NMR in Biomedicine

WILEY

NMR IN BIOMEDICINE

RESEARCH ARTICLE OPEN ACCESS

Intra-MRI Head Motion Tracking and Correction: A Quantitative In Vivo Evaluation Framework

Zakaria Zariry[1,2] | Franck Lamberton[3,4] | Robert Frost[5,6] | Thomas Gaass[7] | Thomas Troalen[8] | Holly Rayson[1,2] | Jakob M. Slipsager[7] | Nathalie Richard[1,2] | Andre van der Kouwe[5,6] | James Bonaiuto[1,2] | Bassem Hiba[1,2]

[1]CNRS, Institute of Cognitive Sciences Marc Jeannerod, Bron, France | [2]Université Claude Bernard, Lyon 1, Villeurbanne, France | [3]Cermep, CNRS/UAR 3453, Lyon, France | [4]SFR Santé Lyon-Est, CNRS UAR3453, INSERM US7, Université Lyon 1, Lyon, France | [5]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts, USA | [6]Department of Radiology, Harvard Medical School, Boston, Massachusetts, USA | [7]TracInnovations, Ballerup, Denmark | [8]Siemens Healthcare SAS, Courbevoie, France

Correspondence: Bassem Hiba (bassem.hiba@isc.cnrs.fr)

Received: 11 February 2026 | **Revised:** 20 July 2026 | **Accepted:** 24 July 2026

Keywords: brain MRI | fat navigator | markerless optical system | motion artifacts | motion correction | movement tracking

ABSTRACT

Despite numerous magnetic resonance imaging (MRI) head motion mitigation strategies, the lack of rigorous evaluation limits their optimization and clinical adoption. We propose an in vivo framework combining a visual instruction system for reproducible head motion with reference standard interpose displacement estimation to assess intra-MRI tracking accuracy and precision. Its utility is demonstrated by comparing a markerless optical system (MOS) and a fat-signal navigator (FatNav). Six participants underwent 3T T1-weighted brain MRI with a FatNav module, performing visually guided 2° and 4° head rotations around the *X*- and *Z*-axes using MOS feedback. T1-weighted images were acquired at seven distinct head poses. MOS and FatNav motion estimates were compared against rigid registration of the T1-weighted images, which served as the reference standard. MOS- and FatNav-corrected images for the three successive head rotations were also compared using the structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), and a focus measure. FatNav accuracy was inferior for translations ($p < 0.001$) and 2°–4° rotations but improved to match MOS for subtle pitch$^+$ and yaw$^+$, even surpassing it for subtle yaw$^-$. Meanwhile, MOS precision was higher for yaw$^+$ than yaw$^-$ ($p < 0.001$) but inferior to FatNav for pitch$^+$ ($p = 0.041$). MOS better restored T1-weighted image fidelity, yielding higher SSIM, PSNR, and focus ($p < 0.01$). Notably, the framework detected a subtle improvement in FatNav performance with neck masking, an effect uncaptured by conventional image quality metrics. In conclusion, while image quality metrics suggested superior overall correction with MOS, our framework provided a more detailed characterization of in vivo performance differences.

1 | Introduction

Motion artifacts represent a prevalent source of degradation in magnetic resonance imaging (MRI). When left uncorrected, they can obscure subtle pathological changes and potentially lead to the misinterpretation of clinical images [1, 2]. Furthermore, these artifacts incur substantial, undesirable costs for hospitals [3]. Andre et al. estimated that repeated sequences are present in 19.8% of all MRI examinations performed on inpatient and/or emergency department patients and that 16.4% of all attempted sequences contained moderate or severe motion artifacts and should have been repeated based on their criteria [4]. This issue is particularly critical for high-resolution brain MRI [5], which has gained renewed interest with the advent of 7-T scanners [6], and for imaging certain populations such as children [7].

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *NMR in Biomedicine* published by John Wiley & Sons Ltd.

A multitude of approaches have been proposed to mitigate head motion artifacts during MRI acquisitions. One class of techniques relies on so-called "navigators". While some of these methods acquire sparse k-space regions, others rely on short acquisitions, timed before or after the echo-train at each repetition time (TR) of the imaging sequence, generating a series of low-resolution three-dimensional images of the head. These images are then registered, either prospectively or retrospectively, to estimate and compensate for head motion at each TR. The fat navigator (FatNav) strategy employs spectrally selective RF excitation to acquire subcutaneous fat images, which serve as a head motion navigator [8, 9].

An alternative strategy involves external devices designed to monitor the head's position and orientation relative to the scanner. For example, electroencephalography sensors [10], electromagnetic sensors [11], and optical tracking cameras [12, 13] have been proposed to track and correct rigid head movements during MRI. In parallel, advanced image acquisition and reconstruction techniques [14], often data-driven and increasingly powered by artificial intelligence [15, 16], are also under active investigation.

Given the proliferation of head motion correction methods in MRI, there is an urgent need for rigorous evaluation strategies that not only allow radiologists and neuroimaging researchers to appreciate the image quality benefits of each technique but also guide developers in refining their approaches [17]. Several research groups have utilized head phantoms affixed to mechanical positioning systems capable of precise rotational and translational displacements to establish the prescribed motion parameters as a gold-standard for motion tracking validation. Errors in phantom motion tracking are then directly quantified by comparing the estimated rotations/translations values against the gold-standard [18, 19]. However, the complex anatomy and composition of the human head, and its actual movements governed by complex biomechanical constraints, are not easily replicable using phantoms. Moreover, external motion sensors like optical cameras may produce different estimates for an identical movement when performed by a human head compared to a phantom, due to inherent differences in appearance and texture. Finally, Gumus et al. found that head motion tracking errors measured in vivo exceeded those detected in offline phantom tests [20].

In the absence of an established in vivo gold-standard for head motion estimation accuracy, most evaluation strategies rely on indirect comparisons between motion-corrected images and presumed motionless reference images [21, 22]. Although widely used in vivo to assess motion artifact mitigation, these approaches fail to directly quantify the precision and accuracy of head motion estimation. Moreover, evaluations based on image quality metrics, morphometric consistency [7, 18, 23], or the preservation of clinically relevant features [24], may be subject to biases stemming from the quality of the reference images and the limited specificity of the employed image quality metrics [22]. Consequently, recent efforts have shifted toward developing more direct, quantitative evaluation strategies, as understanding the fundamental accuracy and precision limits of tracking systems is essential for optimizing these corrections [25, 26].

Despite substantial progress in head motion tracking and correction techniques in MRI, the absence of controlled and standardized in vivo evaluation protocols thus remains a major bottleneck, hindering not only the optimization and objective comparison of these methods but also both their extension beyond anatomical imaging and their routine adoption in clinical practice.

To bridge this gap, the present study proposes an in vivo evaluation framework capable of quantifying both the precision and accuracy of head motion tracking, incorporating a FreeSurfer registration-based reference standard and an intra-MRI visual instruction system designed to enforce controlled, reproducible head movements across subjects. To validate the proposed evaluation framework, we applied it to the in vivo assessment and direct comparison of two fundamentally different head motion tracking approaches, one based on a markerless optical system (MOS) [18] and the other on the 3D FatNav strategy [9], in six healthy volunteers. This comparative analysis served as a primary use case to demonstrate the ability of the framework to objectively quantify motion-estimation accuracy and precision across distinct tracking technologies under controlled and reproducible motion conditions.

Importantly, the conclusions drawn from the proposed parameter-level evaluation framework were further corroborated by conventional image-based assessment methods, including standard image quality and reconstruction fidelity metrics. This consistency confirms that the proposed framework yields results that are coherent with established evaluation practices, while providing complementary and more direct insights into motion-estimation accuracy.

Beyond cross-system comparison, we further demonstrate that the proposed framework enables the quantitative validation of targeted improvements within a given tracking method. Specifically, we assessed the impact of applying a neck-mask to FatNav images [27].

2 | Materials, Methods, and Participants

2.1 | MRI Materials and Participants

MRI acquisitions were conducted on a MAGNETOM Prisma 3-T scanner (Siemens Healthineers, Erlangen, Germany) equipped with a 64-channel receive coil. Six healthy volunteers (31.5 ± 6.2 years) were recruited in accordance with the Declaration of Helsinki, and all participants gave written informed consent which was approved by the regional ethics committee for human research (CPP Sud Mediterranee V-2021-A00042-39).

A magnetization-prepared rapid gradient echo (MP-RAGE) pulse sequence, incorporating a Fat navigator module [9] following each TR, was employed for sagittal T1w-image acquisition. Simultaneously with the MP-RAGE acquisition, head movement parameters were recorded every ~33 ms using the markerless optical device Tracoline in combination with the TracSuite v3.3 software [28] (TracInnovations, Ballerup, Denmark). The Tracoline probe was positioned adjacent to the head coil to capture part of the facial surface through the coil's aperture in front of the right eye (Figure S1A). This system utilizes a structured infrared light approach to acquire three-dimensional point clouds of the subject's face at a rate of approximately 30 frames

per second. These point clouds are subsequently co-registered with a reference model to compute rigid body pose estimates. To prevent tracking biases caused by nonrigid movements (e.g., blinking), the periocular region is typically excluded manually from the reference point cloud for certain subjects.

Prior to each MRI session, a MOS/MRI scanner cross-calibration procedure [29] was performed to determine the coordinate transformation between the optical system and the MRI scanner. This involved a short 17-s MRI acquisition of the subject's face, from which a 3D point cloud was reconstructed and subsequently registered to the point cloud generated by the optical camera (Figure S1B). Highly integrated and automated within Tracoline, these procedures require minimal user intervention.

