HAL
archives-ouvertes.fr

Speech Cine SSFP with optical microphone synchronization and motion compensated reconstruction

Pierre-André Vuissoz, Freddy Odille, Yves Laprie, Emmanuel Vincent, Gabriela Hossu, Jacques Felblinger

▶ **To cite this version:**

Pierre-André Vuissoz, Freddy Odille, Yves Laprie, Emmanuel Vincent, Gabriela Hossu, et al.. Speech Cine SSFP with optical microphone synchronization and motion compensated reconstruction. ISMRM Workshop on Motion Correction in MRI, Jul 2014, Tromso, Norway. 2014. <hal-00994526>

HAL Id: hal-00994526
https://hal.inria.fr/hal-00994526

Submitted on 6 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Speech cine SSFP with optical microphone synchronization and motion compensated reconstruction

Pierre-André VUISSOZ[1,2], Freddy ODILLE[1,2], Yves LAPRIE[3,4,5], Emmanuel VINCENT[3,4,5], Gabriela HOSSU[2,7], Jacques FELBLINGER[2,6]

[1]IADI, Université de Lorraine, Nancy, France
[2]U947, Inserm, Nancy, France
[3]LORIA, CNRS, Nancy, France
[4]Inria, Nancy, France
[5]LORIA, Université de Lorraine, Nancy, France
[6]CHU de Nancy, Nancy, France
[7]CIC-IT 1433, INSERM, Nancy, France

PURPOSE

Dynamic imaging of the vocal tract is important for modeling speech through the acoustic-articulatory relation.

Previous work:

- X-ray video fluoroscopy : ionizing radiation;
- Real-time MRI [1]: limited SNR and spatial resolution;
- Cine MRI with acoustic device gating [2]: needs highly reproducible motion.

This work:

- Ungated acquisition with acoustic device recording;
- Motion-compensated cine reconstruction.

MATERIAL & METHODS

- **Data acquisition**
  - Ungated balanced SSFP: 1 sagittal slice, 256x256 matrix, TR/TE=3.9/1.7 ms, 5 mm slice thickness, 45° flip angle, 30 cm FOV, 65 temporal phases, approx. 1 min acquisition duration.
  - The subject was asked to repeat a sentence until the sequence stopped. The protocol comprised 10 short sentences providing a good coverage of the tongue movement in French language [3].
  - Acoustic signal recording using an optical microphone (FOMRI III, Optoacoustics, Yehuda Israel). The scanner's acquisition window signal was also recorded by the device for synchronization with MR data.
- **Acoustic signal processing**
  - Denoising of acoustic signals [4] to eliminate gradient noise;
  - Phonetic segmentation to annotate the beginning of each phoneme ;
  - Creation of an acoustic phase signal to indicate the temporal position of each k-space sample in the sentence
- **Cine image reconstruction**
  - Piecewise linear scaling is used to combine all occurrences of a sentence based on the manual segmentation of the acoustic signals.
  - Motion-corrected sliding window reconstruction using cine-GRICS [5]: a sliding window of 50 ms was chosen and a template sentence is reconstructed with 128 frames (temporal resolution <10 ms).

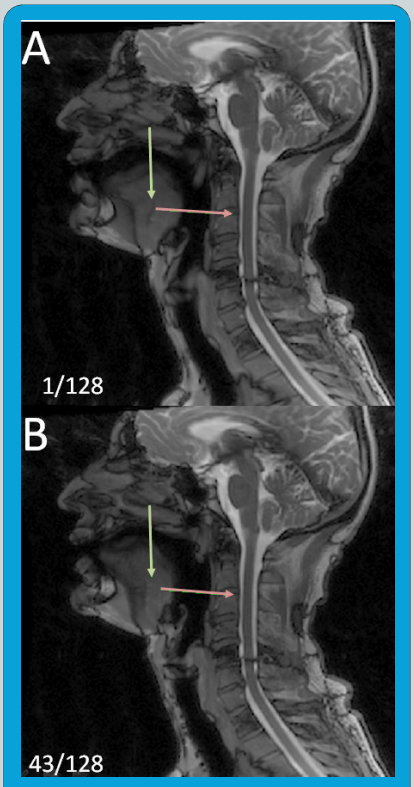

Fig.1 A sagittal slice with two time frames (A) and (B) from a reconstructed template sentence cine loop.

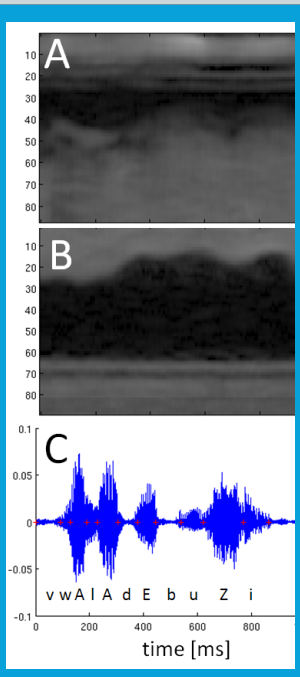

Fig.2 Time-motion display of the cine loop for (A) hard palate to tongue dorsum and (B) tongue back to pharyngeal wall; (C) the corresponding sound signal with manual annotations of phonemes (red crosses).

RESULTS

- Efficient gating and artifact suppression as illustrated in Fig. 1 with two images of the subject pronouncing "Voilà des bougies".
- Characteristic distances can be measured over time such as those between the tongue dorsum and the hard palate or between the tongue back and the pharyngeal wall (see Fig. 2).

DISCUSSION AND CONCLUSION

- A limitation of the balanced SSFP sequence is the possible banding artifacts due to the strong B0 gradient at the air tissue interface.
- Each cine loop enables the delineation of the vocal tract with sufficient spatial and temporal resolution enabling the acquisition of a personalized speech model within an MR examination of half an hour.

REFERENCES [1] Narayanan et al., J Acoust Soc Am, 115(4):1771 (2004) ; [2] Frauenrath et al., Act Acus, 94(1):148 (2008); [3] Maeda, Actes X JEP, p152, Grenoble (1979) ; [4] Ozerov et al., IEEE TASLP, 20(4) :1118 (2012); [5] Vuissoz et al., JMRI, 35 :340 (2012).