Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Choreography Controlled (ChoCo) brain MRI artifact generation for labeled motion-corrupted datasets

Oscar Dabrowski^{a,*}, Sébastien Courvoisier^{b,c}, Jean-Luc Falcone^a, Antoine Klauser^{b,c}, Julien Songeon^{b,c}, Michel Kocher^d, Bastien Chopard^a, François Lazeyras^{b,c}

^a Computer Science Department, Faculty of Sciences, University of Geneva, Switzerland

^b Department of Radiology and Medical Informatics, Faculty of Medicine, University of Geneva, Switzerland

^c CIBM Center for Biomedical Imaging, MRI HUG-UNIGE, Geneva, Switzerland

^d Biomedical Imaging Group (BIG), School of Engineering, EPFL, Lausanne, Switzerland

ARTICLE INFO

Keywords: MRI Motion-tracking Motion calibration Motion artifact Motion estimation In-vivo EPI Spin-Echo

ABSTRACT

MRI is a non-invasive medical imaging modality that is sensitive to patient motion, which constitutes a major limitation in most clinical applications. Solutions may arise from the reduction of acquisition times or from motion-correction techniques, either prospective or retrospective. Benchmarking the latter methods requires labeled motion-corrupted datasets, which are uncommon. Up to our best knowledge, no protocol for generating labeled datasets of MRI images corrupted by controlled motion has yet been proposed. Hence, we present a methodology allowing the acquisition of reproducible motion-corrupted MRI images as well as validation of the system's performance by motion estimation through rigid-body volume registration of fast 3D echo-planar imaging (EPI) time series. A proof-of-concept is presented, to show how the protocol can be implemented to provide qualitative and quantitative results. An MRI-compatible video system displays a moving target that volunteers equipped with customized plastic glasses must follow to perform predefined head choreographies. Motion estimation using rigid-body EPI time series registration demonstrated that head position can be accurately determined (with an average standard deviation of about 0.39 degrees). A spatiotemporal upsampling and interpolation method to cope with fast motion is also proposed in order to improve motion estimation. The proposed protocol is versatile and straightforward. It is compatible with all MRI systems and may provide insights on the origins of specific motion artifacts. The MRI and artificial intelligence research communities could benefit from this work to build in-vivo labeled datasets of motion-corrupted MRI images suitable for training/testing any retrospective motion correction or machine learning algorithm.

1. Introduction

MRI is an essential imaging modality in medicine, which suffers unfortunately from a great sensitivity to movements that can deteriorate image quality. In general, relatively long acquisition times (in the order of minutes) are required to obtain high resolution structural MRI images. Hence, there are situations where motion during scan is unavoidable.

Many researchers have devoted their efforts towards motion impact mitigation by reducing acquisition times or by designing new motion insensitive a.k.a. "motion robust" pulse sequences as described in [1]. The latter can leverage non-Cartesian k-space sampling strategies (e.g., spirals, radial), such as the PROPELLER [2] or SNAILS [3] methods, which usually over-sample the center of k-space.

Several methods have been proposed for motion-tracking that use either external devices or navigators [4]. Widespread systems are based on cameras coupled with special markers [5,6]. One main drawback of such methods is the need for calibration, which can be quite challenging [7,8] especially with "out-of-bore" camera systems [9]. An alternative to visual systems is the use of electro-magnetic (EM) pickup coils [10,11]. These methods require modifications of the pulse sequence to generate gradient pulses needed for the calibration of the device in the reference frame of the MRI system [12]. Navigator echoes offer an attractive means for either prospective or retrospective motion correction [13,14], as they use the MRI intrinsic coordinate system. In particular, the FIDNav technique can be used, which leverages multi-coil antennae to obtain motion information [15]. Much like EM tracking, these navigator-modified sequences increase scan duration and the method may not be compatible with every type of pulse sequence.

https://doi.org/10.1016/j.ejmp.2022.09.005

Received 14 March 2022; Received in revised form 26 August 2022; Accepted 12 September 2022 Available online 19 September 2022



Original Paper



^{*} Correspondence to: University of Geneva, route de Drize 7, 1227 Carouge, Switzerland. *E-mail address:* oscar.dabrowski@unige.ch (O. Dabrowski).

^{1120-1797/© 2022} Associazione Italiana di Fisica Medica e Sanitaria. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).