2.2 | Motion-Correction and Image Reconstruction

After acquisition, the Fourier-space data were transferred to a dedicated server for motion correction and image reconstruction. The RetroMocoBox toolbox [30] was then used to reconstruct the 3D fat navigator images, before rigidly realigning them to estimate the head movement parameters at each TR, relative to the first TR. Two sets of motion parameters (three rotations and three translations) were computed for each MP-RAGE image based on (1) the original 3D fat navigator images and (2) 3D-fat navigator images after neck masking.

The neck region was masked on fat images to account for its nonrigid movement relative to the skull vault (Figure S2) [22]. A semi-automatic neck-masking procedure was applied for each subject. First, 3D fat images from all poses were registered to a reference fat image, which was acquired simultaneously with the k-space center sampling of the T1w-image at the reference pose. A 3D mask was then manually defined over the neck area of this reference image and propagated to exclude the neck region across the subject's entire dataset. Finally, an inverse transformation was applied to each 3D fat image to restore its original orientation prior to head motion estimation.

A custom Python pipeline was developed in which missing k-space data were first reconstructed using the GRAPPA algorithm [31]. To correct for motion, the fully sampled k-space data were adjusted using rigid transformations applied line by line. These transformations were estimated from either MOS or FatNav motion tracking curves, which were interpolated in time to match each line of Fourier-space data. Motion estimates were mapped onto the k-space trajectory via nearest-neighbor temporal interpolation and, for parallel-imaging, cubic interpolation of the estimates flanking each GRAPPA-synthesized line, both implemented within our open-source reconstruction framework. The final T1w-image was reconstructed using a nonuniform inverse Fourier transform.

2.3 | Experimental Setup and Head Motion Guidance

To maintain uniform and reproducible head movements among participants, custom Python-based software was developed to provide visual instructions and feedback within the MRI environment. The 2D visual instruction animation was projected onto a screen at the back of the MRI tunnel via an external projector positioned outside the MRI room, facing a waveguide. Participants viewed the screen through a mirror mounted atop the 64-channel coil.

The animation featured a white background with a red square that dynamically reflected the subject's head position, as measured in real time by the MOS system, and a square with a blue frame indicating the target position (Figure 1). A fixed threshold of 0.1° between successive estimation points was applied to the visual feedback to filter out minor tracking noise and avoid presenting a jittery or distracting display to the participant. The participant is instructed to align the red square to the blue one by adjusting their head rotations. The red square's lateral displacement corresponded to rotations around the *Z*-axis, while its vertical movement was linked to rotations around the *X*-axis. Additionally, a thin black border surrounding the red square rotated according to the *Y*-axis rotation. A text message appeared 5 s before each movement to alert participants and prevent abrupt or forceful motions. To maintain participant comfort and prevent excessive instruction overload, we restricted visual guidance to head rotations only.

Participants performed head rotations rightward (yaw^-), upward ($pitch^+$), and leftward (yaw^+) of 2° or 4° around a single primary axis (*X* or *Z*). Associated minor head rotations were observed along the secondary axes of each rotation (primary rotation around the *X*-axis accompanied by minor secondary rotations around the *Y*- and *Z*-axes, while primary rotation around the *Z*-axis involved minor secondary rotations around the *X*- and *Y*-axes). Similarly, head rotations induce translations in all three spatial directions (Figure S3).

2.4 | MRI Experiments

2.4.1 | Experiment 1: Estimation Accuracy and Precision of MOS and FatNav

This experiment aimed to estimate movement parameters in degrees and millimeters for various stereotypical head poses relative to a reference position. The T1w-image acquired at each head pose was registered to the reference position image using FreeSurfer's *mri_robust_register* tool [32]. This procedure yields a transformation matrix that describes the ideal rigid alignment between the two head positions, establishing the reference standard.

The head locations for each motion-corrected T1w-image were determined based on the fat navigator location closest to the Fourier space center sampling or the average MOS locations recorded during the ~330 ms around this sampling.

Each participant underwent seven separate MP-RAGE scans with an isotropic spatial resolution of 1 mm and acquisition parameters optimized to limit its duration to 4 min 25 s (Table 1).

Before each T1w-scan, participants adjusted their heads to a target position displayed on the screen and remained still. The

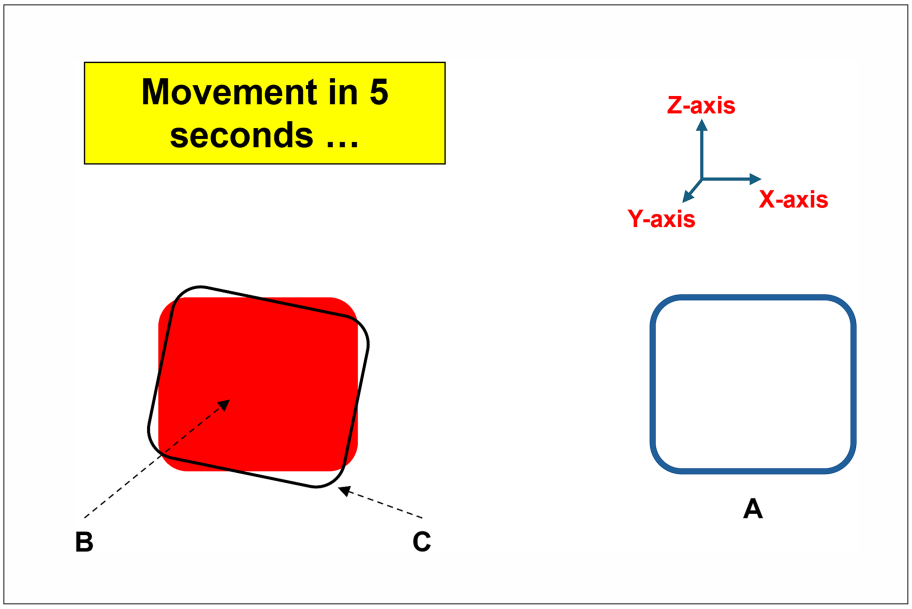


FIGURE 1 | Visual instructions for head rotation guidance. Vertical and horizontal displacements of Squares A and B relative to the isocenter represent rotations around the *X*- and *Z*-axes, respectively. Square B displays real-time position (via MOS), while Square A indicates the target. Rotations of Square C relative to B reflect motion around the *Y*-axis. The example illustrates a 4° rightward rotation command, requiring the participant to align Square B horizontally with Square A.

first scan, with the head straight and centered, served as a reference. Before the next three scans, participants performed yaw$^-$, pitch$^+$, and yaw$^+$ rotations of 2° relative to the reference position. These rotations were then repeated with a 4° amplitude for the final three scans.

The head motion tracking accuracy of MOS and FatNav was assessed by computing the rotational and translational errors for each head position against the reference standard, expressed as the mean absolute error across the six subjects (Figure 2).

To compute the precision of the MOS and FatNav methods, the residual standard deviations of the errors (defined as the reference standard minus the system estimate) were quantified for each head rotation and averaged across the six subjects.

To account for T1w-image distortion variations related to head positioning, rigid, affine, and nonlinear registrations were performed using ANTs [33] between each T1w-image and its corresponding reference image. The distortion-corrected image was then transformed back into its native space using the inverse of the rigid transformation. Finally, the FreeSurfer rigid registration was applied to obtain a distortion-corrected reference standard for head motion estimation.

2.4.2 | Experiment 2: Image Quality Recovery by MOS and FatNav

For each participant, three additional T1w-images (6 min each) were acquired, with settings adjusted for optimal gray to white-matter contrast (Table 1). The participant was instructed to keep their head still at the reference position for the first image. During the second acquisition, the participant was visually instructed to center their head before executing three head rotations of 2° relative to the reference position: yaw$^-$, pitch$^+$, and yaw$^+$, at 1 min 20 s, 2 min 30 s, and 4 min, respectively, after the scan began. The third acquisition followed the same head rotations with amplitude increased to 4°. T1w-images were each reconstructed three times using the retrospective motion-correction pipeline with motion estimates derived from MOS, FatNav, and neck-masked FatNav respectively. The reference T1w-image was corrected for any unintended head movement occurring during its acquisition using MOS estimates.

A brain extraction was carried out using the BET [34] tool. Then, each T1w-image was aligned with the reference image using *mri_robust_register* to facilitate quantitative quality assessments based on three metrics: structural similarity index measure (SSIM) [35], peak signal-to-noise ratio (PSNR), and focus measure (variance of Laplacian) [36]. Finally, a visual quality evaluation of motion-corrected images was performed by three independent blinded experts who evaluated three key criteria: sharpness, noise, and ghosting artifacts, using a scoring system from 1 (*poor*) to 5 (*excellent*).

2.5 | Statistical Analysis

Pairwise statistical tests were performed on the motion estimation errors. To evaluate performance differences between MOS and FatNav, and between the two FatNav variants, a linear mixed-effects model was employed, including tracking method as a fixed effect, and subject-specific intercepts as random effects, treating different angles (2° and 4°) and directions (yaw and pitch) as repeated measures. These analyses were

TABLE 1 | Parameters of sagittal T1-weighted MP-RAGE pulse sequence and its FatNav module.

