

Spraakherkenning en synthese 2009

David Weenink

December 8, 2009

Inleiding

Kort overzicht van de behandelde stof per college. Dit document zal in de loop van het semester elke week aangepast worden.

De behandelde stof zal voor een groot deel te vinden zijn in Jurafsky and Martin [2009] en in Weenink [2009a].

College: Woensdagen van 9.00-11.00u in P.014 / E.20

Practicum: Donderdagen van 11.00-13.00u in P1.26

1 Week 1, 3 september: Het spraaksignaal

De stof: Levelt [1994], Jurafsky and Martin [2009, ch 7.1-7.3], Weenink [2009a, Ch. 1, 2]

De presentatie: SpraakherkenningEnSynthese_week_01.pdf

2 Week 2, 10 september: Meten aan spraak F0

Stof: Overzicht spraakproductie in Jurafsky and Martin [2009, ch. 7.4], toonhoogteanalyse in Boersma [1993].

De presentatie: SpraakherkenningEnSynthese_week_02.pdf

3 Week 3, 16 september: Meten aan spraak: spectrum, formant

Stof: Jurafsky and Martin [2009, ch 7.4], Weenink [2009a, ch. 3, 4, 9]

De presentatie: als week 2.

4 Week 4, 23 september: Corpora

De stof: Jurafsky and Martin [2009, ch 7.5]

De presentatie: SpraakherkenningEnSynthese_week_04.pdf

5 Week 5, 30 september: TTS en ASR; SpeakGoodChinese

Stof: Jurafsky and Martin [2009, sectie 3.11], Weenink et al. [2007]

De presentatie: SpraakherkenningEnSynthese_week_05.pdf

6 Week 6, 7 oktober: Synthese

Stof: Jurafsky and Martin [2009, ch 8]

De presentatie: SpraakherkenningEnSynthese_week_06.pdf

19 oktober, tentamen 14-17.00u

Stof: Literatuur en presentaties van week 1 t/m 6.

De uitwerkingen van de eerste drie opgaves staan hier.

7 Week 7: 28 oktober, meer synthese

Stof: Jurafsky and Martin [2009, ch 8.5, 8.6]

De presentatie: [SpraaakherkenningEnSynthese_week_07.pdf](#)

8 Week 8: 4 november, Formant synthese

Stof 1-ste deel: Weenink [2009a, ch 8, KlattGrid], Weenink [2009b], Klatt [1980], Klatt and Klatt [1990]

2-de deel: Bespreking tussentoets van 19 oktober j.l.

De presentatie: [SpraaakherkenningEnSynthese_week_08.pdf](#)

9 Week 9: 11 november, Inleiding Spraakherkenning, DTW

Stof: MFCC's en DTW, Jurafsky and Martin [2009, ch 9.3] en praat scripting Weenink [2009b, app. C & D]

Bespreking opgaves 4, 5 en 6.

De presentatie: [SpraaakherkenningEnSynthese_week_09.pdf](#)

10 Week 10: 18 november, Inleiding spraakherkenning, HMM

Stof: ASR Jurafsky and Martin [2009, ch. 5.5, ch. 9] HMM's

De presentatie: [SpraaakherkenningEnSynthese_week_10.pdf](#)

11 Week 11: 1 december, ASR en TTS

Stof: De Wachter et al. en Tokuda et al. [2002]

De presentatie: [SpraaakherkenningEnSynthese_week_11.pdf](#)

16 december, tentamen

De stof van de colleges week 7 t/m 11 en de opgaves.

Belangrijk

Om een voldoende voor dit vak te scoren moet:

1. het gemiddelde van beide tentames voldoende zijn,
2. de opgaves ingeleverd zijn en als voldoende beoordeeld. Zoals afgesproken moeten alle opgaves ingeleverd zijn twee weken nadat ze op BlackBoard zijn gezet. Omdat veel mensen dit niet op tijd gaan halen hebben jullie hiervoor nog gelegenheid tot twee weken na het tentamen!

References

Paul Boersma. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. 17:97–110, 1993. URL http://www.fon.hum.uva.nl/paul/papers/Proceedings_1993.pdf.

M. De Wachter, M. Matton, K. Demuyne, P. Wambacq, R. Cools, and D. Van Compernelle. Template-Based Continuous Speech Recognition. *IEEE TRANSACTIONS ON AUDIO, SPEECH & LANGUAGE PROCESSING*, 15(4):1377–1390. ISSN 1558-7916. URL http://www.fon.hum.uva.nl/david/ba_shs/2009/DeWachterEtAl-2002.pdf.

Daniel Jurafsky and James H. Martin. *Speech and Language Processing*. Prentice Hall, 2nd edition, 2009.

Dennis H. Klatt. Software for a cascade/parallel formant synthesizer. 67:971–995, 1980. URL http://www.fon.hum.uva.nl/david/ba_shs/2009/klatt-1980.pdf.

- Dennis H. Klatt and Laura C. Klatt. Analysis, synthesis, and perception of voice quality variations among female and male talkers. 87:820–857, 1990. URL http://www.fon.hum.uva.nl/david/ba_shs/2009/klatt-1990.pdf.
- W.J.M. Levelt. *The Skill of Speaking*, volume 1 of *International perspectives on psychological science*, pages 89–103. Lawrence Erlbaum Associates, 1994. URL <http://hdl.handle.net/2066/15531>.
- Keiichi Tokuda, Heiga Zen, and Alan W. Black. An HMM-based speech synthesis system applied to English, 2002. URL <http://www.cs.cmu.edu/~awb/papers/IEEE2002/hmmenglish.pdf>.
- D. Weenink, G. Chen, Z. Chen, S. de Konink, D. Vierkant, E. van Hagen, and R.J.J.H. van Son. Learning tone distinctions for Mandarin Chinese. In *Proceedings of Interspeech 2007*, pages 950–953, Antwerp, Belgium, August 2007. URL http://www.fon.hum.uva.nl/david/ba_shs/2009/weenink-et-al-sgc-2007.pdf.
- David Weenink. Speech signal processing with praat, 2009a. URL <http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf>.
- David Weenink. The KlattGrid acoustic speech synthesizer. *Proceedings Interspeech 2009*, pages 2059–2062, 2009b. URL http://www.fon.hum.uva.nl/david/ba_shs/2009/is2009_klattgrid.pdf.