Fig. 1. Experimental setup. The subject wears plastic glasses with a sticker attached in front of their pupil, restricting the green dot tracking to head motion only and avoiding any eye tracking. During a controlled motion experiment, the subject is told to keep the green dot in the center of the pinhole by moving the head. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Regardless of the chosen method, whether calibration is done by choreographed head movements or by simulation, there is a need to obtain controlled head motion acquisitions together with reproducibility validation for motion correction quality assessment [16]. However, to the best of our knowledge, no dedicated system is available for the investigation of controlled motion impact on MRI scans in a straightforward, inexpensive and versatile (i.e., compatible with any scanner or pulse sequence, provided that a screen is available to display the visual stimulus) manner.

Motion correction is a complex problem, which depends on many parameters such as pulse sequence, type of motion (where it occurs in k-space, abrupt vs progressive motion, etc.) and also on the motion detection/estimation methods. Therefore, it is often difficult to know the exact link between an observed artifact and a given motion [17]. So far, researchers in the field have mostly focused on "free motion" and there is a need for labeled motion-controlled datasets acquired within ranges that are most likely to occur during clinical scans. This kind of labeled data should allow the evaluation of motion correction techniques in a more tractable and systematic manner.

We propose a general and straightforward method allowing reproducible head choreographies to be performed by subjects inside any MRI scanner equipped with a video display system. The choreographic control is based on the visual pursuit of a target solely by the subject's head displacement. The performance of the controlled motion system is validated by estimating motion during a fast time series acquisition. In practice, each subject undergoes a target-tracking training session during a fast echo planar imaging (SMS-EPI) [18,19] acquisition, from which subject motion is estimated by rigid-body volume registration. Then, the same type of motion is performed during standard structural brain imaging in order to create corrupted images based on controlled head motion. Any sequence can be used to produce corrupted images with known motion patterns, provided that the choreography control video is synchronized with the start of the sequence and the availability of an MRI-compatible screen (as in a standard fMRI setup).

Previous work has shown that abrupt head motion can be as fast as 200 ms [20] and EPI repetition times are usually of the order of the second. In order to improve the accuracy of motion estimation of EPI time series, especially with rapid motion occurring within EPI repetition time (TR), a spatio-temporal upsampling and interpolation method is proposed. It leverages the interleaved simultaneous multislice (SMS) scheme used in conjunction with EPI. Methods leveraging the SMS scheme have been proposed, such as [18], but the usage of Kalman filtering may hinder detection of fast abrupt motion and no implementation is freely available. We provide a Matlab implementation for our method at: https://gitlab.unige.ch/Oscar.Dabrowski/te mporally-improved-volume-registration-by-data-driven-subvolume-inte rpolation-of-simultaneous-multislice-epi-time-series. Our motion-control method enables the creation of labeled datasets of motion-corrupted MRI images that could be used for motion correction algorithm benchmarking or serving as the basis for motion artifact simulation, for a better understanding of the phenomenon. The protocol is designed to allow reproducible motion artifacts within a known range of parameters and following a precise timing. Among many applications, the dataset obtained following this procedure may be used as a test set to assess the performance of any retrospective motion correction technique.

2. Methods

2.1. System setup

Motion instructions were provided to the subjects via an LCD screen placed at the back of the MRI and viewed by a mirror mounted on the standard head coil. Subjects wore a modified pair of plastic glasses to restrict their field-of-view (Fig. 1). The idea of motion tracking with this equipment is to mimic the same principle as following a moving target with binoculars or a spyglass. One eye is masked, and for the other, a pinhole forces an eye-pinhole-marker alignment. It has been shown that head and eye motion both contribute to tracking during a visual pursuit [19]. Using glasses with a sight (masking crosshair marker), naturally constrains motion to the head only, removing any eye tracking contribution. This simple device alleviates the experiment from any variation in instruction comprehension or implementation such as subject ability to not move the eyes while tracking.

To generate a suitable target pursuit movie for the head choreography, the distance in pixels x_{px} on the screen is expressed as a function of the head angle θ with respect to the rest position as shown in Fig. 1:

$$x_{px}(\theta) = D \cdot \frac{w_{px}}{w_{cm}} \cdot \tan(\theta)$$
(1)

where D = d + d' is the distance between the subject's eye and screen, w_{cm} is the screen width in centimeters and w_{px} is the horizontal resolution in pixels.

The impact of the distance measurement inaccuracy is clearly negligible compared to the subject's angle reproduction precision (human error), considering that a ± 1 cm distance error translates to an error of approximately ± 0.0065 degrees.

As a proof of concept of the system, a choice had to be made with respect to the motion parameters and pulse sequences to use. The widely used 2D (T1-w-SE) and 3D (T1-w MP-RAGE) sequences were chosen as well as EPI for motion estimation. Two motion types were chosen: "abrupt" (fast) and "motion and hold" (slower), with 3 different angles $\theta \in \{1^{\circ}, 2^{\circ}, 3^{\circ}\}$ representative of the range of motion which may occur during a clinical scan.

