33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Cluster Sizes in Interleaved Silent Steady State (ISSS) Imaging

Kwaku Akrofi, *Member, IEEE,* Bradley P. Sutton, *Member, IEEE,* Cheng Ouyang, *Student Member, IEEE,* and Fatima T. Husain

***Abstract*—The effect of increasing the number of scans in the "cluster" of an interleaved silent steady state (ISSS) fMRI imaging scheme from 1 to 2, and then to 3 was examined by a fixed-effects analysis of an auditory short-term memory task with four subjects. Compared to a cluster size of 1, a cluster of 2 scans improved sensitivity at detecting brain activity and statistical power, while a cluster of 3 scans further improved statistical power but seemed not to improve sensitivity beyond that achieved with a cluster of 2 scans. The findings reveal that cluster size is a vital parameter for an ISSS imaging scheme.**

I. Introduction

FUNCTIONAL *magnetic resonance imaging* (fMRI) has been widely used in studies of the human auditory system. Acoustic MRI scanner noise (about 120 dBA) generated by the fast switching gradients of echo-planar imaging (EPI) - the predominant fMRI imaging method - is, however, a serious problem in auditory studies for several reasons. Due to perceived changes in loudness and intelligibility, experimental auditory stimuli may be difficult or nearly impossible for subjects to hear [1], [2]. Even if a subject hears the stimuli, these perceived changes may alter the pattern of brain activation [1]. One reason for this is that the noise generates responses in the central auditory pathway that reduce the dynamic range of the responses to the stimuli [3]. Blood oxygenation level dependent (BOLD) responses to the noise would result in further brain activation changes [4]. When imaging patients with hearing loss or *tinnitus* (ringing in the ears), scanner noise can result in unpredictable effects that are another cause for concern. For example, scanner noise may mask the internal noise of some tinnitus patients and may not mask the internal noise of other tinnitus patients, leading to different brain activation patterns within the same patient population.

Beyond the use of ear protection and active noise cancellation (ANC), several temporal acquisition or "sampling" schemes have been designed to reduce the effects of scanner noise on experimental stimuli. The simplest and most widely used is ***sparse sampling*** (SS) [1]–[5], which inserts a long delay (usually about 8-10 s) between each volume image acquisition. The delay provides a period of (relative) silence, as no switching gradients are active for its entire duration. The stimulus is presented during the delay, and its onset is placed at a point in time that would allow the BOLD response for the stimulus to start during the delay and peak at the time of the image acquisition. While SS has proved to be more robust in detecting auditory brain activity than conventional "continuous" sampling [3] (and hence its widespread use), it does result in a considerable loss of statistical power because fewer images are acquired in a given period of time due to the delay between image acquisitions.

Clustered sampling (CS) [2], [6] gains more statistical power, relative to SS, by acquiring a "cluster" of images, instead of one, between delays. Each image in the cluster, however, has different T1 saturation levels, complicating statistical analysis of the statistical fMRI analysis.

Interleaved silent steady state (ISSS) sampling [1], [4] acquires image volumes every repetition time (TR), but unlike continuous sampling, certain imaging gradients are disabled for some TRs (dummy scans) and enabled for others (true scans), thus creating a relatively silent period and a cluster of images just like CS. Unlike CS, the T1 saturation levels in all images in all clusters are the same in ISSS. The slice selection gradient and RF pulses are played out every TR in order to maintain T1 steady state. Sinusoidal ramps may be used in place of linear ones in the slice-selection gradients [1], phase-encoding gradients [7] or readout echo trains [7] to make the scanner less noisy during true scan acquisition in ISSS schemes.

Several studies, e.g. [1]–[3], [7] have compared the ISSS and/or CS to SS and/or continuous sampling for statistical power and noise reduction. However, the relationship between cluster size and statistical power has yet to be studied. In this paper, we examine that relationship by comparing the extent and details of the activation maps generated for an auditory task when cluster size is increased from 1 to 3.

II. Procedures

A. Subjects

Four subjects participated in the study. Two were male,

Manuscript received April 15, 2011. This work was supported in part by the Campus Research Board of the University of Illinois at Urbana-Champaign (UIUC) and the Tinnitus Research Consortium.

