DANSA

A compact statistical model of coarticulatory dynamics

Veena Singampalli and Philip JB Jackson

Centre of Vision, Speech and Signal Processing (CVSSP)
University of Surrey, UK

11 January 2007

Overview

- Dealing with context-dependency:
  - speech isn't like “beads on a string”, it's a dance!
  - coarticulation applicable in synthesis & robust ASR
  - compact models for efficient exploitation of data
- Building a model of coarticulation:
  - Introduction
  - Critical, dependent & redundant articulation
  - Modelling dependencies dynamically
  - Preliminary evaluation of generated trajectories
  - Summary

Related contributions

- Statistical trajectory-modelling techniques:
  - Tokuda (trajectory HMM), Frankel & King (LDM)
  - Ostendorf, Russell, Holmes (segment models)
  - Deng et al. (dynamic Bayesian networks)
- Articulatory kinematics:
  - Öhman, Mermelstein ("knots and beads on a string")
  - Lindblom (targets & smoothing)
  - Kaburagi, Honda, Dang, Löfqvist (phone distributions, context-dependent weights)
  - Blackburn & Young (statistics of articulator inertia)
- Biomechanical dynamics:
  - Coker (articulatory priority)
  - Saltzman, Tremblay & Ostry (task dynamics)

Articulatory correlations & rôles

Three main kinds of correlation in articulatory data:

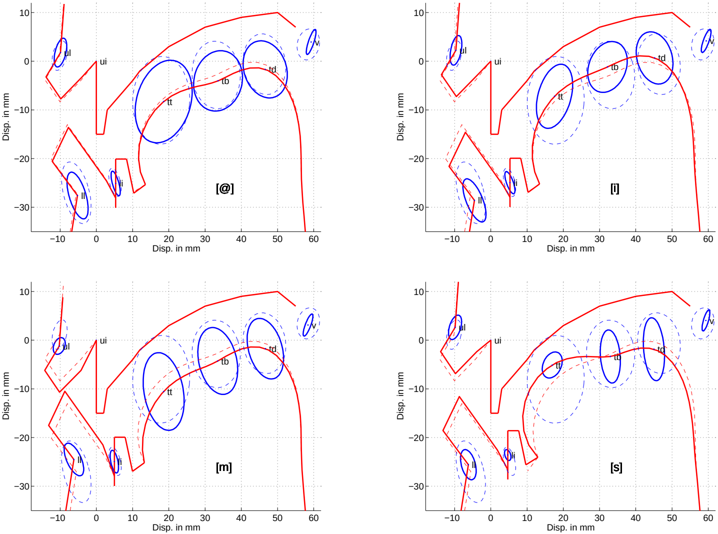
1. correlation for each articulator in **space**
2. correlations **between** articulators
3. correlations of trajectories over **time**

While speaking, articulators continually change their rôle:

- **critical** articulator
  - specified position (or gesture) is essential to phone
- **dependent** articulator
  - movement follows that of correlated critical articulator(s)
- **redundant** articulator
  - irrelevant, only constrained by surrounding context

Study relations 1 & 2 statistically to reveal coarticulation model's structure, based on Mocha data.

Phone distributions



Critical-artic. detection algorithm

1. Initialise model to grand distributions (Gaussian)
2. Compute phone distributions
3. Calculate intra- and inter-articulator correlations:
    - calculate grand correlations and test for significance
    - set low and insignificant values to zero

