PCA regression for continuous estimation of head pose in PET/MR

Ashley G. Gillman *Student Member, IEEE*, Alaleh Rashidnasab, Richard Brown, Nicholas Dowson, Benjamin Thomas, Francesco Fraioli, Stephen Rose, and Kris Thielemans, *Senior Member, IEEE*,

***Abstract*—With the availability of improved hardware and local point-spread function modelling, the presence of patient motion has become a major barrier to further improvements in the quality of PET images and their clinical efficacy. Although numerous approaches to compensate for patient motion have been proposed and are even commercially available, the additional hardware and extended setup time can preclude their routine clinical use. The MR modality on combined PET and MR (PET/MR) scanners can be used to correct motion with almost no additional setup time but currently must replace other MR acquisitions that may be required for clinic use. To overcome these problems, principal component analysis (PCA) and other data-driven techniques have been demonstrated to be able to reliably provide a signal related to patient motion based on raw PET data. Typically, these signals are used to split the PET acquisition into a discrete set of approximately motion-free time segments. This work introduces an approach where the PCA-signals are used as direct surrogates for the motion and regressed against rigid head motion parameters, enabling continuous pose estimation. A proof-of-concept is presented in which the approach is applied to upsample a low temporal resolution MR motion estimate. This proof-of-concept uses rapid echo planar imaging (EPI) data together with PET-derived motion signals. In a comparison of four techniques, nearest neighbour (NN) and linear temporal interpolation and linear and radial basis function (RBF) regression of pose against the PCA surrogate, we demonstrate that the model can be used to accurately interpolate pose continuously throughout the scan.**

***Index Terms*—PET, Motion Correction, PET/MR, Motion Tracking, Data-driven, PCA**

The availability of improved PET hardware and software, has led to the ability to pursue increasingly quantitative accuracy in PET. Motion correction has been shown to improve quantitation, resulting in demonstrable change in patient management [1]. MR-based motion correction in PET/MR are amongst the most promising methods available. However, only one MR sequence can be acquired at a time and typical MR protocols consist of a consecutive set of diagnostic sequences. If diagnostic sequences are not acquired simultaneously with the PET acquisition, they will increase the scan time and cost.

Data-driven PET tracking techniques are a class of methods for tracking the progression of motion directly from PET raw data. The tracking signal produced is known as a motion surrogate, $\boldsymbol{\tau}(t)$. The *sensitivity* technique [2] monitors movement of the emission distribution with respect to the stationary sensitivity profile of the scanner through the instantaneous count rate. Moment-based techniques monitor the central tendency of the distribution (and its moments) [3, 4]. The *PCA* technique [5] monitors changes in the distribution of counts in sinogram-space over time, and attributes changes to motion. However, for head motion these techniques have only been used to detect when (fast) movement occurs, and split the data into time frames as input to the multiple acquisition frame method [6]. This approach is not suitable for slow head motion and implies some remaining intra-frame motion.

The contribution of this work is to provide a data-driven, high temporal resolution pose estimate by regressing pose against the motion surrogate. Improved temporal resolution has been demonstrated to correlate with enhanced quality of motion correction [7], and is therefore a desirable property. Similar approaches have been applied in respiratory motion, modelling voxel-wise displacements [8], however the application to head motion is novel.

I. Methodology and materials

Two subjects suffering epileptic seizures underwent a 45 minute FDG PET scan [9]. Interspersed through each were four arterial spin labelling (ASL) scans, each 101 frames with TR $= 2860$ ms at a voxel size of $3.56 \times 3.56 \times 6$ mm. A magnetisation-prepared rapid acquisition with gradient echo (MPRAGE) image was acquired for anatomy, at a voxel size of $0.52 \times 0.52 \times 1.10$ mm. Head pose estimates were derived by registration of the ASL series to the first ASL frame in each session, using NiftyReg's `reg_aladin` [10]. Each transform matrix was converted into a signal of 3 translations and 3 Euler angles of rotation for the pose estimate, $\boldsymbol{\rho} \in \mathbb{R}^6$.