K. Akrofi is with the Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign (UIUC), Champaign, IL 61820 USA (Phone: 217-333-2230; fax: 217-244-2235 e-mail: kakrofi@illinois.edu).

B. P. Sutton is with the Department of Bioengineering, UIUC, Urbana, IL 61802 USA (e-mail: bsutton@illinois.edu).

C. Ouyang is with the Department of Bioengineering, UIUC, Urbana, IL 61802 USA (e-mail: couyang2@illinois.edu).

F. T. Husain is with the Department of Speech and Hearing Science, UIUC, Champaign, IL 61820 USA (e-mail: husainf@illinois.edu).

two were female, all were right-handed and had normal hearing. Their ages were 22, 23, 29 and 32. All subjects volunteered to participate in this study and gave written informed consent. Imaging data were collected at the Biomedical Imaging Center at the Beckman Institute of the University of Illinois at Urbana-Champaign.

B. fMRI Details

All volume images were acquired on a Siemens Allegra 3 T head scanner, with the following imaging parameters: slice thickness, 4 mm; inter-slice gap, 0.4 mm; number of slices, 32; slice orientation, axial oblique; field of view (FoV) read, 220 mm; FoV phase, 220 mm; TR, 2000 ms (ISSS), 12000 ms (SS); TE, 25 ms (ISSS), 30 ms (SS); matrix size (per slice), 64 × 64; flip angle, 90°; bandwidth, 2894 Hz/Pixel.

C. Sampling Schemes

Three sampling schemes were examined in this study: sparse sampling (SS), ISSS with two true scans per cluster (ISSS2) and ISSS with three true scans per cycle (ISSS3). Note that SS is essentially the same as an ISSS scheme with one true scan per cycle (ISSS1). Fig. 1 shows the timing details of these sampling schemes.

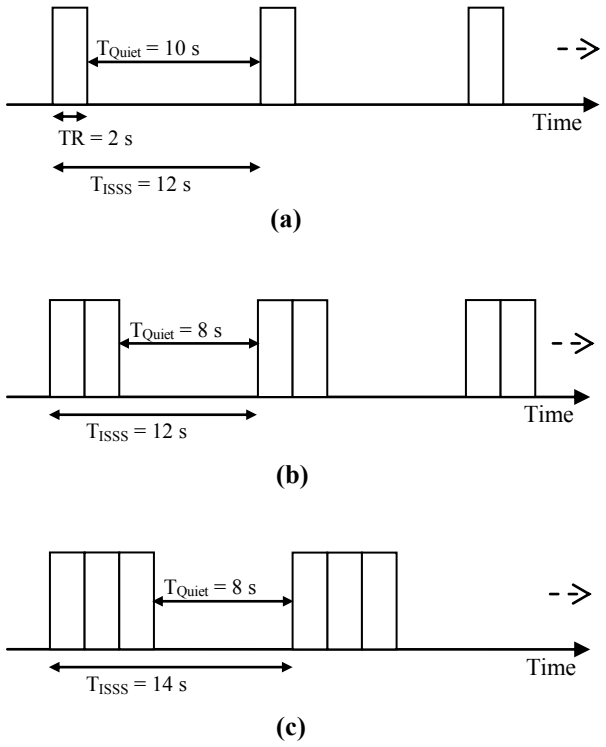

Fig. 1. Sampling schemes: **(a)** Sparse sampling, **(b)** ISSS2, and **(c)** ISSS3. TR = 2 s for all three schemes. Although (a) is not an ISSS scheme, T_{ISSS} describes the duration of a cycle. In (b) and (c), the T_{Quiet} period is not a time delay, but actually 4 dummy scans.

For ISSS3, T_{ISSS} - the duration of a cycle of true and dummy scans - was set to 14 s so that T_{Quiet} - the time taken to acquire successive dummy scans - would be 8 s. A shorter T_{Quiet} would have been too short to ensure that BOLD response of the subject would peak while the true scans were being acquired.

To ensure quiet scanner operation during ISSS dummy scan acquisition, the following EPI gradients were disabled throughout T_{Quiet}: phase preparation, negative readout, positive readout, and phase blip gradients.

