

A Non-Parametric Articulatory-to-Acoustic Conversion System for Silent Speech using Shared Gaussian Process Dynamical Models

Jose A. Gonzalez^{*1}, Phil D. Green^{†1}, Roger K. Moore^{‡1}, Lam A. Cheah^{§2}, and James M. Gilbert^{¶2}

¹Department of Computer Science, University of Sheffield

²School of Engineering, University of Hull

As part of our ongoing work to develop a silent speech interface (SSI) system for post-laryngectomy speech rehabilitation, this work presents a technique for articulatory-to-acoustic conversion using a non-parametric, statistical approach based on shared Gaussian process dynamical models (SGPDMs). In the proposed technique, simultaneous recordings of articulatory and acoustic data are used to learn a mapping between both domains using a SGPDM, which is a non-parametric model providing a shared low-dimensional embedding of the articulatory and acoustic data as well as a dynamic model in the latent space. The learned model is then used for generating an audible speech signal from captured articulatory data. In this work, articulator motion data from the lips and tongue is captured using a technique known as permanent magnet articulography, in which a set of magnets are attached to the articulators and the variations of the magnetic field generated while the user 'speaks' are sensed by a number of magnetic sensors located around the mouth. Preliminary results show that the proposed mapping is able to synthesise high-quality speech from PMA data for certain restricted tasks, but further research is needed before the technique can be applied to a real-life scenario.

*j.gonzalez@sheffield.ac.uk

†p.green@sheffield.ac.uk

‡r.k.moore@sheffield.ac.uk

§l.cheah@hull.ac.uk

¶j.m.gilbert@hull.ac.uk