Due to the restricted field of view, all relevant information regarding head choreographies was displayed in the central region delineated



Fig. 2. (a) Three consecutive groups of 4 SMS slices are combined in to a partial volume acquired each third of a TR (b) spatially interpolated into a full volume (c) suitable for rigid-body registration (w.r.t. first full volume) with SPM.

by the pinhole (Fig. 1). Head choreography movies were designed with specific timings corresponding to fast *abrupt* and slower *motion and hold* motion types in up/down and right/left directions. A set of symbols shown in Fig. 1 informs the subjects on which movements to perform. A countdown timer is displayed during rest periods and two seconds before motion onset, a crosshair appears on the targeted angle to avoid a "surprise" effect, which could potentially lead to "overshooting". To obtain accurate timings for the occurrence of motion artifacts, the video must start synchronized with the beginning of the acquisition sequences. The MRI scanner device-specific sync setup used here is described in Appendix D.

2.2. Subject preparation

Before sending volunteers inside the MRI scanner, their interpupillary distance is checked to match with the glasses pinhole (adjustments were rarely needed). Then, subjects are placed on the MRI table and equipped with a mirror-mounted head coil. Glasses vertical position is adjusted on subject's faces and the glasses are taped on their forehead and temples. Finally, subjects are sent to the isocenter, and final left/right adjustments of the screen is performed following their feedback.

2.3. Experimental protocol

All experiments were performed at University Hospitals of Geneva on a 3T MRI system (MAGNETOM Prisma fit, Siemens Healthineers, Erlangen, Germany) with a 20-channel receiver head coil. Six volunteers participated in the study.

After the acquisition of no-motion T1-w Spin-Echo and T1-w 3D MP-RAGE sequences for reference, the protocol consists of an EPI (training) sequence used for motion estimation. Spin-Echo and T1 3D MP-RAGE sequences were then acquired for each type of motion.

The parameters of the SMS-EPI time series were TR = 600 ms, TE = 30 ms at a resolution of 64×64 pixels per slice, 36 axial slices per volume with a voxel size of $3.0 \times 3.0 \times 3.3$ mm, number of repeated acquisitions = 645. Other parameters were a $2 \times$ GRAPPA acceleration factor with 4 simultaneous slice excitations and TA = 6 m 27 s.

Spin-Echo sequence parameters consisted of axial slices with a resolution of 256 \times 256 pixels, TR = 750 ms (externally triggered with an acquisition window of 800 ms), TE = 12 ms and TA = 3 m 1 s. Voxel size was 0.9 \times 0.9 \times 4.0 mm.

T1 3D MP-RAGE sequences were acquired with sagittal orientation with a resolution of $288 \times 288 \times 208$ pixels, TR = 1930 ms (externally triggered with an acquisition window of 1950 ms), TE = 2.36 ms, TA = 3 m 36 and IT = 972 ms. Voxel size was: 0.9 mm isotropic.

2.4. Sub-TR volume-to-volume registration for accurate abrupt motion estimation

Two different situations may arise for practitioners using the proposed protocol. Either the motion is slow or it is fast with respect to TR period. In the latter case, motion will be undersampled, therefore misestimated. The phenomenon can be understood in a similar way – although not strictly identical – than the Nyquist–Shannon theorem that requires a sampling period (TR) being two times shorter than the fastest motion period to be estimated. In our experiments TR = 600 ms, and fast head motion "period" was estimated to be around 800 ms, hence a TR = 400 ms would be needed at least. In the current application, we performed a threefold increase in temporal upsampling, which would approximately correspond to TR = 200 ms.

The goal of EPI acquisitions is to estimate head motion in relation with the instructions displayed on the screen. As just mentioned, rapid motion may be too fast to be properly sampled by EPI with TR = 600 ms, which leads to its underestimation. In order to have more accurate motion estimation of rapid movements, the SMS scheme can be leveraged to perform temporal upsampling of the EPI acquisitions, which, in turn, leads to spatial undersampling which can be dealt with interpolation.