Acquisition parameter	Experiment 1	Experiment 2
Isotropic resolution (mm)	1	1
Field-of-view (mm^3)	256×256×192	256×256×192
Inversion time (ms)	900	1070
Echo time (ms)	3.3	3.3
Repetition time (ms)	2180	2500
Flip angle (°)	8	8
Bandwidth (Hz/pixel)	250	240
Parallel imaging: Acceleration factor	3×1	2×1
Parallel imaging: nb. of ref. lines	40	32
Turbo factor	208	192
Nb. of 3D fat navigator images	112	143
FatNav - resolution (mm)	4	
FatNav - TE/TR (ms)	1.43/3.4	
FatNav - flip angle (°)	7	
FatNav - acceleration factor	4×4	
FatNav - bandwidth (Hz/pixel)	1950	
Total acquisition time	4 min 9 s	6 min 2 s

conducted using R (version 4.4.1) [37] and *lme4* (version 1.1.36) [38]. Fixed effects were assessed using Type II Wald *F* tests with Kenward-Roger estimated degrees of freedom (*car* version 3.1.3) [39]. Pairwise Tukey-corrected follow-up tests were run using estimated marginal means from the *emmeans* package (version 1.10.6) [40]. A significance threshold of $p<0.05$ was adopted. Where applicable, effect sizes (standardized mean difference, Cohen's *d*) were calculated to quantify the magnitude of the observed differences.

Precision statistical assessments across the six subjects were also performed using linear mixed-effects models, fitted independently for each experimental condition, which comprised movement direction and the principal rotation axes (yaw, Rz; pitch, Rx). Each model incorporated rotation magnitude and tracking system as fixed effects, with subject-specific variations accounted for via a random intercept.

For a comparative precision analysis of MOS and FatNav, two nested models were compared per condition, a homoscedastic null model, and a heteroscedastic alternative allowing per-system residual variance and evaluated by likelihood ratio test (df = 1, $\alpha=0.05$). Directional symmetry of precision (yaw^+ vs. yaw^- on Rz) was assessed independently for each system using the same framework, with movement direction replacing system in the variance structure.

3 | Results

3.1 | Head Motion Reproducibility

The head motion control strategy demonstrated high reproducibility across participants. For instructed 2° rotations, the reference standard values were 2.09° ± 0.39° for yaw^-, −1.88° ± 0.43° for $pitch^+$, and −1.98° ± 0.05° for yaw^+. Similarly, for a 4° target head rotation, the reference standard values measured were 4.13° ± 0.33°, −4.01° ± 0.32°, and −4.08° ± 0.22° in the respective directions. In addition to the expected rotations around the primary axes (*Z* for yaw and *X* for pitch), associated translations of the head and minor rotations around the secondary axes (*X* and *Y* for yaw and *Y* and *Z* for pitch) were also observed (Table S1).

3.2 | Accuracy and Precision of Head Motion Estimation

A global analysis of Experiment 1 estimation errors covering the 108 rotations (6 subjects × 6 head positions × 3 axes) and the 108 translations indicates that both MOS and FatNav with neck masking provided head movement estimates closely approximating the reference standard (rotational and translational mean errors: 0.24° ± 0.24° and 0.31 ± 0.29 mm for MOS and 0.28° ± 0.24° and 0.65 ± 0.68 mm for FatNav, respectively). While both methods performed similarly in estimating head rotations, MOS provided more accurate estimates of translational shifts ($p<0.001$; Figure 3).

For a more granular evaluation, performance differences between MOS and FatNav were independently assessed for yaw^-, $pitch^+$, and yaw^+ rotations, with a separate analysis of large-amplitude rotations (2° and 4°) around the primary axes on one side and smaller amplitude rotations around the secondary axes on the other side. Figure 4 and Table 2 show results from Experiment 1, demonstrating that the mean errors for primary rotations were 0.21° ± 0.23° for MOS and 0.50° ± 0.22° for FatNav (8.4% ± 11.4% and 17.5% ± 8.1%, respectively). MOS showed mean errors of 0.26° ± 0.25° for secondary rotations and 0.31 ± 0.29 mm for translations, whereas FatNav had mean errors of 0.17° ± 0.17° and 0.66 ± 0.68 mm, respectively.

Mixed model *emmeans* ($\alpha=0.05$) results confirmed that MOS reduced mean primary rotation errors by 38% for $pitch^+$ ($p=0.023$), 77% for yaw^+ ($p<0.001$), and 53% for yaw^- ($p=0.004$) compared to FatNav, demonstrating significantly better performance. MOS also demonstrated a superior performance on translation estimation, reducing mean errors of yaw^+ by 59% ($p=0.002$) and of $pitch^+$ by 60% ($p<0.001$), compared to FatNav.

In contrast, FatNav exhibited better accuracy than MOS for small rotations around a secondary axis during yaw^-, reducing mean errors by ~50% ($p=0.0138$, Figure 4A and Table 2).

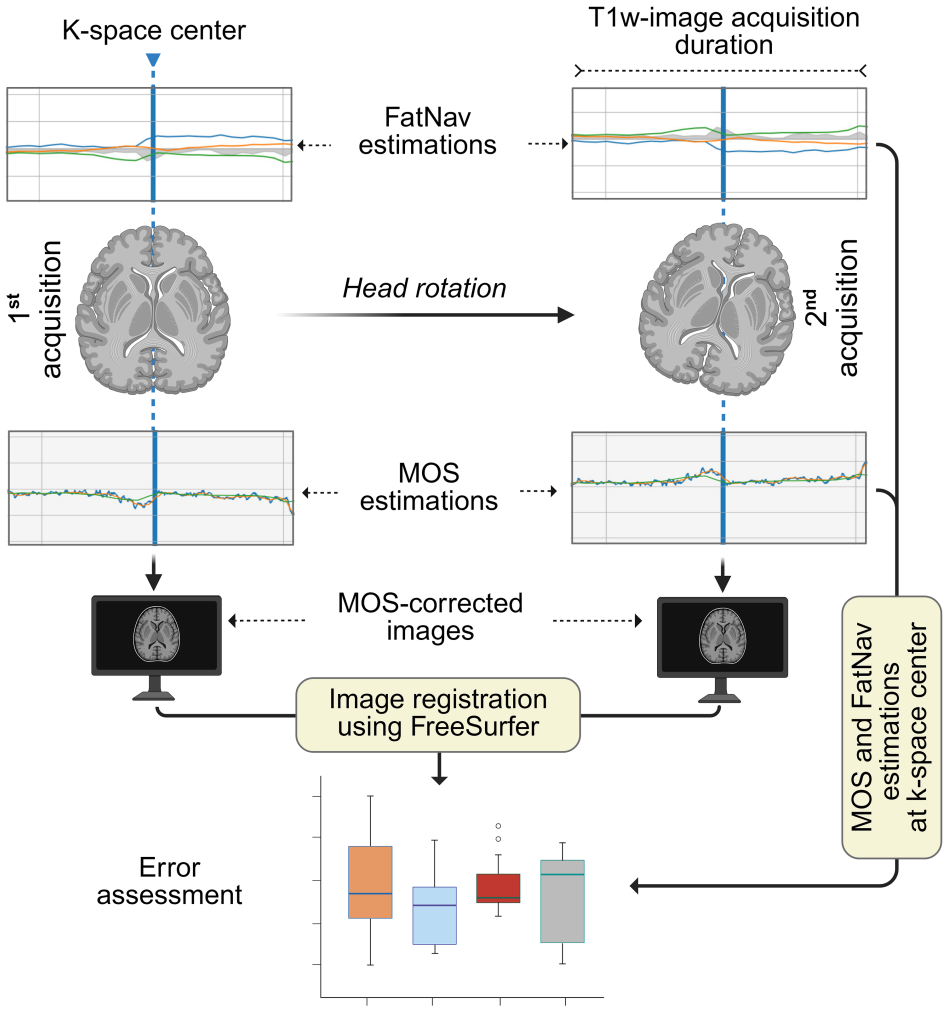


FIGURE 2 | Procedure for assessing errors in head motion estimation by MOS and FatNav. The transformation matrices considered for MOS and FatNav are those measured during Fourier-space center sampling. FatNav and MOS head motion estimations are compared with the reference standard, obtained by registering each T1w-image to its reference.

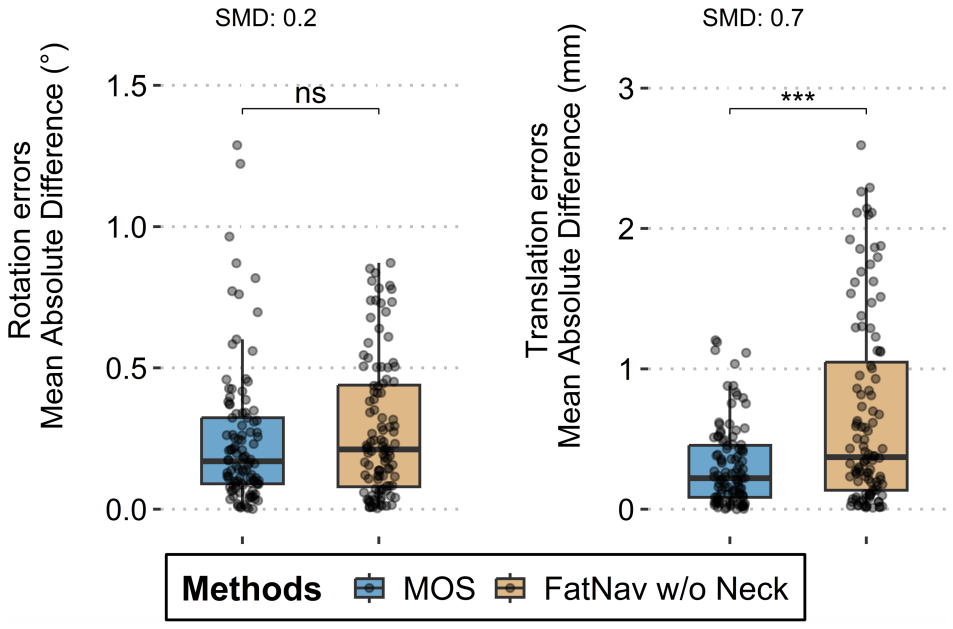


FIGURE 3 | Global comparison of head motion rotation and translation estimation accuracy between MOS and FatNav with neck-mask.

