K-t BLAST/k-t FOCUSS in real time imaging of the soft palate during speech

Marzena Wylezinska[1], Andreia Freitas[1,2], Malcolm Birch[1], and Marc Miquel[1,2]

[1]*Clinical Physics, Barts Health NHS Trust, London, United Kingdom,* [2]*William Harvey Research Institute, Queen Mary University of London, London, United Kingdom*

Introduction: The assessment of velopharyngeal closure is carried out routinely in the management of cleft patients [1]. In clinical practice these tests are most commonly performed using x-ray videofluoroscopy and nasendoscopy [1]. Recently MRI has been suggested as an alternative technique for dynamic imaging of the vocal tract [2-5] thus, eliminating x-ray exposure. However, early real-time MR studies of the soft palate during speech have been limited by insufficient frame rates [2,3]. Recent developments using spiral advanced real-time imaging techniques were proposed [4], but these may not be available to standard clinical centres. Another approach used to accelerate dynamic imaging has been to exploit spatial-temporal data correlation such as k-t BLAST/k-t SENSE [5,6]. In particular, k-t BLAST protocol has been proved useful for accelerated cardiac imaging [5] and has been implemented on clinical scanners. The aim of this work was firstly to investigate the performance of k-t BLAST acquisition and reconstruction when applied to real-time imaging of speech in healthy subject. We were particularly interested to determine if this method could be use clinically to visualize soft palate motion and velopharyngeal closure during speech. We have also investigated the use of k-t FOCUSS algorithm to reconstruct dynamic speech data and compare it with commercially available k-t BLAST reconstruction.

Methods: MRI was performed on a 1.5T Achieva scanner in conjunction with a 16-channel neurovascular coil (Philips Healthcare, Best, the Netherlands). Real-time images were acquired in the mid-sagittal plane covering the head and neck to below the level of the epiglottis. 3 subjects were imaged while completing a speech task consisting of counting (1-10), nonsense words such as " za-na-za" (vowel sound-nasal-vowel) and sustained phonation (/a/, /i/). Audio was simultaneously recorded using a fibre-optic MR microphone system (FOMRI II, Opto-accoustics) and retrospectively synchronised to the images. Balanced SSFP (bSSFP) was acquiered with the following parameters: TE/TR=1.6/3.2ms, 30° flip angle, spatial resolution 1.9×1.9mm^2, single 10mm slice with k-t BLAST acquisition scheme. It consisted of two stages: an acceleration stage in which data are undersampled resulting in aliasing and a training stage in which low resolution estimate of dynamic changes are obtained. Two different dynamic scans modes were used: with training mode interleaved into undersampling stage and with training data acquired in one block acquired before undersampling stage. Each subject performed the same speech task for each undersampling patterns resulting in different frame rates. For interleaved training and undersampling factors of 3, 5, 8, and 16 the frames rates were 7, 10.75, 14.45 and 19.78 fps respectively. For the undersampling factors of 8 and 16, non-interleaved training was also used resulting in frame rates of 23.45 and 46.94fps respectively. In addition, images with full k-space sampling were also acquired at a rate of 2.83fps. Data were then reconstructed using commercial k-t BLAST reconstruction and k-t FOCUSS (used with permission from bisp.kaist.ac.kr [8]). Dynamic images were synchronized with speech samples to allow for assessment of temporal fidelity.

Analysis: The aim was to assess how acceleration scheme affects image quality and if temporal fidelity is preserved. Image quality was evaluated by means of the mean-squared error (MSE) between the fully sampled and undersampled images within the region of interest (ROI) selected over the soft palate. Signal-to-noise (SNR) was also measured in the soft palate. Analysis was performed using MATLAB (Mathworks, Natick, MA). Since visualization of velopharyngeal closure was of special interest, images of the palate in the elevated (while sustain "Ah") and relaxed (nasal breathing) were used for this evaluation. Temporal fidelity in depicting palatal movement during accelerated acquisitions was assessed using intensity-time profiles created from an image intensity profile placed along the primary direction of motion in the soft palate.

Results: *Figure 1* shows calculated MSE for used acceleration factors measured over soft palate ROI in the elevated position in the reference fully-sampled images and those for different acceleration. For modest accelerations no aliasing artefacts were observed and there was a reasonable correspondence between the fully sampled images and both k-t BLAST and k-t FOCUSS. *Table 1* shows measured SNR in the soft palate region. We also verified that for the higher acceleration factors MSE was increased and SNR decreased for both reconstruction methods. Intensity–time profile of the soft palate during counting were very similar for all frame rates used, however there was a discrepancy for the lowest frame rate (2.82 fps) during the nonsense words (later part of the speech sample as illustrated in *Figure 2*). Also images acquired with higher acceleration factors (above 8) showed increased blurring as visible on the lowest profile corresponding to ~19 fps.

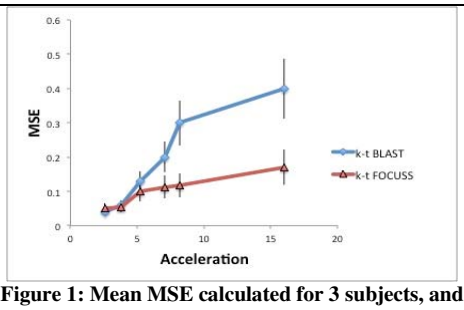

Figure 1: Mean MSE calculated for 3 subjects, and for the frames in elevated position; SD shown in bars.

Table 1 SNR measurement in the soft palate for the different acceleration factors, Mean (sd).

Frame rate (fps)	7	10.75	14.45	19.71	23.45	46.9
k-t BLAST	5.95(2.01)	5.6(2.10)	5.2(1.85)	4.1(1.50)	3.8(1.55)	3.0(1.82)
k-t FOCUSS	6.1(1.81)	6.09(1.95)	5.9(1.91)	5.2(1.49)	4.51(1.45)	3.31(1.63)

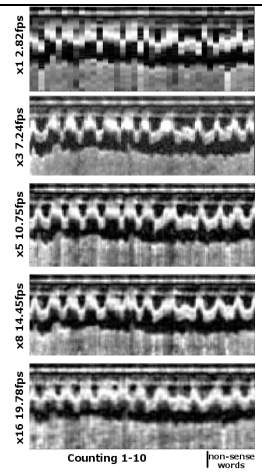

Figure 2. Intensity–time profile of the soft palate.

Discussion: Results suggest that k-t BLAST scheme is a promising tool to accelerate MRI acquisition for speech assessment. K-t FOCUSS reconstruction allowed to improve quality of the reconstructed images. We found that for moderate acceleration factors, resulting in around 15fps, the quality of dynamic images did not deteriorate substantially and temporal depiction of the movement of the soft palate was preserved for the speech samples. However, for higher acceleration factors images revealed increased blurring and SNR dropped. The results also revealed differences in temporal signal variations for the very low frame rate detected for some parts of speech sample suggesting minimal frame rate requirement and should be further investigated. **References:** [1] Rudnil et al *Curr Opin Otolaryngol Head Neck Surg* 2008:16; 530. [2] Anagnostara et al *JMRI* 2001;14;193. [3] Beer et al *JMRI* 2004: 20; 791[4] Narayanan *J.Acounst Soc Am*. 2004 115;1771.[5].Kozerke et al *Magn Reson Med* 2004; 52(1);19-26 [6] Hansen et al *Proc ISMRM* 2006,3187;[7]Jung et al *Magn Reson Med* 2009, 61; 103-116; [8]**Acknowlegements**: Barts Charity and Cleft Charity for financial support; k-t FOCUSS software was kindly provided by H. Jung.