The method is based on splitting entire EPI acquisitions into multiple time series (TS) according to the slice ordering of the SMS acquisition. For the proof of concept presented here, 36 slices/volume were acquired during each TR, with an SMS factor of 4. We separated the time series into 3 TS (TS1, TS2, TS3) corresponding to 12 slices per volume. This choice represents a good compromise between spatial and temporal resolution. Since incomplete spatial coverage impairs realignment algorithms, each subvolume was spatially interpolated. The proposed interpolation algorithm computes optimal affine pairwise transforms between adjacent slices, i.e., from slice n to slice n + 1 and from slice n to slice n - 1, referred to as "forward" and "backward", respectively. The interpolation is obtained by weighted averages of forward and backward transformed slices in a given subvolume according to hand-crafted weight matrices constructed for each time series. The latter matrices specify the location of existing slices and the weights to use. A qualitative example of interpolation is provided in Fig. 2. Each time series was processed separately. They were each realigned with respect to their first - fully interpolated - corresponding volume with SPM [21]. The latter uses an "intensity-based" registration algorithm, i.e., based on a MSE metric between volumes.

A fully upsampled (×3) time series can be reconstructed by temporally interleaving the motion parameter vectors. A zoom on two motion blocks shows a comparison of the realignment results between the original (non-upsampled, non-interpolated) acquisitions and the same acquisitions processed with our method (Fig. 3). Rigid-body volume registration was performed with respect to a temporally close volume to avoid baseline offset between time series that results in apparent oscillation of the timeline as observed in Fig. 3(a). In practice motion estimation is performed by considering each time series independently and is not affected by the choice of the reference volume.



Fig. 3. Zoom on the two last motion blocks of an EPI training session. (a) Angles estimated with respect to each local baseline (TS1,2,3-b). Maximum angle amplitude is taken for abrupt motion and maximum average for motion&hold. (b) Temporally interleaved 200 ms upsampled signal from (a) (realigned w.r.t. volume#525) superimposed on volume-to-volume, non-upsampled, 600 ms, time series. (1D motion parameter signal (blue) is linearly interpolated to allow visual comparison with upsampled signal (red).) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. Data postprocessing

The interpolation and upsampling method described previously was applied to each subject's EPI data and motion parameters were obtained for each time series separately.

Motion parameters were computed considering local baselines, i.e., resting positions before each motion onset (Fig. 3). Ideally, the baselines should be at zero degrees, but unavoidable variations do occur due to deviations in the spatial content of the slices, when registration is performed with respect to the initial volume (instead of a temporally closer volume, as mentioned in Section 2.4). For the "abrupt" motion type, the maximum peak value among each time series was considered. For the "motion and hold" type, the maximum average value of the "plateau" points, among each time series, was taken.

Descriptive statistics are presented in the form of boxplots and curves displaying signed errors between estimated angles with respect to target values. Boxplots aggregate all subjects' data together to describe overall error whereas curves describe the temporal error-wise behavior of each subject. Both boxplots and curves take baselines into account for error computation.

Representative artefacted Spin-Echo were obtained from raw multicoil k-space combination using the ESPIRiT algorithm [22]. Representative MP-RAGE DICOM images are also provided to allow qualitative estimation of motion artifacts.

2.6. Dataset labeling

The screen defines a 2D projection of the corresponding 6 DoF that subjects are performing. Motion estimation obtained by rigid-body volume registration of EPI training sessions gives empirical labels for acquisitions performed thereafter. Data labeling can be done by extracting the corresponding average motion parameters from the training acquisition once subjects are trained, (after 2.5 min, see results). Hence, in the context of training a supervised deep learning model [23], it is possible to build a labeled dataset { $\mathbf{x}^{(i)}, \mathbf{y}^{(i)}$ }, where *i* is the dataset element index, and where each MRI image $\mathbf{x}^{(i)}$ is associated with its empirical target $\mathbf{y}^{(i)} \in \mathbb{R}^{n \times 6}$, a matrix of 6 motion parameters for each of the *n* phase encoding lines (e.g., in the case of a classical Spin-Echo acquisition).

3. Results and discussion

3.1. Choreography training sessions and motion estimation

For each participant, the experiment begins with a target tracking training session during repeated SMS EPI acquisitions that allows the investigation of subjects' spatio-temporal motion accuracy using rigidbody registration, as well as their learning behavior. The EPI training sessions also provides the labels for the dataset of corrupted MRI images. As described in the last methods subsection, the stimulus parameter space (2D) maps to the spatial parameter space of a subject's head (6-DoF), that corresponds to the "response function" for each subject. This includes the theoretical 6 DoF motion as well as any experimental fluctuations (e.g., internal constraints such as subject neck/back pain as well as external parameters such as excessively inflated air cushion padding, etc.). In this work, we rely on intra-session repeatability to ensure consistent motion control that can be applied reliably in subsequent structural MRI acquisitions. Thus, the training session provides the real motion labels for the data, which are more accurate than "theoretical" target motion labels. Consequently, our approach aims at accurately estimating the motion even if the subject did not perform the motion accurately (as long as the subject performance is consistent within the session). The EPI training session is therefore required for each subject and each session not only to validate motion reproduction but also to obtain the empirically labeled dataset we aim at.