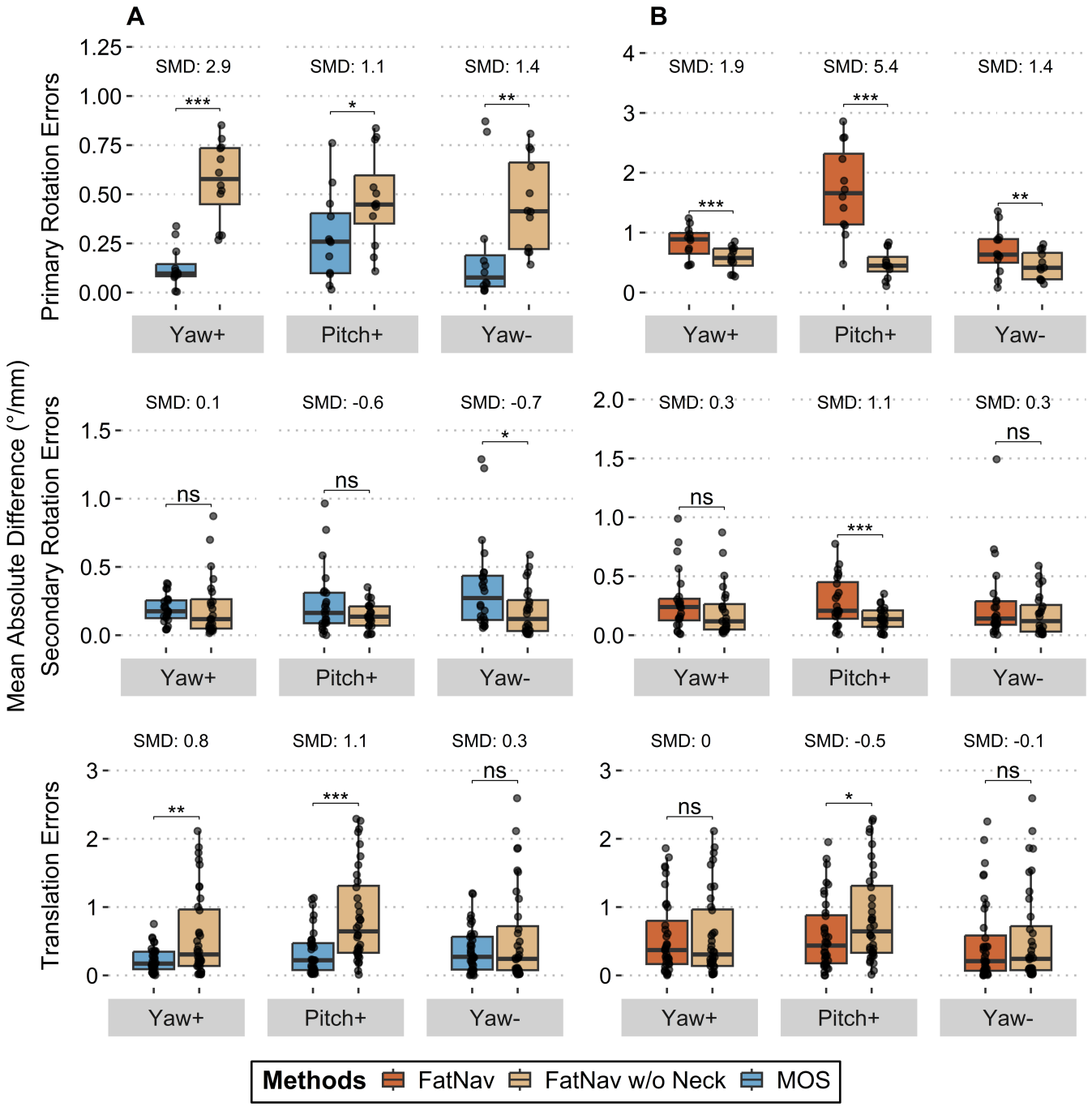


FIGURE 4 | Accuracy of head motion estimation for yaw$^-$, pitch$^+$, and yaw$^+$ head movements, with a separate analysis of large-amplitude rotations (2° and 4°) around the primary rotational axes on one side and smaller amplitude rotations around the secondary rotational axes on the other side. (A) MOS vs. FatNav with neck-mask and (B) FatNav with and without neck-mask. SMD: Standardized mean difference.

Interestingly, MOS's ability to estimate yaw$^+$ rotation was significantly more accurate than for yaw$^-$ ($p < 0.05$). Notably, we observed that FatNav errors increased with larger rotation amplitudes (Spearman correlation, $R = 0.70$; $p < 0.001$; see Figure 5), whereas MOS errors were amplitude-independent ($R = -0.06$).

The comparison of the two FatNav variants reveals that neck masking significantly reduced the mean errors of primary and secondary pitch$^+$ rotations by 48% and 32%, respectively (Figure 4B).

Measurement precision analyses for head rotations of 2° and 4° demonstrated that the average precision across the six subjects ranged between 0.16° and 0.36° for both MOS and neck-masked FatNav. FatNav demonstrated a significantly lower residual SD than MOS precision during pitch$^+$ rotations (0.24° vs. 0.36°, $p = 0.041$; Figure 6), indicating superior FatNav precision. These analyses also revealed asymmetric precision for MOS between rightward and leftward rotations (SD: 0.34° vs. 0.16°, $p < 0.001$; Figure 6), indicating lower MOS precision during rightward (yaw$^-$) rotations.

TABLE 2 | Reference standard values, along with the tracking accuracy (expressed as the intersubject mean error ± standard deviation) of MOS and FatNav for the primary yaw$^+$, pitch$^+$, and yaw$^-$ rotations.

Method			Yaw$^+$	Pitch$^+$	Yaw$^-$
Primary rotation	Reference standard	For 2°	1.98° ± 0.05°	1.88° ± 0.43°	2.09° ± 0.39°
		For 4°	4.08° ± 0.22°	4.01° ± 0.32°	4.13° ± 0.33°
	MOS	Error (°)	0.13 ± 0.10	0.28 ± 0.23	0.21 ± 0.31
	FatNav	Error (°)	0.57 ± 0.20	0.47 ± 0.24	0.45 ± 0.23
Secondary rotation	Reference standard	For 2°	0.4° ± 0.33°	0.33° ± 0.36°	0.49° ± 0.37°
		For 4°	0.62° ± 0.57°	0.43° ± 0.4°	0.66° ± 0.51°
	MOS	Error (°)	0.19 ± 0.1	0.23 ± 0.24	0.35 ± 0.33
	FatNav	Error (°)	0.21 ± 0.22	0.14 ± 0.09	0.18 ± 0.18
Translation	Reference standard	For 2°	0.65 ± 0.56 mm	0.59 ± 0.37 mm	0.6 ± 0.54 mm
		For 4°	1.03 ± 1.12 mm	0.94 ± 0.45 mm	1.04 ± 0.91 mm
	MOS	Error (mm)	0.22 ± 0.18	0.33 ± 0.33	0.37 ± 0.33
	FatNav	Error (mm)	0.58 ± 0.62	0.86 ± 0.68	0.55 ± 0.70

Note: Errors in estimating the associated translations and unintended rotations around the secondary axes are also presented (Experiment 1).

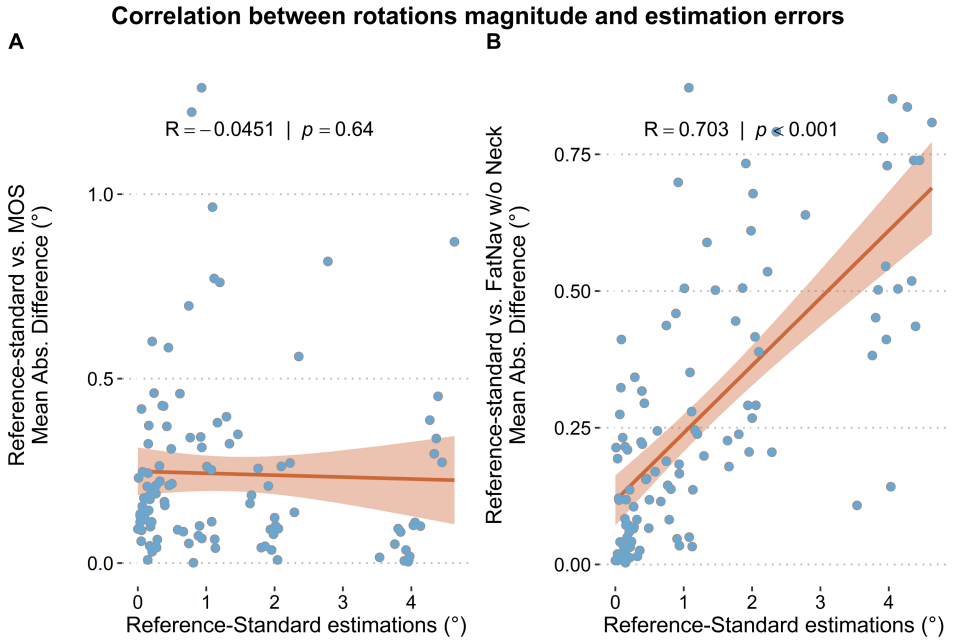


FIGURE 5 | Spearman's rank correlation between reference standard head rotation values and the estimation errors derived from MOS and FatNav with neck-mask.

3.3 | Image Distortion Effects on Head Motion Reference Standard

To evaluate distortion in T1w-images and its impact on reference standard values, we computed the proportion of brain voxels with a displacement norm greater than half the voxel size (0.5 mm) from elastic and affine registration fields. Table S2 confirms mild T1w-image distortions that vary with the orientation and magnitude of head displacement relative to the reference position. The impact of this distortion is minimal, as head motion estimates from both distortion-corrected and uncorrected reference standards are highly correlated (Figure S4). Furthermore, comparisons between MOS and FatNav remain largely unaffected when using distortion-corrected reference standard values (Figure S5).