Univariate correlation matrix

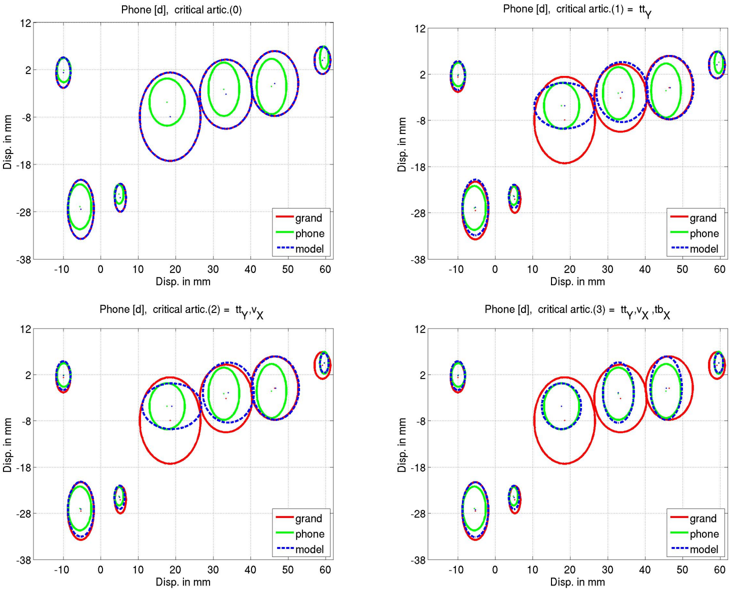
1D grand correlation matrix

	ul_x	ul_y	ll_x	ll_y	li_x	li_y	tt_x	tt_y	tb_x	tb_y	td_x	td_y	v_x	v_y
ul_x	1.00	.53	.34	-.15	.00	-.17	-.18	.00	-.16	.00	-.22	.00	.00	.00
ul_y	.53	1.00	.27	-.31	.00	.00	.00	.29	.00	.19	-.15	.00	.13	.00
ll_x	.34	.27	1.00	-.70	.61	-.55	.00	-.31	-.19	.00	-.17	.14	.11	.00
ll_y	-.15	-.31	-.70	1.00	-.49	.65	.00	.32	.14	.00	.10	-.10	-.18	.00
li_x	.00	.00	.61	-.49	1.00	-.71	.00	-.43	.00	-.36	.00	.00	.12	.00
li_y	-.17	.00	-.55	.65	-.71	1.00	.00	.60	.12	.42	.00	.00	-.12	.00
tt_x	-.18	.00	.00	.00	.00	.00	1.00	.00	.90	.00	.82	.00	.24	.19
tt_y	.00	.29	-.31	.32	-.43	.60	.00	1.00	.18	.53	.11	.00	.00	.00
tb_x	-.16	.00	-.19	.14	.00	.12	.90	.18	1.00	.00	.92	-.24	.14	.00
tb_y	.00	.19	.00	.00	-.36	.42	.00	.53	.00	1.00	.00	.75	.00	.00
td_x	-.22	-.15	-.17	.10	.00	.00	.82	.11	.92	.00	1.00	-.21	.00	.00
td_y	.00	.00	.14	-.10	.00	.00	.00	.00	-.24	.75	-.21	1.00	.00	.23
v_x	.00	.13	.11	-.18	.12	-.12	.24	.00	.14	.00	.00	.00	1.00	.81
v_y	.00	.00	.00	.00	.00	.00	.19	.00	.00	.00	.00	.23	.81	1.00

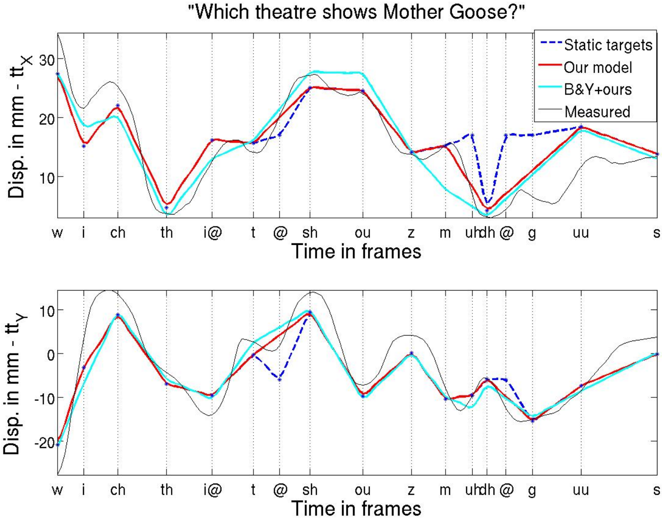
Critical-artic. detection algorithm

1. Initialise model to grand distributions (Gaussian)
2. Compute phone distributions
3. Calculate intra- and inter-articulator correlations:
   - calculate grand correlations and test for significance
   - set low and insignificant values to zero
4. Compute modified Kullback-Leibler divergences:
   - calculate std. error for phone and model pdfs
   - compute KL divergence including std. errors
5. Select articulator with max. divergence
   - test whether the value is above divergence threshold
   - set model pdf to phone pdf for critical articulator
6. Update model distributions:
   - for each dependent articulator, apply conditional mean and covariance based on critical articulator(s)
7. Repeat from step 4.

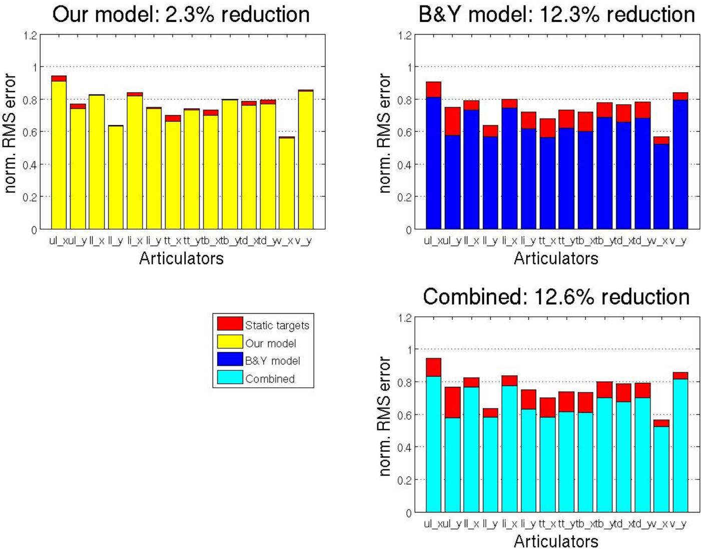
Univariate model convergence



Generation of trajectories



Average model error



Summary

- Coarticulatory model of speech dynamics
  - Tested for significant correlations (1D & 2D)
  - Detected critical articulator(s) for each phone, considering dependencies
  - Modified trajectory interpolation through redundancies
  - Preliminary resuls of model gave an improvement
- Future investigation
  - Development of trajectory interpolation (e.g., to 2D)
  - Detailed evaluation of the model's performance
  - Model incorporation into multi-level SEGREC system

Questions? 

www.ee.surrey.ac.uk/Personal/P.Jackson/Dansa