Fig. 4 shows rotational motion estimations for a given subject during a training session after the corresponding upsampled and interpolated EPI acquisition has been realigned with SPM. The first row represents the targeted movements that should be performed by the subject.

Qualitatively, we observe that the subjects motion patterns match the targets well. The "Up/Down" motion angle amplitudes are generally smaller than the "Right/Left" ones (Figs. 4 and 5). According to subject's feedback, this can be explained by the "Up/Down" motion being more difficult to perform accurately.

The other two rows show the effective movements made by the subject as estimated by rigid-body EPI volume registration. Volumes acquired during each EPI sequence are realigned with respect to the first volume to obtain rigid-body motion parameters allowing subject motion estimation. It can be observed that the volunteer (S2) "overshoots" the target angles at the beginning of the acquisition but stabilizes relatively fast, i.e., after about the first two motion blocks ("motion and hold" and "abrupt"). To prevent potential overshooting, a countdown timer is displayed 2 s before each motion starts and a crosshair appears on the targeted angle. This "anticipation timing performance" process proved to improve time response and motion accuracy [24].

3.2. Intra-subject motion repeatability and baseline estimation

The starting position is by convention the zero-degree position. It can be observed that the baseline fluctuates over time because subjects



Fig. 4. Displayed target rotation (top) and estimated performed angle for right/left (middle) and for up/down (bottom) motions. Motion estimation is performed on 3 times upsampled and interpolated multi-volume EPI. Baseline variation for "motion&hold" and "abrupt" motion types are highlighted in frames (a) and (b), respectively.



Fig. 5. Signed error in degrees, i.e., the difference between measured angles and targets, for each type of motion. Movements in the Right/Left direction exhibit less error than in the Up/Down direction.

need to get accustomed to the expected dot position and speed. After about one trial of each motion type, subjects tend to reach a more stable baseline. Hence, they undergo a *training* phase. Baseline variation factors may include slight position readjustments during the experiment (e.g., if the subject felt uncomfortable or swallowing, coughing, etc.).

Different deviations from the ideal baseline can be observed among different subjects. However, it can be observed that the baseline offsets within each group of (fast) abrupt movements are relatively constant (Fig. 4, frame a and b). When the duration between motion events is longer, i.e., in motion and hold blocks, the offset varies more. This observation could be explained by proprioceptive memory that is known to allow subjects to retain accurate motion memory during a repetitive movement reproduction task, when it is performed within short delays between trials [25]. Therefore, to obtain a reliable angle amplitude measurement, we estimate a *local baseline* before each motion onset. The actual angles are computed by taking the maximum peak amplitude for abrupt motion and the mean angle amplitude for sustained motion with respect to the local baseline for each time series (Fig. 3).

3.3. Temporal resolution increase by volume splitting and fast motion estimation

Taking the maximum angle value is justified because temporal upsampling removes mixtures of different movements occurring within one TR by splitting them into separate time series. Therefore, the highest peak is the best motion estimation available.

Abrupt peak detection and mean angle computation between motion and hold peaks was performed automatically with a sliding window approach (similar to some well-known peak-detection algorithms as in [26]), but with prior knowledge of the target, which allowed us to directly place the search window around the correct positions and avoid outliers.

With temporal upsampling and spatial interpolation, a significant abrupt peak amplitude improvement is obtained as shown in Fig. 3b. The typical amplitude gain is shown. As expected, motion amplitude estimation for "motion and hold" sequences does not change much with upsampling, since head positions are maintained for periods longer than the duration of one TR (acquisition of one volume). We recall that the goal of temporal upsampling is to improve the estimation of fast motion (occurring while a volume is acquired), which, in turn, allows to validate motion reproduction.

3.4. Subject motion accuracy and reproducibility

It is important to point out that we care about *intra-session* reproducibility, i.e., that a subject is able to perform consistently within a session, since the motion pattern estimated by EPI time series realignment is applied to structural imaging performed during the same session. It likely that the response function cannot be applied in another session, since the estimated motion parameters even for the same subject is prone to change significantly due to experimental conditions (e.g., different FoV, and other parameters mentioned in Section 3.1).