D. Stimuli

For each sampling scheme, subjects were asked to perform two types of auditory short-term memory tasks: ***low attention*** (**LO**) and ***high attention*** (**HI**). In a LO task, a pure tone is played for 0.5 s and, 1 s later, another one is played for 0.5 s; the subject responds SAME if the tones are identical and DIFFERENT if they differ in pitch (frequency). In a HI task, there is a 0.5 s pure tone followed by a 0.5 s time delay, then a second 0.5 s pure tone, then a 1 s delay, and then a third 0.5 s pure tone. The first and second tones *never* have the same frequency. If the third tone has the same frequency as either of the previous two, the response is SAME; otherwise the response is DIFFERENT.

22 pure tones, ranging from 500 Hz to 1 kHz in frequency, were shuffled and used for the LO and HI tasks. In all, for each sampling scheme, 70 tasks in total were performed: 30 were LO, 30 were HI and 10 were rest periods with no sound stimuli.

All tones were played through Resonance Technology MRI-compliant headphones (Model RTC2K), by a computer running Presentation 14.7 (www.neurobs.com) in a Windows XP environment. Two sets of response buttons were strapped to each hand of a subject; only index finger buttons were required for this study. Subjects were instructed to use their right index finger button to indicate a DIFFERENT response and their left index finger button to respond SAME.

E. Statistical fMRI Analysis

SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) was used to analyze the fMRI data.

For a given subject and sampling scheme, preprocessing was as follows. Scans were realigned to the mean scan, which is itself computed during the realignment process. For ISSS data, only true scans were realigned. Each dummy scan was replaced with a copy of the mean scan. A two-step coregistration process was used: a low-resolution anatomical T2 image was coregistered to the mean scan, followed by coregistration of a high-resolution MPRAGE T1 image to the T2 image. The realigned scans – for ISSS, both realigned true scans and copies of mean scan – were then normalized to a standard MNI T1 template with the MPRAGE image as the source image. Normalized images were then smoothed at 8 mm FWHM. After this preprocessing procedure, the general linear model (GLM) was specified as follows.

The design matrix had two conditions for each subject: LO and HI. To correct for the effects of head motion, 6 regressors generated by the realignment process were also

included in the design matrix. For ISSS, a dummy regressor, which had a value of 1 for dummy scans (i.e. mean scan copies) and 0 for true scans, was included. Lastly a baseline regressor, a constant of 1, was automatically added by SPM8. The choice of basis function was the canonical HRF with no derivatives for SS and finite impulse response (FIR) set for ISSS. Table 1 lists the specifications of the FIR sets for ISSS2 and ISSS3. The order was set to the number of true scans per cycle while window length was equal to $T_{ISSS} - T_{Quiet}$. The coefficients of the FIR set bins were samples of the canonical HRF and the 2nd, 4th, and (for ISSS3) 6th seconds. For each sampling scheme, a fixed-effects statistical analysis of all subjects was conducted to determine significant brain response for the LO condition compared to the HI condition.

TABLE I
FIR BASIS SET DETAILS

FIR Set Characteristics	ISSS2	ISSS3
Order	2	3
Acquisition Window Length	4 s	6 s
Bin Coefficients	0.0115	0.0115
	0.0203	0.0203
		0.0154

III. Results

The parameters of the design matrices described in Section IIE were estimated with SPM8 to obtain general linear models (GLMs). The activations were computed for a *t*-contrast to indicate regions where the activation for the HI task was greater than the activation for the LO task. Figs. 2, 3, and 4 show the results of the fixed-effects analysis for SS, ISSS2, and ISSS3 respectively. In all three figures, $p < 0.005$ (uncorrected). Table 2 lists the degrees of freedom (*df*) in Figs, 2, 3 and 4.

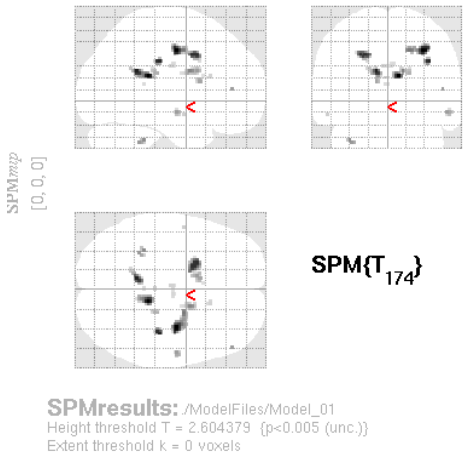

Fig. 2. Statistical parametric map (SPM) for the contrast HI>LO for the SS sampling scheme depicted on a glass brain.