3.4 | Image Quality Recovery by MOS and FatNav Corrections

Figure 7 illustrates the impact of 2° head rotations on image quality in Experiment 2, alongside the retrospective motion

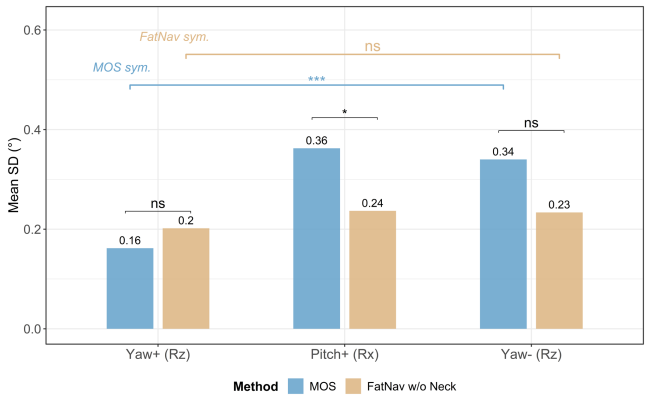


FIGURE 6 | Precision of rotation estimation using the markerless optical system (MOS) versus the fat navigator-based (FatNav) approach for 2° and 4° head rotations (Yaw$^+$, Pitch$^+$, and Yaw$^-$), including a comparison of directional precision (right vs. left movement) for both systems.

correction capabilities of MOS and neck-masked FatNav. For reference, the figure also displays head rotations and translations along the *X*-, *Y*-, and *Z*-axes, highlighting the discrepancies between MOS and FatNav measurements at each time point during a 6-min T1w-acquisition.

In blinded image quality evaluations, three evaluators assigned significantly higher scores to MOS-corrected images compared to FatNav corrected images (3.58 ± 1.13 and 2.52 ± 0.94, respectively; $p < 0.001$; Figure 8).

The analyses of image quality metrics further reinforced MOS's consistent superiority in restoring T1w-image fidelity (Figure 8). SSIM comparisons between motion-free and motion-corrected images revealed higher agreement for MOS (0.98 ± 0.008) than for FatNav (0.96 ± 0.007; $p < 0.001$). MOS also exhibited a 10.9% improvement in PSNR relative to FatNav (34.28 ± 1.61 dB vs. 32.08 ± 1 dB; $p < 0.001$), while focus measure analyses corroborated MOS's enhancement of image sharpness (196.8 ± 45.33 vs. 185.81 ± 40.22; $p = 0.009$).

Notably, three T1w-images were excluded from this analysis due to anomalous SSIM values, resulting from co-registration failures and thus a misalignment between reference images and motion-corrected images (Figure S6).

Finally, neither visual nor quantitative evaluations of image quality revealed statistically significant performance differences between the two FatNav variants with and without neck masking (Figure S7).

4 | Discussion

This study introduces an in vivo approach to evaluate the accuracy and the precision of intra-MRI head motion measurements and, consequently, the effectiveness of associated correction methodologies. The custom-developed software provided visual instruction with real-time feedback via MOS to standardize head positioning and ensure reproducible motion across participants. While Tracoline was utilized in this study to provide real-time feedback, the proposed methodology is fully compatible with any alternative real-time head tracking system. Unlike other recent visual instruction systems, where subjects were instructed to follow a cross moving across the screen with their nose [41], the real-time visualization of head position is more accurate and removes the need for any calibration procedure, as both the current and target head positions are presented at the same scale on the projection screen. The head movement guidance system exhibited low intersubject variability ($5.6\% \pm 4.3\%$ for 2° and $2.8\% \pm 2.1\%$ for 4°). Head translational shifts (up to 4 mm along the *X*-axis during yaw rotations) as well as subtle rotations around secondary nontargeted axes, consistent with natural head motion, were also observed and quantified (Tables 2 and S1).

Our framework combines a visual guidance system with a retrospective motion correction reconstruction pipeline, generating motion-free brain images across multiple head poses. Through corresponding head motion estimation and a dedicated statistical analysis component, the framework utilizes rigid and non\rigid registration to calculate interpose transformation matrices (the reference standard), thereby enabling the simultaneous quantification of motion tracking accuracy and precision across a sufficient number of subjects.

Involuntary head motion and/or magnetic susceptibility variations from head positioning can introduce artifacts in T1w-images, potentially compromising rigid registration and thus reference standard reliability; the next section examines these errors and their impact on reference standard validity.

During Experiment 1, volunteers were instructed to maintain their heads at the target position by monitoring the visual feedback system. This approach minimized subject movement during brain imaging at different head poses, thereby enhancing the robustness of the reference standard against small, involuntary motions. Furthermore, although the measured motion estimates were minimal, they were nonetheless retrospectively corrected during image reconstruction.

To evaluate potential bias from using MOS-corrected T1w-images in the comparison between MOS and FatNav, analyses were repeated using T1w-images reconstructed either without correction or with FatNav motion correction. Results demonstrated that performance discrepancies between MOS and FatNav remained consistent across all three conditions. The sole exception occurred during minor "pitch$^+$" movements, where FatNav's superior accuracy in estimating secondary rotation became statistically significant ($p < 0.05$) in both the uncorrected and FatNav-corrected datasets (Figure S8). This potentially suggests a minor degradation in MOS accuracy during subtle pitch$^+$ head rotations.

The distortion analyses of T1w-images previously presented demonstrate that reference standard estimates derived from both distortion-corrected and uncorrected T1w-images are highly correlated (Figure S4). Moreover, the comparison outcomes between MOS and FatNav remain largely consistent when using the distortion-corrected reference standard values (Figures 4 and S5). These results indicate that distortion effects are minimal within the experimental conditions of this study and that a reference standard based on rigid registration provides sufficiently

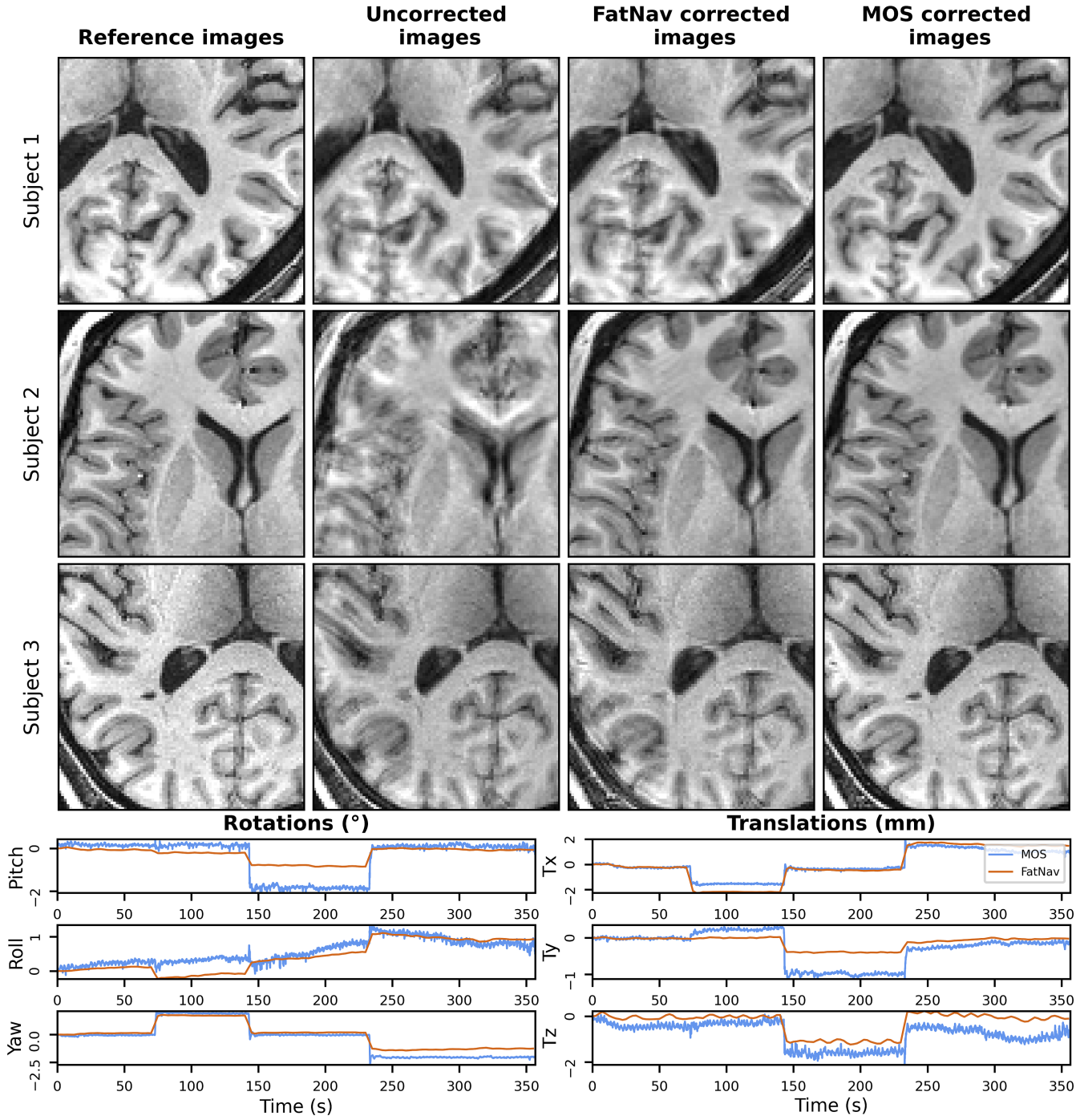


FIGURE 7 | Visual difference between brain T1w-images of 3 volunteers without and with 2° head rotations, and the same images retrospectively corrected from motion by MOS and by FatNav with neck-mask. The curves show MOS and FatNav estimates of rotations and translations in *X*, *Y*, and *Z*-axes executed by the volunteer during one of the T1w-scans. Curves were baseline-corrected to zero for clarity of visualization.

accurate estimations for head rotations in the range of 2°–4° at 3 T. Nonetheless, the use of distortion-corrected reference standard values may offer improved reliability in scenarios involving larger head movements and/or higher magnetic field strengths.

