Frequency Diffeomorphisms for Efficient Image Registration

Miaomiao Zhang, Ruizhi Liao, Adrian V. Dalca,
Esra A. Turk, Jie Luo, P. Ellen Grant, and Polina Golland

Presented by Rishi Edupalli

October 14, 2026

Outline

1. Background: LDDMM and FLASH
2. Frequency diffeomorphisms
3. Results and conclusions

Background: LDDMM

Diff(Ω): smooth, invertible maps of the image domain $\Omega \subset \mathbb{R}^d$; e: the identity map. Distance:

$$\operatorname{dist}(e, \phi) = \int_0^1 \langle \mathcal{L} v_t,\ v_t \rangle \, dt.$$

The path of maps and its inverse are generated by

$$\frac{d\phi_t}{dt} = v_t \circ \phi_t, \qquad \frac{d\phi_t^{-1}}{dt} = -D\phi_t^{-1} \cdot v_t.$$

A geodesic is fixed by its initial velocity v_0 through EPDiff:

$$\frac{\partial v_t}{\partial t} = -\mathcal{K}\left[(Dv_t)^T m_t + Dm_t\, v_t + m_t \operatorname{div}(v_t)\right].$$

$\mathcal{L}$: symmetric positive-definite operator, e.g. $(-\alpha\Delta + e)^c$; $m_t = \mathcal{L} v_t$: momentum; $\mathcal{K} = \mathcal{L}^{-1}$: smoothing; D: Jacobian.

Background: LDDMM

With source S, target T and $\psi = \phi^{-1}$, minimize

$$E(v_0) = \frac{\lambda}{2}\,\mathrm{dist}(S \circ \psi_1, T) + \frac{1}{2}\|v_0\|_V^2$$

by gradient descent, repeating:

1. Integrate EPDiff forward from v_0 and the inverse flow to get ψ_1:

$$\frac{\partial v_t}{\partial t} = -K\left[(Dv_t)^T m_t + Dm_t\, v_t + m_t \,\mathrm{div}(v_t)\right], \qquad \frac{d\psi_t}{dt} = -D\psi_t \cdot v_t.$$

2. Warp the source, $S \circ \psi_1$, and compute the gradient at $t = 1$.
3. Carry the gradient back to $t = 0$ (adjoint equations) and update v_0.

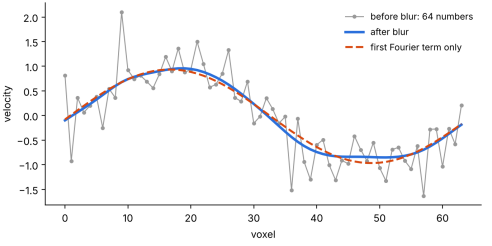
- Each step is computed on the full image grid.
- This makes registration computationally expensive.

$\mathrm{dist}(\cdot,\cdot)$: image dissimilarity (SSD, NCC or MI); $\lambda > 0$: weight.

Background: FLASH

$$\frac{\partial v_t}{\partial t} = -\mathcal{K}\left[(Dv_t)^T m_t + Dm_t\, v_t + m_t \,\mathrm{div}(v_t)\right]$$

1. Velocity is stored at each voxel: $256^3 \approx 17\text{M}$ values.
2. $\mathcal{K}$ is a blur, applied at every step of EPDiff.
3. A blur removes fine detail, so the velocity is always smooth.
4. A smooth field is a Fourier series with only a few slow terms:
$$v(x) \approx \sum_{\xi \text{ small}} \widetilde{v}(\xi)\, e^{2\pi i \xi \cdot x}.$$
5. So store the series coefficients, not voxels: $32^3 \approx 33\text{K}$ values.



ξ: how quickly the motion changes across the image; $\widetilde{v}(\xi)$: how much of that wave.

Background: FLASH

The Fourier transform takes v from position x to frequency ξ (the series coefficients):

$$\mathcal{F}[f](\xi) = \int f(x)\, e^{-2\pi i\, \xi\cdot x}\, dx, \qquad \mathcal{F}^{-1}[\widetilde{f}](x) = \sum_{\xi} \widetilde{f}(\xi)\, e^{2\pi i\, \xi\cdot x}.$$

In Fourier space: derivative $\to$ multiply by $\widetilde{D}(\xi)$; Laplacian (so $\mathcal{L}, \mathcal{K}$) $\to$ multiply by $\widetilde{\mathcal{L}}(\xi)$, $\widetilde{\mathcal{K}}(\xi)$; product $\to$ convolution $\star$.

So geodesic shooting is computed entirely in the bandlimited space $\widetilde{V}$:

$$\frac{\partial \widetilde{v}_t}{\partial t} = -\widetilde{\mathcal{K}}\Big[(\widetilde{D}\widetilde{v}_t)^T \star \widetilde{m}_t + \widetilde{\nabla}\cdot(\widetilde{m}_t \otimes \widetilde{v}_t)\Big].$$

$\widetilde{D}(\xi) = (i \sin 2\pi\xi_j)_j$; $\widetilde{\mathcal{L}}(\xi) = \big[-2\alpha \sum_j (\cos 2\pi\xi_j - 1) + 1\big]^c$, $\widetilde{\mathcal{K}} = \widetilde{\mathcal{L}}^{-1}$.