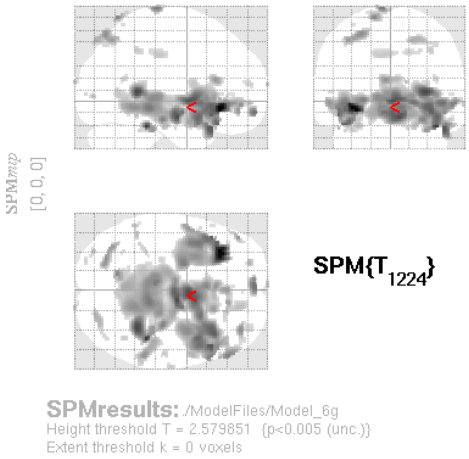

Fig. 3. SPM for the contrast HI>LO for the ISSS2 sampling scheme depicted on a glass brain.

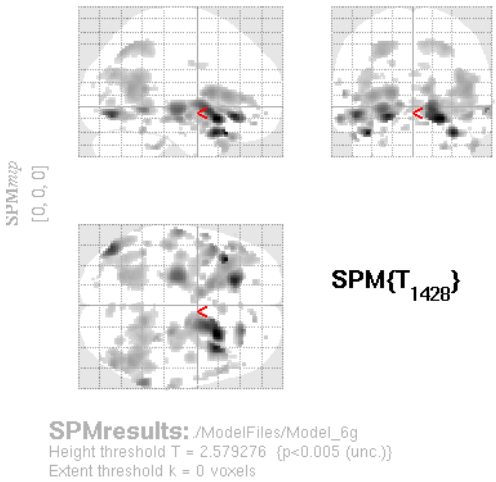

Fig. 4. SPM for the contrast HI>LO for the ISSS2 sampling scheme depicted on a glass brain.

TABLE 2
DEGREES OF FREEDOM

Sampling Scheme	SS	ISSS2	ISSS3
Degrees of Freedom (df)	174	1224	1428

IV. Discussion

Figs. 2-4 show that ISSS2 and ISSS3 produced more extensive activation patterns than SS (which is essentially ISSS1), potentially demonstrating that there is a more extensive functional brain network involved in processing the auditory attention tasks than can be inferred from SS alone. ISSS2 and ISSS3, however, appear to have yielded roughly the same amount of activation. The pattern produced by ISSS3, from Fig. 4, appears more "focused" or "sharper". We are yet to determine the functional implications, if any, of this sharpness.

Table 2 shows a stronger correlation between *df* and

cluster size. The GLM model for ISSS3 has a larger *df* than that of ISSS2, which in turn has a larger *df* than SS. *df* is the number of independent parameters in the model. The higher the *df*, the more the statistical power. Consequently, a more complex analysis (e.g. a model with 6 conditions) might yield an estimable model with ISSS3, but not with SS.

The *df* of ISSS2 and ISSS3 are both an order of magnitude larger than that of SS. This is misleading. *df* is directly proportional to the total number of scans, N. ISSS2 and ISSS3 have more dummy scans, D, than true scans, M. SS has no dummy scans. Only true scans contribute meaningful information to fMRI time series. Since $M_{ISSS3} = 3M_{SS}$ and $M_{ISSS2} = 2M_{SS}$, in actual fact, $(df_{ISSS3} \approx 3df_{SS}) > (df_{ISSS2} \approx 2df_{SS}) > df_{SS}$. While a cluster size of 2 or more does increase *df*, the increase is not 10-fold as suggested by Table 2. Fig. 5 shows the relationship between cluster size and actual *df* of our data.

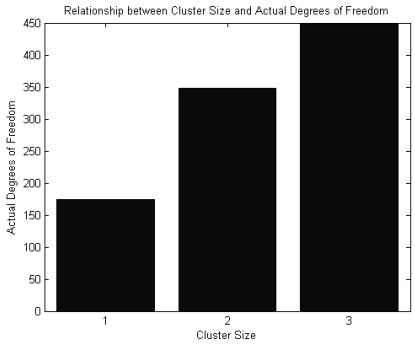

Fig.5 The relationship between actual cluster size and *df*. Doubling cluster size does not increase *df* by an order of magnitude.