In Figs. 5 and 6, errors with respect to the target are reported by type of motion (U/D or R/L, abrupt or motion&hold) and by target angle amplitude. In Fig. 5 the data acquired for all subjects is aggregated in each boxplot to show the error distributions for each target angle and motion type. The first measurement for each subject was discarded due to the effect of transient adaptation to the task. Hence, each boxplot represents the distribution of 6 subjects * 4 measurements = 24 data points per target angle and per motion type. In Fig. 6 the 5 measurements (i.e., including the first) for each motion block (abrupt or motion&hold) are shown separately. Each color corresponds to one subject. Measures are connected by a line to make the graph easier to read. The dashed line is the average measure for each subject at each point and the gray shaded area shows one standard deviation above and below the mean. The error is computed by taking local baselines into account, for each subject. It is the difference between the target angle and the effective angle. Each row shows results for the same type of motion and each column corresponds to a given angle.

For all subjects, motion accuracy error increases with angle amplitude. It can also be observed that subjects were able to quickly generalize the learning process. Considering S2: we can infer that there was a "strong learning effect" during the first two Right/Left motion blocks (Fig. 4). Afterwards, subject angle accuracy is quite stable even for the first Up/Down motion blocks, although there is a slight improvement near the end of the acquisition. Moreover, rapid convergence can be observed in Fig. 6, mostly in the first row, i.e., motion&hold in Right/Left direction, as the standard deviation decreases quickly. That direction (Right/Left) corresponds to the first training session and suggests that most of the learning was achieved there. Indeed, in the third row, i.e., for motion&hold in the Up/Down direction, no such phenomenon is visible, and the standard deviation is very low. This leads us to the conclusion that subjects are able to generalize well the task to perform from only a few training blocks. Therefore, other researchers interested in this method can reduce training time and perform a single training session with a mixture of the different motion types, according to their needs.



Fig. 6. Error (difference between target and effective subject angle) per angle and per motion type for all subjects (baseline taken into account). The black dashed line is the mean of each group of measurements; shaded area is the standard deviation for each group. The standard deviation is in general positively correlated with angle.

3.5. Generating motion artifacted images

Representative images acquired with the proposed protocol are presented in Fig. 7 to illustrate the type of motion artifacts produced. The angles used for head choreographies were chosen empirically, based on the range of angles observed in clinical exams, and arbitrarily set to three angles covering this range, namely 1,2, and 3 degrees. The assumption being that it is unrealistic for a given subject – except in young pediatric patients – to move much more than this in a padded head coil. The proposed head choreographies used the same motion patterns than the EPI training sessions.

Classical T1 Spin-Echo is widely used in clinical settings for Gduptake scans. Hence, studying such kind of pulse sequence presents a practical interest and has the advantage of following a simple sequential and Cartesian k-space sampling scheme. Examples of artifacts produced during a 2D Spin-Echo acquisition are shown in Fig. 7. Motion artifacts are qualitatively similar for all subjects and take the form of "ringing artifacts" in the phase encoding direction (R/L).

An example of empirical labeling with 6 DoF is shown in Fig. 8. The average k-space magnitude over all slices of the corresponding acquisition (Spin-Echo) is shown next to the label to highlight the corrupted regions. Motion impacts mostly the phase-encoding direction, since this is the slowest one to acquire. The effect of motion can be seen as dark lines appearing in k-space magnitude at phase encoding lines corresponding to motion onset. This well-known artifact [27] depends on the position of affected lines in k-space as well as the type of motion.

The 3D MP-RAGE acquisition presents more complex patterns as it uses two phase encoding directions, namely phase encoding per se (A/P) and slice encoding (R/L). Therefore, visual inspection of kspace magnitude is not as informative as for a classical Spin-Echo. Representative examples of motion artifacts observed in 3D T1 MP-RAGE acquisitions are shown in Fig. 7. Mild ringing artifacts can be observed in the frontal regions of the sagittal and axial slices (see



Fig. 7. Sample of acquired Spin echo (axial) artifacts (a, b, c) and MP-RAGE (coronal) (d, e, f). It can be observed that artifacts are qualitatively similar (ringing/ghosting with similar spacing) and clearly visible for all subjects, which is what we aimed to achieve. MP-RAGE artifacts are milder than those observed on Spin echo acquisitions.

Appendix) and towards the upper right parts on the coronal views. The artifacts are qualitatively similar for all subjects, which suggests that head choreographies could be well reproduced. In addition to being visually similar, the artifacts appear in the expected regions.



Fig. 8. Spin-Echo k-space magnitude combined from raw data with the ESPIRiT algorithm [22] for one subject (all slices averaged). For classical Spin-Echo, the full k-space is acquired sequentially line by line (in 180 [s] in this case). Right: screen target and dataset label (6 DoF). (Head choreography timing performed by subject closely matches k-space corruption.)



