009.
- 6. ladarola, I. (2009), EMOVO, database di parlato emotivo per l'italiano, Atti del 4° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Arcavacata di Rende (CS), 3-5 dicembre 2007.
- 7. Kuroda I., Fujiwara O., Okamura N. & Utsuki N. (1976), Method for determining pilot stress through analysis of voice communication, Aviation, space, and environmental medicine, 47, pp. 528-533.
- 8. Scherer, K.R., (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, 99 (2), 143–165.
- 9. S. Ramakrishnan (2012). Recognition of Emotion from Speech: A Review, Speech Enhancement, Modeling and Recognition Algorithms and Applications, Dr. S Ramakrishnan (Ed.), InTech, DOI: 10.5772/39246.
- 10. Costantini, G., Paoloni, A., Todisco, M., (2013). Quantifying the Value of Subjective and Objective Speech Intelligibility Assessment in Forensic Applications. In WSEAS Transactions on Systems, Volume 12, Issue 11, pp. 561-572, November 2013.