Notably, unlike MOS, FatNav accuracy depended on the amplitude of the head rotation (Figure 5). This finding is consistent with the simulation study by Andersen et al., which demonstrated that navigation errors generally increase with the magnitude of motion across various navigator types [42]. 3D FatNav-based motion tracking utilizes highly accelerated acquisitions of fat-signal reconstructed via GRAPPA. Since calibration data are typically acquired only at the scan's onset, a spatial mismatch can occur during significant head displacements. As demonstrated by Marchetto et al., this discrepancy between calibration and imaging data degrades fat navigator quality, leading to a progressive deterioration in motion estimation accuracy [43].

While our framework revealed that MOS generally provided superior accuracy to FatNav in estimating both translational and

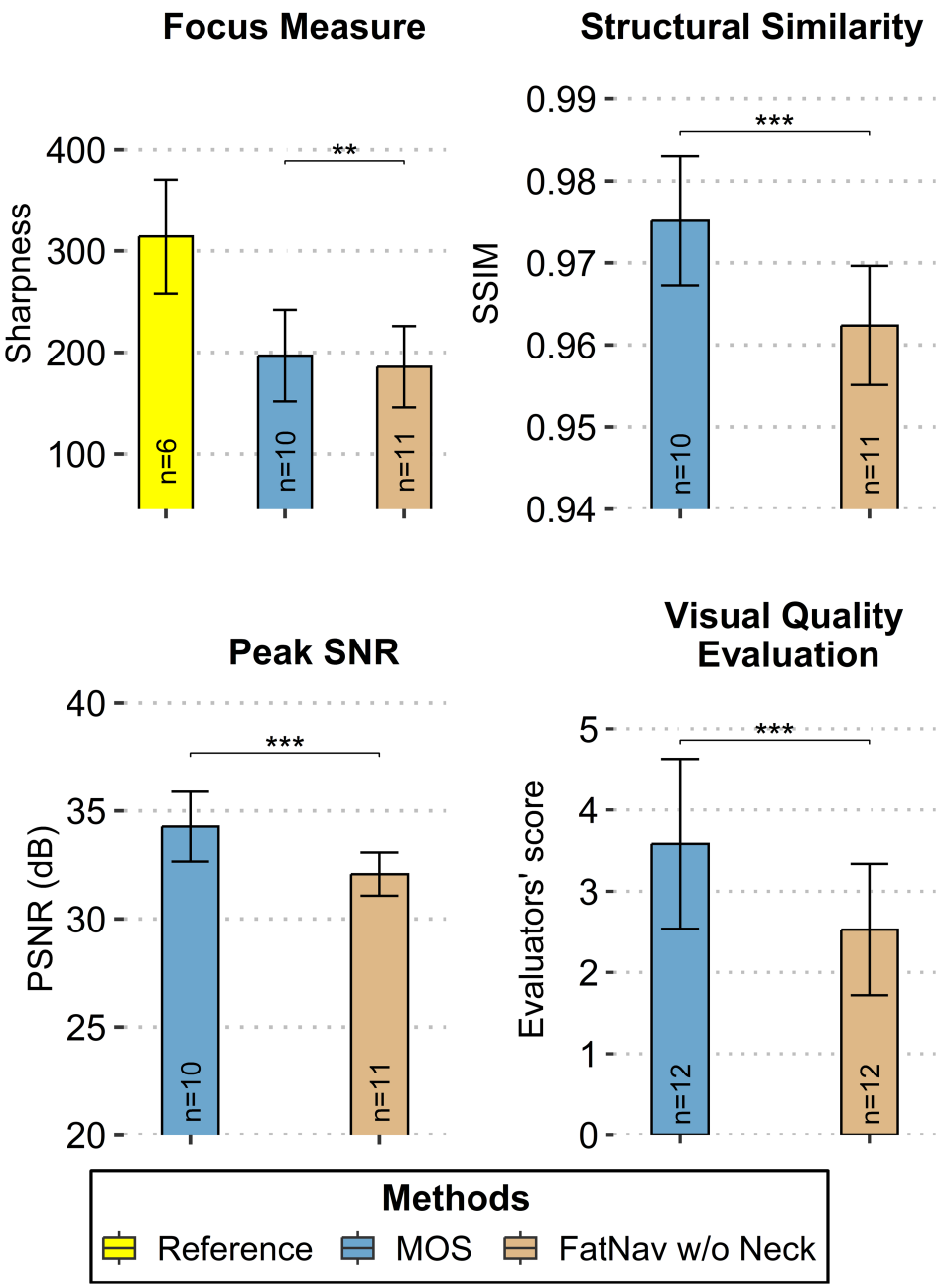


FIGURE 8 | Visual and quantitative quality assessment of T1w-images collected in Experiment 2 and retrospectively corrected for motion effects using MOS and FatNav with neck-mask.

large rotational motions, MOS appears less accurate when estimating yaw$^-$ rotations ($p < 0.05$), performing worse than FatNav for subtle movements along this specific axis.

Precision analyses for 2° and 4° head rotations showed that, across the six subjects, average values ranged from 0.16° to 0.36° for both MOS and FatNav. However, MOS exhibited lower precision than FatNav during pitch$^+$ rotations (Figure 6).

Beyond merely comparing head motion tracking approaches, the proposed methodology enables a comprehensive performance evaluation of each method across all spatial directions. Specifically, MOS appears to estimate yaw$^+$ rotations more accurately than yaw$^-$ rotations (Figure 4), with the lowest precision observed for yaw$^-$ rotations (Figure 6). This discrepancy could be attributable to facial features shifting in and out of the MOS field-of-view during motion. Consequently, the overlapping facial surface available for temporal alignment may vary depending on movement orientation, thereby impacting MOS tracking estimates. To the best of our knowledge, this study provides the first in vivo experimental characterization of such tracking dynamics at this level of detail.

Prior to this study, the metrological metrics of precision and accuracy for within-MRI head motion tracking have been evaluated primarily under idealized conditions. For instance, Maclaren et al. benchmarked tracking precision intra-bore during active scanning using a stationary water phantom,

yielding sub-millimeter temporal standard deviations (10 μm for translations and 0.01° for rotations). Conversely, accuracy was assessed ex-bore using a mechanical rotary table, reporting errors of < 90 μm for translations and < 0.06° for rotations [26]. By transitioning to a dynamic in vivo paradigm, our framework simultaneously characterizes both parameters under real clinical conditions. It encapsulates complex physiological and anatomical confounders, such as facial morphology, head size, and RF coil positional offsets, to deliver a highly realistic performance evaluation.

In Experiment 2, participants performed a step-wise head motion paradigm during the T1w-scan, transitioning between four positions similar to those in Experiment 1. Specifically, yaw$^+$, pitch$^+$, and yaw rotations were performed at 1 min 20 s, 2 min 30 s and 4 min after the start of the acquisition, respectively. This specific design was chosen to distribute the k-space sampling across distinct, stable poses, maximizing motion effects in the uncorrected data while preserving the central Fourier-space sampling at the pitch$^+$ position.

The temporal dynamics of this step-wise paradigm are particularly relevant when comparing tracking methodologies with differing temporal resolutions. While FatNavs excel at tracking slow drifts, they can struggle during rapid transitions due to their lower sampling rate, as shown by Gretsch et al. [23]. However, because the three position changes in our study were extremely brief compared to the extended stationary plateaus, the FatNav system had ample time to acquire uncorrupted fat navigator images for each pose. Consequently, FatNav was not disadvantaged by the choice of this movement paradigm.

The head rotation and displacement curves recorded during the 6 min of acquisition duration (Figure 7) reveal distinct discrepancies between MOS and FatNav estimates, which are consistent with the static estimation errors quantified in Experiment 1. Marchetto et al. documented similar discrepancies between MOS and FatNav estimates across both continuous and step-wise paradigms (specifically in their Figures 4 and 5) [22]. Because they lacked an in vivo reference standard, they could not determine the source of these variations. In contrast, the discrepancies observed in our protocol can be attributed predominantly to the modalities' intrinsic accuracy limits in static pose estimation and co-registration, rather than purely temporal constraints.

Quantitative and visual assessments of T1w-image quality confirmed that MOS was more effective in correcting head motion applied in this study (Figure 8). This finding is consistent with the results of Experiment 1, as MOS exhibits greater accuracy in estimating translational head displacements as well as high-amplitude rotations, which degrade perceived image quality more severely than smaller rotations. In our study, the selected image quality metrics, applied to the skull-stripped brain region after percentile normalization, were in good agreement with expert visual assessments.

Importantly, the head motion tracking evaluation approach we are proposing using a reference standard demonstrated sufficient sensitivity to detect improvement in FatNav estimates when a neck-mask was applied to the fat navigators. This improvement, undetectable by the image quality-based approach (Figure S7), was most pronounced for *X*-axis rotations (pitch$^+$), inducing elastic deformation at the head–neck junction (Figure 4B).