Fig. 9. Slices from three representative 3D T1 MP-RAGE acquisitions in sagittal and axial directions corrupted by "motion and hold" R/L choreography. The artifacts are qualitatively similar among different subjects. Mild ringing-like artifacts can be observed near the frontal lobe.

The proposed method is usable as such for repetitive movements because it can be considered as an "anticipation timing performance" task and subjects response time can be neglected [24]. Our system would allow generating head choreography tracking movies based on real motion data (estimated from EPI sequences or by other means) but in that case, subjects' temporal accuracy should be measured.

To summarize, volume registration confirms correct head choreography reproduction and provides empirical labels for the dataset. The same choreography performed by different subjects resulted in qualitatively similar artifacts, indicating that the method may be useful in artifact simulation validation. Our experimental protocol providing well-defined motion patterns and amplitudes will help to better understand the effects of motion on MRI images and will provide a practical basis to compare or validate motion models. The method can easily be extended to more complex motion patterns according to specific needs.

3.6. Limitations of the system

The proposed approach is inherently limited to head motion. The screen is a 2D projection of the 6-DoF movement of a subject's head. Left/right or up/down motion on screen will be translated to a combination of these 6-DoF (mainly yaw and pitch, respectively). Roll motion is difficult to reproduce accurately and would require adaptations of the system. A rotating crosshair could be displayed on the moving dot and customized optics could be adjusted on the glasses.



Fig. 10. Motion estimation plots for temporally upsampled and interpolated EPI volumes for all subjects. This plot is provided to allow a global view of all subject's performance side by side.





Fig. 11. (a & b) Raw volume-to-volume realignment output from SPM. (c) 3-DoF matrix for in-plane rotation with (t_x, t_y) offset from image center showing the preservation of the rotation and residual translations (a similar equation can be derived for 6DoF).

It would not be practically feasible to constrain a subject to do pure translations, but this type of motion is not realistic in the scope of a clinical scan, except for a slight translation in the axis of the scanner (z axis), most likely due to subject neck relaxation in the beginning of the scan (cf. Appendix C).

4. Conclusion

A versatile, straightforward and inexpensive methodology to control head movements in an MRI scanner is proposed. A method to improve registration in case of motion faster than TR has been developed. The apparatus and protocols provided allows systematic and reproducible



Fig. 12. Schematic view of the hardware setup used in our experiments. For pulse sequence providing trigger pulses, the scanner controls the start of the video. When triggering is unavailable, optical pulses are sent to the scanner from Arduino Uno. (In the worst case, manual triggering would always be possible.)

motion-controlled experiments, which, in turn, opens means to better understand the effects of motion and provides a practical basis for simulation. The method can also be used for generating realistic motion-corrupted datasets suitable for benchmarking any retrospective motion-correction technique or training supervised machine learning algorithms.

The proposed methodology can be adapted by researchers based on their needs. For example, a different range of angles and motion types could be investigated depending on the application. In particular, it can help to study the impact of motion on different k-space regions.

Acknowledgments

We acknowledge access to the facilities and expertise of the CIBM Center for Biomedical Imaging founded and supported by Lausanne University Hospital (CHUV), Switzerland, University of Lausanne (UNIL), Switzerland, École polytechnique fédérale de Lausanne (EPFL), Switzerland, University of Geneva (UNIGE), Switzerland and Geneva University Hospitals (HUG), Switzerland.

Appendix A. Other experimental results

See Fig. 9.

Appendix B. EPI data of all subjects

See Fig. 10.

Appendix C. Raw realignment data

See Fig. 11.

Appendix D. Ad hoc setup for video synchronization

See Fig. 12.