PCA is used for motion surrogate extraction, as described by Thielemans et al. [5], except, 5 eigenvectors are retained during dimensionality reduction in order to obtain $\boldsymbol{\tau}(t) \in \mathbb{R}^5$. The intent is for the regression to learn which of the transformed components are representative of motion. Two methods are investigated for finding $f : \boldsymbol{\tau} \to \boldsymbol{\rho}$. The first is linear regression, $f_{\text{linear}}(\boldsymbol{\tau})$. The second technique relies only the fact that that time points with similar-valued surrogates

Manuscript received January 7, 2020. This work was supported by the CCP PETMR exchange grant, funded by EPSRC (grant EP/M022587/1) and the NIHR UCLH Biomedical Research Centre. The work of A. G. Gillman was supported by an Australian Government Department of Education and Training Research Training Program stipend and a Commonwealth Scientific and Industrial Research Organisation (CSIRO) Top Up Scholarship. *(Corresponding author: A. G. Gillman.)*

A. G. Gillman and S. Rose are with the Australian e-Health Research Centre, CSIRO, Brisbane, Australia, and also with with the University of Queensland, Brisbane, Australia (e-mail: ashley.gillman@csiro.au).

A. Rashidnasab was with University College London Institute of Nuclear Medicine, London, UK. She is now with Dunnhumby, London, UK.

R. Brown, B. Thomas, F. Fraioli and K. Thielemans are with University College London Institute of Nuclear Medicine, London, UK.

N. Dowson is with Optellum, Oxford, UK.

should correspond to a similar-valued pose. Gaussian RBF interpolation imposes few assumptions on the relationship between the endogeneous and exogeneous variables [11]. The mapping from surrogate to pose,

$$f_{\text{RBF}}(\boldsymbol{\tau}) := \sum_{j=1}^{n} \lambda_j K(\boldsymbol{\tau}, \boldsymbol{\tau}_j) \tag{1}$$

Briefly, a kernel is defined between a new surrogate observation, $\boldsymbol{\tau}$, and the j^{th} training example, $\boldsymbol{\tau}_j$, as a Gaussian RBF, $K(\boldsymbol{\tau}, \boldsymbol{\tau}_j) = \exp\left(-\frac{||\boldsymbol{\tau}-\boldsymbol{\tau}_j||^2}{2\sigma^2}\right)$, where σ is set to the mean distance between neighbours in $\boldsymbol{\tau}$. An $n \times n$ kernel matrix, A (where n is the number of training points), is subsequently constructed, defined as $A_{i,j} = K(\boldsymbol{\tau}_i, \boldsymbol{\tau}_j)$. Finally, λ is defined by $\lambda_j = A^{-1}\boldsymbol{\rho}_j$.

We demonstrate the above technique in a proof-of-concept, upsampling a low temporal resolution pose estimate. The ASL pose estimate is downsampled to simulate low resolution training data, by factors $d = 10, 11, ..., 29, 30$ ($d = 20$ is a sampling rate of 57.2 s), to avoid unrepresentative results where sampling aligns with motion events. The fully sampled $\hat{\boldsymbol{\rho}}_{\text{MR}}$ was the gold-standard estimate. We compared against two simple temporal interpolation techniques: NN linear interpolation. These are not aware of the motion surrogate.