One limitation of this study is the extended acquisition time required for T1w-images at multiple head poses (4min-9sec × 7 ~ 29 min). To ensure accurate reference standard estimates via FreeSurfer registration, we prioritized high-quality 1-mm resolution images. However, in agreement with Gumus et al., who acquired 2-mm brain images every 26 s [20], downsampling our 1-mm T1w-images to 2 mm and repeating the registrations did not significantly alter reference standard motion estimates. This suggests that T1w-images with shorter acquisition times could maintain the reference standard's accuracy while increasing the sampling of head positions, reducing movements related to inattention, and mitigating magnetic susceptibility effects.

The in vivo framework presented in this study enables direct quantification of errors associated with motion tracking techniques but cannot capture all facets of motion-related image correction. Notably, it does not permit assessment of how correction frequency influences the quality of reconstructed images [18]. Although we interpolated motion parameters for each Fourier-space line using the nearest available head motion estimate, the higher temporal resolution of MOS (30 Hz vs. ~0.4 Hz for FatNav) may confer an advantage in correcting relatively fast motions. Furthermore, this in vivo approach cannot fully account for potential estimation errors arising during rapid, large-amplitude head movements, which may distort both the fat navigator images and the facial surface tracked by the MOS infrared camera. This limitation could be mitigated by reacquiring Fourier-space data affected by substantial motion during motion-free periods at the end of the acquisition [44, 45].

The proposed framework is thus well-suited for optimizing MRI motion tracking methods or conducting comparative performance benchmarks. However, it cannot independently predict final image quality, as degradation is fundamentally determined by the characteristics of uncorrected residual errors. A comprehensive evaluation of head motion correction strategies therefore requires benchmarking direct motion estimation accuracy and precision against a reliable in vivo reference standard, paired with complementary assessments of final image quality.

5 | Conclusion

This study introduces an in vivo framework combining visual motion instructions with a rigid registration-based reference standard to quantitatively evaluate and compare the accuracy and precision of intra-MRI head motion tracking techniques. Our approach demonstrated that the MOS offers superior overall performance for head translations and large rotations, as corroborated by image quality metrics. Notably, while FatNav achieved comparable accuracy to the MOS for small yaw$^+$ and pitch$^+$, it explicitly outperformed the MOS for small yaw$^-$ head rotations. This in vivo evaluation approach effectively assessed head motion tracking performance in each spatial direction and showed high sensitivity to the subtle improvements provided by

neck masking in FatNav data. These findings underscore the utility of this framework for advancing and refining head motion tracking and correction strategies.

Author Contributions

Z. Zariry contributed to the methodology, development of the reconstruction script, data acquisition, data processing, statistical analysis, and drafting of the manuscript. F. Lamberton contributed to data acquisition and reviewed the manuscript. R. Frost and A. van der Kouwe contributed to the study conception, data processing design, and reviewed and edited the manuscript. T. Gaass and J. M. Slipsager provided technical support and reviewed the manuscript. T. Troalen provided technical support and MRI platform implementation. H. Rayson provided statistical analysis support and reviewed and edited the manuscript. N. Richard provided image processing support and reviewed the manuscript. J. Bonaiuto contributed to statistical analysis, funding acquisition, study conception, and reviewed the manuscript. B. Hiba contributed to the study conception, methodology, drafting of the manuscript, project supervision, and funding acquisition.

Acknowledgments

This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (ERC Consolidator Grant 864550), LABEX CORTEX funding (ANR-11-LABX-0042) from the Université de Lyon operated by the French National Research Agency (ANR), the National Institute of Biomedical Imaging and Bioengineering (R21EB029641 and R01EB035560), and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01HD110152).

Conflicts of Interest

T.G. and J.M.S. are TracInnovations employees, and T.T. is an employee of Siemens Healthineers. All other authors declare having no competing financial and/or nonfinancial interests.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The retrospective head motion correction pipeline is publicly available on GitHub [pyMocoKit]. The remaining code (real-time head motion guidance and statistical analysis scripts) will be progressively added to the following dedicated public repositories (https://github.com/zzariry/moco_eval_framework, https://github.com/zzariry/MRImove). In the meantime, these codes can be requested from the corresponding author.

References

1. E. L. Gedamu and A. Gedamu, "Subject Movement During Multislice Interleaved MR Acquisitions: Prevalence and Potential Effect on MRI-Derived Brain Pathology Measurements and Multicenter Clinical Trials of Therapeutics for Multiple Sclerosis," *Journal of Magnetic Resonance Imaging* 36, no. 2 (2012): 332–343, https://doi.org/10.1002/jmri.23666.

2. J. P. Sun, C. X. Bu, J. H. Dang, et al., "Deep Learning-Based Reconstruction Enhanced Image Quality and Lesion Detection of White Matter Hyperintensity Through in FLAIR MRI," *Asian Journal of Surgery* 48, no. 1 (2025): 342–349, https://doi.org/10.1016/j.asjsur.2024.09.156.

3. J. M. Slipsager, S. L. Glimberg, J. Søgaard, et al., "Quantifying the Financial Savings of Motion Correction in Brain MRI: A Model-Based Estimate of the Costs Arising From Patient Head Motion and Potential Savings From Implementation of Motion Correction," *Journal of Magnetic Resonance Imaging* 52, no. 3 (2020): 731–738, https://doi.org/10.1002/jmri.27112.

4. J. B. Andre, B. W. Bresnahan, M. Mossa-Basha, et al., "Toward Quantifying the Prevalence, Severity, and Cost Associated With Patient Motion During Clinical MR Examinations," *Journal of the American College of Radiology* 12, no. 7 (2015): 689–695, https://doi.org/10.1016/j.jacr.2015.03.007.

5. M. F. Callaghan, O. Josephs, M. Herbst, M. Zaitsev, N. Todd, and N. Weiskopf, "An Evaluation of Prospective Motion Correction (PMC) for High Resolution Quantitative MRI," *Frontiers in Neuroscience* 9 (2015): 9, https://doi.org/10.3389/fnins.2015.00097.

6. X. Zong, S. Nanavati, S. Hung, T. Li, and W. Lin, "Effects of Motion and Retrospective Motion Correction on the Visualization and Quantification of Perivascular Spaces in Ultrahigh Resolution T2-Weighted Images at 7T," *Magnetic Resonance in Medicine* 86, no. 4 (2021): 1944–1955, https://doi.org/10.1002/mrm.28847.

7. B. Gal-Er, Y. Brackenier, A. F. Bonthrone, et al., "Verifying the Concordance Between Motion Corrected and Conventional MPRAGE for Pediatric Morphometric Analysis," *Frontiers in Neuroscience* 19 (2025): 1534924, https://doi.org/10.3389/fnins.2025.1534924.

8. M. Engström, M. Mårtensson, E. Avventi, O. Norbeck, and S. Skare, "Collapsed Fat Navigators for Brain 3D Rigid Body Motion," *Magnetic Resonance Imaging* 33, no. 8 (2015): 984–991, https://doi.org/10.1016/j.mri.2015.06.014.

9. D. Gallichan, J. P. Marques, and R. Gruetter, "Retrospective Correction of Involuntary Microscopic Head Movement Using Highly Accelerated Fat Image Navigators (3D FatNavs) at 7T," *Magnetic Resonance in Medicine* 75, no. 3 (2016): 1030–1039, https://doi.org/10.1002/mrm.25670.

10. M. Laustsen, M. Andersen, R. Xue, K. H. Madsen, and L. G. Hanson, "Tracking of Rigid Head Motion During MRI Using an EEG System," *Magnetic Resonance in Medicine* 88, no. 2 (2022): 986–1001, https://doi.org/10.1002/mrm.29251.

11. O. Afacan, T. E. Wallace, and S. K. Warfield, "Retrospective Correction of Head Motion Using Measurements From an Electromagnetic Tracker," *Magnetic Resonance in Medicine* 83, no. 2 (2020): 427–437, https://doi.org/10.1002/mrm.27934.

12. M. Zaitsev, C. Dold, G. Sakas, J. Hennig, and O. Speck, "Magnetic Resonance Imaging of Freely Moving Objects: Prospective Real-Time Motion Correction Using an External Optical Motion Tracking System," *NeuroImage* 31, no. 3 (2006): 1038–1050, https://doi.org/10.1016/j.neuroimage.2006.01.039.

13. J. Maclaren, B. S. R. Armstrong, R. T. Barrows, et al., "Measurement and Correction of Microscopic Head Motion During Magnetic Resonance Imaging of the Brain," *PLoS One* 7, no. 11 (2012): e48088, https://doi.org/10.1371/journal.pone.0048088.

14. L. Cordero-Grande, G. Ferrazzi, R. P. A. G. Teixeira, J. O'Muircheartaigh, A. N. Price, and J. V. Hajnal, "Motion-Corrected MRI With DISORDER: Distributed and Incoherent Sample Orders for Reconstruction Deblurring Using Encoding Redundancy," *Magnetic Resonance in Medicine* 84, no. 2 (2020): 713–726, https://doi.org/10.1002/mrm.28157.

15. N. Yoshida, H. Kageyama, H. Akai, et al., "Motion Correction in MR Image for Analysis of VSRAD Using Generative Adversarial Network," *PLoS One* 17, no. 9 (2022): e0274576, https://doi.org/10.1371/journal.pone.0274576.

16. S. Chatterjee, A. Sciarra, M. Dünnwald, S. Oeltze-Jafra, A. Nürnberger, and O. Speck, *Retrospective Motion Correction of MR Images Using Prior-Assisted Deep Learning* (arXiv, 2020): arXiv:2011.14134, https://doi.org/10.48550/arXiv.2011.14134.