References

- Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. J Magn Reson Imaging 2015;42(4):887–901.
- [2] Pipe JG. Motion correction with propeller MRI: Application to head motion and free-breathing cardiac imaging. Magn Reson Med: Off J Int Soc Magn Reson Med 1999;42(5):963–9.
- [3] Liu C, Bammer R, Kim D-h, Moseley ME. Self-navigated interleaved spiral (snails): Application to high-resolution diffusion tensor imaging. Magn Reson Med: Off J Int Soc Magn Reson Med 2004;52(6):1388–96.
- [4] Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: A review. Magn Reson Med 2013;69(3):621–36.
- [5] Qin L, van Gelderen P, Derbyshire JA, Jin F, Lee J, de Zwart JA, Tao Y, Duyn JH. Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system. Magn Reson Med: Off J Int Soc Magn Reson Med 2009;62(4):924–34.
- [6] Kyme AZ, Aksoy M, Henry DL, Bammer R, Maclaren J. Marker-free optical stereo motion tracking for in-bore MRI and PET-MRI application. Med Phys 2020;47(8):3321–31.
- [7] Aksoy M, Forman C, Straka M, Çukur T, Hornegger J, Bammer R. Hybrid prospective and retrospective head motion correction to mitigate cross-calibration errors. Magn Reson Med 2012;67(5):1237–51.
- [8] Singh A, Zahneisen B, Keating B, Herbst M, Chang L, Zaitsev M, Ernst T. Optical tracking with two markers for robust prospective motion correction for brain imaging. Magn Reson Mater Phys Biol Med 2015;28(6):523–34.
- [9] Zaitsev M, Dold C, Sakas G, Hennig J, Speck O. Magnetic resonance imaging of freely moving objects: Prospective real-time motion correction using an external optical motion tracking system. Neuroimage 2006;31(3):1038–50.
- [10] Afacan O, Wallace TE, Warfield SK. Retrospective correction of head motion using measurements from an electromagnetic tracker. Magn Reson Med 2020;83(2):427–37.
- [11] Gholipour A, Polak M, Van Der Kouwe A, Nevo E, Warfield SK. Motion-robust MRI through real-time motion tracking and retrospective super-resolution volume reconstruction. In: 2011 annual international conference of the IEEE engineering in medicine and biology society. IEEE; 2011, p. 5722–5.
- [12] Norbeck O, van Niekerk A, Avventi E, Rydén H, Berglund J, Sprenger T, Skare S. T1-FLAIR imaging during continuous head motion: Combining propeller with an intelligent marker. Magn Reson Med 2021;85(2):868–82.
- [13] Fu ZW, Wang Y, Grimm RC, Rossman PJ, Felmlee JP, Riederer SJ, Ehman RL. Orbital navigator echoes for motion measurements in magnetic resonance imaging. Magn Reson Med 1995;34(5):746–53.
- [14] Welch EB, Manduca A, Grimm RC, Ward HA, Jack Jr. CR. Spherical navigator echoes for full 3D rigid body motion measurement in MRI. Magn Reson Med: Off J Int Soc Magn Reson Med 2002;47(1):32–41.
- [15] Wallace TE, Afacan O, Waszak M, Kober T, Warfield SK. Head motion measurement and correction using FID navigators. Magn Reson Med 2019;81(1):258–74.
- [16] Herbst M, Maclaren J, Lovell-Smith C, Sostheim R, Egger K, Harloff A, Korvink J, Hennig J, Zaitsev M. Reproduction of motion artifacts for performance analysis of prospective motion correction in MRI. Magn Reson Med 2014;71(1):182–90.
- [17] Maclaren J, Armstrong BS, Barrows RT, Danishad K, Ernst T, Foster CL, Gumus K, Herbst M, Kadashevich IY, Kusik TP, et al. Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain. PLoS One 2012;7(11):e48088.
- [18] Hoinkiss DC, Erhard P, Breutigam N-J, von Samson-Himmelstjerna F, Günther M, Porter DA. Prospective motion correction in functional MRI using simultaneous multislice imaging and multislice-to-volume image registration. NeuroImage 2019;200:159–73.
- [19] Franchak JM, McGee B, Blanch G. Adapting the coordination of eyes and head to differences in task and environment during fully-mobile visual exploration. PLoS One 2021;16(8):e0256463.
- [20] Fang Y, Nakashima R, Matsumiya K, Kuriki I, Shioiri S. Eye-head coordination for visual cognitive processing. PLoS One 2015;10(3):e0121035.
- [21] Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE. Statistical parametric mapping: The analysis of functional brain images. Elsevier; 2011.
- [22] Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M. ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where sense meets grappa. Magn Reson Med 2014;71(3):990–1001.
- [23] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016, http: //www.deeplearningbook.org.
- [24] Thomas JR, Gallagher JD, Purvis GJ. Reaction time and anticipation time: Effects of development. Res Q Exerc Sport 1981;52(3):359–67.
- [25] Strimpakos N. The assessment of the cervical spine. Part 1: Range of motion and proprioception. J Bodyw Mov Ther 2011;15(1):114–24.
- [26] Palshikar G, et al. Simple algorithms for peak detection in time-series. In: Proc. 1st Int. Conf. advanced data analysis, business analytics and intelligence, vol. 122, 2009.
- [27] Wood ML, Shivji MJ, Stanchev PL. Planar-motion correction with use of k-space data acquired in Fourier MR imaging. J Magn Reson Imag 1995;5(1):57–64.