For error measurement, the MPRAGE was segmented with FreeSurfer 6.0. [12]. The centre-of-mass of each region was used to form a point cloud. The error is assessed on point-wise displacement for each point between the truth and estimated transforms, aggregated by mean-of-maximum displacement (MMD) given in Eq. (2), where P_i is the i^{th} of I points in the point cloud, and $\hat{X}_t$ and X_t are the estimated and truth affine transforms at the t^{th} of T time points. This choice of error aggregation is motivated by the fact that localisation of misalignment is unimportant, and that the spatial maximum will account for the observed artefact in a given time frame.

$$\epsilon_{\text{MMD}} = \frac{1}{|T|} \sum_{t \in T} \max_{i \in I} \|P_i X_t - P_i \hat{X}_t\|^2 \tag{2}$$

II. Results

Of the eight ASL acquisitions, two captured a motion event in which regions of the brain were displaced > 1 mm, dichotomised as "High Motion". The aggregated MMD results are given in Fig. 1. Additionally, Table I includes the results of a Wilcoxon signed-rank test between the distributions in Fig. 1. A Wilcoxon signed-rank test was used for significance testing as the distributions were not normally distributed. RBF-PCA performed best in combined and high motion cases. RBF-PCA performed worse in the low motion case, but with a negligible effect size (< 0.1 mm).

III. Discussion

RBF-PCA consistently provides an improvement in the motion estimate in high-motion cases. Both PCA regression techniques perform worse than temporal interpolation in low-motion cases. However, the effect size for linear-PCA is approximately 0.3 mm, less than a typical PET reconstruction resolution, and for RBF-PCA is less than 0.1 mm.

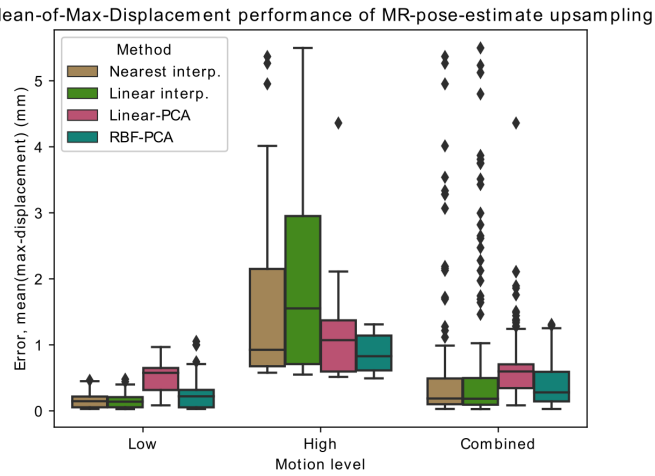

Fig. 1. The distribution of aggregated MMD error (y-axis) for each upsampling method (hue). Results are grouped for scans in which, subjectively, a low or high level of motion occurred, and the Combined results include both.

TABLE I
RESULTS OF A PAIRED WILCOXON SIGNED-RANK TEST ON DISTRIBUTIONS IN FIG. 1. "$\sim$" NO SIG.; "$<$" AND "$>$" LESS THAN OR GREATER THAN, $p < 0.05$; "$\ll$" AND "$\gg$" LESS THAN OR GREATER THAN $p < 0.005$. LOWEST-ERROR FOR EACH MOTION LEVEL IS EMBOLDENED.

Motion	Reference	MMD	Sig.	MMD	Test Method
Low	NN interp.	0.2	$\ll$	0.5	Linear PCA
	NN interp.	0.2	$\ll$	0.2	RBF PCA
	Linear interp.	0.2	$\ll$	0.5	Linear PCA
	Linear interp.	0.2	$\ll$	0.2	RBF PCA
High	NN interp.	1.6	$\sim$	1.1	Linear PCA
	NN interp.	1.6	$\gg$	0.9	**RBF PCA**
	Linear interp.	2.0	$>$	1.1	Linear PCA
	Linear interp.	2.0	$\gg$	0.9	**RBF PCA**
Combined	NN interp.	0.5	$\ll$	0.7	Linear PCA
	NN interp.	0.5	$\gg$	0.4	**RBF PCA**
	Linear interp.	0.6	$\ll$	0.7	Linear PCA
	Linear interp.	0.6	$\gg$	0.4	**RBF PCA**

Linear-PCA was observed to extrapolate pose estimates beyond the range of the training data, unlike RBF-PCA and reference interpolant methods. In some cases this provided accurate estimation, but in others led to gross exaggeration and misestimation. In contrast, RBF-PCA provided a conservative estimate that reduced error compared to interpolant methods, and avoided gross misestimation.