It should be noted that although $df_{ISSS3} \approx (3/2)df_{ISSS2} \approx 3df_{SS}$, the statistical power of ISSS3 was not 3/2 times that of ISSS2 in our study, but $(3/2 \times 12/14) = 1.29$ times instead. This is because, as shown in Fig. 1b-c, T_{ISSS} was 14 s for ISSS3 but 12 s for ISSS2. Increasing T_{ISSS} decreases the statistical power of an experiment. Therefore, cluster size increases may not always yield the expected increases in statistical power, as it is sometimes not feasible to increase cluster size without increasing T_{ISSS}.

In general, the results in this study are in agreement with those from previous studies such as [2] and [4]. In [2], CS was found to have better statistical power, while [4] found ISSS to be more sensitive to detecting brain activity resulting from auditory experiments. Both recommended ISSS over SS.

ISSS does have its drawbacks. MRI scanners are currently not equipped with standard ISSS and CS pulse sequences. It takes considerable time, programming effort and skill, and a strong MR background to program an ISSS sequence. In contrast, an SS sequence simply requires adjustment of scanner parameters.

V. Conclusion

The study examined the effect of increasing the number of volumes obtained within clustered ISSS sequences from 1 to 2, and then to 3. (A sparse sampling scheme was used in place of an ISSS scheme with a cluster size of 1; the two are essentially the same.) Our preliminary results showed that compared to SS, ISSS2 and ISSS3 exhibited more sensitivity to detecting brain activations and increased statistical power. Compared to ISSS2, ISSS3 showed roughly the same sensitivity to brain activation detection but increased statistical power. Overall, ISSS improves the efficiency and sensitivity of auditory fMRI studies, but requires considerable MR scanner programming effort and could demand extra scanning time.

We intend to further examine the effect of increasing number of sparse sampling clusters (for instance, 5). We plan to verify the efficiency and sensitivity of the different fMRI sequences using a larger set of subjects that would include patients with tinnitus and/or hearing loss as well as normal hearing controls. Our study underscores the interaction of scanner noise and brain response and the need to account for it in experimental investigations using auditory stimuli or in patients groups affected by the noise.

Acknowledgment

We would like to express our profound gratitude to Nancy Dodge and Holly Tracy-Hearnley for their help and support with fMRI acquisition and testing of the ISSS pulse sequences, and to Matt Davis for his invaluable help with the ISSS GLM data analysis.

References

[1] S. Schwarzbauer , M. Davis, J. M. Rodd, and I. Johnsrude, "Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI," *NeuroImage*, vol. 29, pp. 774–782, 2006.

[2] T. Zaehle, C. F. Schmidt, M. Meyer, S. Baumann *et al*., "Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks," *NeuroImage*, vol. 37, pp. 1195–1204, 2007.

[3] G. A. Blackman and D. A. Hall, "Reducing the effects of background noise during auditory functional magnetic resonance imaging of speech processing: Qualitative and quantitative comparisons between two image acquisition schemes and noise cancellation," *J. Speech Lang. Hear. Res.*, vol. 5,4 iss. 2, pp. 693–704, April 2011.

[4] K. Mueller, T. Mildner, T. Fritz, J. Lepsien et al., "Investigating brain response to music: A comparison of different fMRI acquisition schemes," *NeuroImage*, vol. 54, pp. 337–343, 2011.

[5] A. A. Stevens and K. E. Weaver, "Functional characteristics of auditory cortex in the blind," *Behav. Brain Res.*, vol. 196, iss. 1, pp. 134–138, Jan. 2009.

[6] D. A. Hall and A. E. Paltoglou, "fMRI of the Central Auditory System," in *fMRI Techniques and Protocols*, Neuromethods, vol. 41, M. Filippi, Ed., Humana Press, 2009, pp. 537–569.

[7] S. Schmitter, E. Diesch, M. Amann, A. Kroll et al., "Silent echo-planar imaging for auditory FMRI," *Magn. Reson. Mater. Phy.*, vol. 21, pp. 317–325, 2008.