17. J. Maclaren, M. Herbst, O. Speck, and M. Zaitsev, "Prospective Motion Correction in Brain Imaging: A Review," *Magnetic Resonance in Medicine* 69, no. 3 (2013): 621–636, https://doi.org/10.1002/mrm.24314.

18. R. Frost, P. Wighton, F. I. Karahanoğlu, et al., "Markerless High-Frequency Prospective Motion Correction for Neuroanatomical MRI,"

Magnetic Resonance in Medicine 82, no. 1 (2019): 126–144, https://doi.org/10.1002/mrm.27705.

19. J. M. Slipsager, S. L. Glimberg, L. Højgaard, et al., "Comparison of Prospective and Retrospective Motion Correction in 3D-Encoded Neuroanatomical MRI," *Magnetic Resonance in Medicine* 87, no. 2 (2022): 629–645, https://doi.org/10.1002/mrm.28991.

20. K. Gumus, B. Keating, N. White, et al., "Comparison of Optical and MR-Based Tracking," *Magnetic Resonance in Medicine* 74, no. 3 (2015): 894–902, https://doi.org/10.1002/mrm.25472.

21. P. L. Bazin, H. E. Nijsse, W. Van Der Zwaag, et al., "Sharpness in Motion Corrected Quantitative Imaging at 7T," *NeuroImage* 222 (2020): 117227, https://doi.org/10.1016/j.neuroimage.2020.117227.

22. E. Marchetto, K. Murphy, S. L. Glimberg, and D. Gallichan, "Robust Retrospective Motion Correction of Head Motion Using Navigator-Based and Markerless Motion Tracking Techniques," *Magnetic Resonance in Medicine* 90, no. 4 (2023): 1297–1315, https://doi.org/10.1002/mrm.29705.

23. F. Gretsch, H. Mattern, D. Gallichan, and O. Speck, "Fat Navigators and Moiré Phase Tracking Comparison for Motion Estimation and Retrospective Correction," *Magnetic Resonance in Medicine* 83, no. 1 (2020): 83–93, https://doi.org/10.1002/mrm.27908.

24. M. Reuter, M. D. Tisdall, A. Qureshi, R. L. Buckner, A. J. W. Van Der Kouwe, and B. Fischl, "Head Motion During MRI Acquisition Reduces Gray Matter Volume and Thickness Estimates," *NeuroImage* 107 (2015): 107–115, https://doi.org/10.1016/j.neuroimage.2014.12.006.

25. J. Maclaren, O. Speck, D. Stucht, P. Schulze, J. Hennig, and M. Zaitsev, "Navigator Accuracy Requirements for Prospective Motion Correction," *Magnetic Resonance in Medicine* 63, no. 1 (2010): 162–170, https://doi.org/10.1002/mrm.22191.

26. J. Maclaren, B. S. R. Armstrong, R. T. Barrows, et al., "Measurement and Correction of Microscopic Head Motion during Magnetic Resonance Imaging of the Brain," *PLoS One* 7, no. 11 (2012): e48088, https://doi.org/10.1371/journal.pone.0048088.

27. E. Avventi, H. Ryden, O. Norbeck, J. Berglund, T. Sprenger, and S. Skare, "Projection-Based 3D/2D Registration for Prospective Motion Correction," *Magnetic Resonance in Medicine* 84, no. 3 (2020): 1534–1542, https://doi.org/10.1002/mrm.28225.

28. TracInnovations, *Motion Tracking* (TracInnovations), accessed January 26, 2025, https://tracinnovations.com/motion-tracking/.

29. J. M. Slipsager, A. H. Ellegaard, S. L. Glimberg, et al., "Markerless Motion Tracking and Correction for PET, MRI, and Simultaneous PET/MRI," *PLoS One* 14, no. 4 (2019): e0215524, https://doi.org/10.1371/journal.pone.0215524.

30. D. Gallichan, *retroMoCoBox/fatnavtools at Master dgallichan/retroMoCoBox* (GitHub, 2026), https://github.com/dgallichan/retroMoCoBox/tree/master/fatnavtools.

31. M. A. Griswold, P. M. Jakob, R. M. Heidemann, et al., "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)," *Magnetic Resonance in Medicine* 47, no. 6 (2002): 1202–1210, https://doi.org/10.1002/mrm.10171.

32. M. Reuter, H. D. Rosas, and B. Fischl, "Highly Accurate Inverse Consistent Registration: A Robust Approach," *NeuroImage* 53, no. 4 (2010): 1181–1196, https://doi.org/10.1016/j.neuroimage.2010.07.020.

33. B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration," *NeuroImage* 54, no. 3 (2011): 2033–2044, https://doi.org/10.1016/j.neuroimage.2010.09.025.

34. S. M. Smith, "Fast Robust Automated Brain Extraction," *Human Brain Mapping* 17, no. 3 (2002): 143–155, https://doi.org/10.1002/hbm.10062.

35. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing* 13, no. 4 (2004): 600–612, https://doi.org/10.1109/TIP.2003.819861.

36. S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of Focus Measure Operators for Shape-From-Focus," *Pattern Recognition* 46, no. 5 (2013): 1415–1432, https://doi.org/10.1016/j.patcog.2012.11.011.

37. R Foundation for Statistical Computing, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024), https://www.R-project.org/.

38. D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software* 67 (2015): 1–48, https://doi.org/10.18637/jss.v067.i01.

39. J. Fox and S. Weisberg, *An R Companion to Applied Regression*, Third ed. (Sage, 2019), https://www.john-fox.ca/Companion/.

40. R. V. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means* (R Foundation, 2017), https://doi.org/10.32614/CRAN.package.emmeans.

41. J. Berglund, A. Van Niekerk, H. Rydén, et al., "Prospective Motion Correction for Diffusion Weighted EPI of the Brain Using an Optical Markerless Tracker," *Magnetic Resonance in Medicine* 85, no. 3 (2021): 1427–1440, https://doi.org/10.1002/mrm.28524.

42. M. Andersen, M. Laustsen, and V. Boer, "Accuracy Investigations for Volumetric Head-Motion Navigators With and Without EPI at 7 T," *Magnetic Resonance in Medicine* 88, no. 3 (2022): 1198–1211, https://doi.org/10.1002/mrm.29296.

43. E. Marchetto and D. Gallichan, "Analysis of the Effect of Motion on Highly Accelerated 3D FatNavs in 3D Brain Images Acquired at 3T," *PLoS One* 19, no. 7 (2024): e0306078, https://doi.org/10.1371/journal.pone.0306078.

44. M. D. Tisdall, A. T. Hess, M. Reuter, E. M. Meintjes, B. Fischl, and A. J. W. Van Der Kouwe, "Volumetric Navigators for Prospective Motion Correction and Selective Reacquisition in Neuroanatomical MRI," *Magnetic Resonance in Medicine* 68, no. 2 (2012): 389–399, https://doi.org/10.1002/mrm.23228.

45. R. Frost, A. T. Hess, T. W. Okell, et al., "Prospective Motion Correction and Selective Reacquisition Using Volumetric Navigators for Vessel-Encoded Arterial Spin Labeling Dynamic Angiography," *Magnetic Resonance in Medicine* 76, no. 5 (2016): 1420–1430, https://doi.org/10.1002/mrm.26040.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** (A) Point clouds of the facial surface captured by the markerless optical system through the coil's aperture in front of the right eye. (B) Cross-calibration procedure: Alignment of the point cloud derived from the 17-s image (blue surface) with the point cloud captured by the MOS camera during the scan (grey surface). **Figure S2:** (A) Raw FatNav image including both the brain and neck regions. (B) FatNav image with the neck region masked. **Figure S3:** Primary and secondary axes of head rotations: The primary rightward (yaw–) and leftward (yaw+) rotations around Z-axis are associated with unintentional small secondary rotations around the Y- and X-axes, and the primary upward (pitch+) rotation around the X-axis is associated with unintentional small secondary rotations around the Y- and Z-axes. **Figure S4:** Comparison of head motion parameters of Experiment 1 derived from the rigid registration-based reference standard and from the reference standard with distortion-corrected T1w-images. **Figure S5:** Accuracy of head motion estimation by MOS and neck-masked FatNav in Experiment 1. Reference standard values were obtained from T1w-images: (A) without distortion correction and (B) with distortion correction. SMD: Standardized mean difference. **Figure S6:** Example of a motion-corrected T1-weighted image excluded from the quantitative analysis due to a co-registration failure. Despite visually successful MOS-based motion correction of the 2° rotation, inaccurate 3D co-registration to the reference image artificially lowered

the computed structural similarity index measure (SSIM ~0.86, quantifying resemblance to the reference image), compared to the uncorrected version of the same image (SSIM ~0.96). Visual inspection confirms the superior quality of the motion-corrected image, illustrating that the exclusion of this dataset from the quantitative analysis was driven by a postprocessing co-registration artifact rather than by poor motion-correction quality. **Figure S7:** Quality assessment of T1-weighted images acquired in Experiment 2 and retrospectively corrected for motion effects using standard FatNav and FatNav with neck-mask. **Figure S8:** Impact of involuntary head motion during T1w-images acquisitions on the reference standard-based evaluation of MOS and FatNav errors. T1w-images of Experiment 1 were reconstructed under three conditions: without motion correction, with MOS motion correction, and with FatNav motion correction. **Table S1:** Tracking accuracy (expressed as the intersubject mean error ± standard deviation) of head translations and minor involuntary rotations, assessed along the X-, Y-, and Z-axes, for all yaw+, pitch+, and yaw− movements of 2° and of 4°. **Table S2:** Mean number and ratio of brain voxels with distortion vector norms greater than 0.5 mm in T1-weighted images from Experiment 1. The distortion field was derived from elastic and affine registrations between each T1w-image and its corresponding reference.