A likely future clinical application of this technique would include short, purpose-specific pose estimation acquisitions, likely EPI series, interspersed through the acquisition.

IV. Conclusion

This work demonstrated that subject pose can be learned from a PCA motion surrogate and used to produce a temporally upsampled pose estimate. Two regression techniques were investigated to learn the relationship between motion surrogate and pose, linear regression and RBF regression. Results were compared to temporal interpolation, blind to the motion surrogate. It was found that RBF regression was able to provide the most accurate results, improving upon temporal regression techniques. However, in cases where the subject moved less that 1 mm, the technique provided slightly worse

results, but with a small mean effect size of < 0.1 mm. This work demonstrates that, at least in some cases, motion can be learned directly from a motion surrogate. The technique is not specific to PET/MR, and pose estimates for training can potentially be obtained from motion-specific-MR, other diagnostic MR, non-attenuation corrected PET reconstruction frames, or a combination of these.

REFERENCES

[1] Richard Manber, Kris Thielemans, Brian F. Hutton, et al. "Clinical Impact of Respiratory Motion Correction in Simultaneous PET/MR, Using a Joint PET/MR Predictive Motion Model". In: *Journal of Nuclear Medicine* 59 (Sept. 2018), pp. 1467–1473.

[2] D. Visvikis, O. Barret, T. Fryer, et al. "A posteriori respiratory motion gating of dynamic PET images". In: *Nuclear Science Symposium and Medical Imaging Conference*. Vol. 5. Oct. 2003, 3276–3280 Vol.5.

[3] Chuan Huang, Yoann Petibon, Marc Normandin, et al. "Fast head motion detection using PET list-mode data". In: *Journal of Nuclear Medicine* 56 (May 2015), pp. 1827–1827.

[4] P. J. Schleyer, J. T. Dunn, S. Reeves, et al. "Detecting and estimating head motion in brain PET acquisitions using raw time-of-flight PET data". In: *Physics in Medicine and Biology* 60 (Aug. 2015), pp. 6441–6458.

[5] K. Thielemans, P. Schleyer, J. Dunn, et al. "Using PCA to detect head motion from PET list mode data". In: *Nuclear Science Symposium and Medical Imaging Conference*. Oct. 2013, pp. 1–5.

[6] Y. Picard and C. J. Thompson. "Motion correction of PET images using multiple acquisition frames". In: *IEEE Transactions on Medical Imaging* 16 (Apr. 1997), pp. 137–144.

[7] Andrew J. Montgomery, Kris Thielemans, Mitul A. Mehta, et al. "Correction of Head Movement on PET Studies: Comparison of Methods". In: *Journal of Nuclear Medicine* 47 (Dec. 2006), pp. 1936–1944.

[8] Chi Liu, Adam M. Alessio, and Paul E. Kinahan. "Respiratory Motion Correction for Quantitative PET/CT Using All Detected Events with Internal—External Motion Correlation". In: *Medical Physics* 38 (2011), pp. 2715+.

[9] Alaleh Rashidnasab, Benjamin A Thomas, Richard Manber, et al. "Hybrid Head Motion Correction in PET/MR Brain Imaging". In: *6th conference on PET-MRI and SPECT-MRI (PSMR)*. Lisboa, Portugal, May 2017.

[10] Marc Modat, David M. Cash, Pankaj Daga, et al. "Global image registration using a symmetric block-matching approach". In: *Journal of Medical Imaging (Bellingham, Wash.)* 1 (July 2014), p. 024003.

[11] D. S. Broomhead and David Lowe. *Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks*. Malvern, UK: Royal signals and radar establishment, Mar. 1988.

[12] Bruce Fischl, David H. Salat, Evelina Busa, et al. "Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain". In: *Neuron* 33 (Jan. 2002), pp. 341–355.