Background: FLASH

1. Shoot velocity in the truncated Fourier space.
2. Build the deformation map on the full image grid.
3. Compare warped source and target on the full grid.
4. Back-propagate and update the low-dimensional initial velocity.

"The algorithm still maps the velocity fields from the low dimensional Fourier space back to the full image domain to perform forward integration at each iteration. The computational complexity of this step thus dominates the entire procedure."

Frequency Diffeomorphisms

Idea: represent the diffeomorphism itself in Fourier space.

"The high frequency components of a diffeomorphism remain stationary throughout the integration process when computing the deformation associated with smooth velocity fields."

- FLASH: Store the **velocity** in a few Fourier coefficients.
- This paper store the **map** itself in a few fourier coefficients.

Frequency Diffeomorphisms

Writing $\psi \triangleq \phi^{-1}$, the inverse flow $\dfrac{d\psi_t}{dt} = -D\psi_t \cdot v_t$ becomes, in Fourier space,

$$\frac{d\widetilde{\psi}_t}{dt} = -\widetilde{D}\widetilde{\psi}_t * \widetilde{v}_t.$$

Computed with FFTs this costs $O(N^d \log N)$: no saving yet, because $\widetilde{\psi}_0 = \widetilde{e}$ carries the identity's high frequencies. Split the map into identity plus displacement:

$$\widetilde{\psi}_t = \widetilde{e} + \widetilde{u}_t.$$

Since $De \cdot v_t = v_t$, we have $\widetilde{D}\widetilde{e} * \widetilde{v}_t = \widetilde{v}_t$, so

$$\frac{d\widetilde{u}_t}{dt} = -\widetilde{v}_t - \widetilde{D}\widetilde{u}_t * \widetilde{v}_t, \qquad \widetilde{u}_0 = 0.$$

$\widetilde{e}$: Fourier transform of the identity; $\widetilde{u}$: displacement; $*$: circular convolution, truncated.

Frequency Diffeomorphisms

"The high frequency components ... that come from the initial condition $\widetilde{\psi}_0 = \widetilde{e}$ remain unchanged, and only low frequency components vary throughout the integration":

$$\widetilde{u}_t(\xi) = \begin{cases} -\int_0^t \left(\widetilde{v}_\tau(\xi) + \widetilde{D}\widetilde{u}_\tau * \widetilde{v}_\tau(\xi)\right) d\tau, & \xi \le \eta, \\ 0, & \xi > \eta. \end{cases}$$

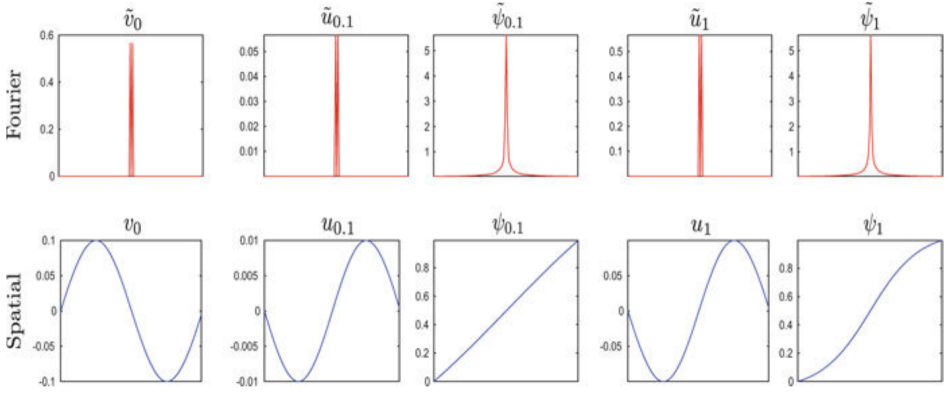


Fig. 1: initial velocity v_0 (a sinusoid), displacement u_t and map ψ_t at $t = 0.1$ and $t = 1$; Fourier (top) and spatial (bottom). η: frequency cutoff.

Frequency Diffeomorphisms

Input: source S, target T. Initialize $\widetilde{v}_0 = 0$. Repeat until convergence:

(a) Shoot the velocity in Fourier space:

$$\frac{\partial \widetilde{v}_t}{\partial t} = -\widetilde{K}\Big[(\widetilde{D}\widetilde{v}_t)^T \star \widetilde{m}_t + \widetilde{\nabla} \cdot (\widetilde{m}_t \otimes \widetilde{v}_t)\Big].$$

(b) Build the map $\widetilde{\psi}_t = \widetilde{e} + \widetilde{u}_t$, evolving only $\xi \leq \eta$:

$$\frac{d\widetilde{\psi}_t}{dt} = -\widetilde{D}\widetilde{\psi}_t * \widetilde{v}_t, \qquad \frac{d\widetilde{u}_t}{dt} = -\widetilde{v}_t - \widetilde{D}\widetilde{u}_t * \widetilde{v}_t, \quad \widetilde{u}_t(\xi) = 0 \text{ for } \xi > \eta.$$

(c_1) Convert $\widetilde{\psi}_1$ to the spatial domain to transport S.

Frequency Diffeomorphisms

(c_2) Compute the gradient at $t = 1$:

$$\nabla_{\widetilde{v}_1} E = \widetilde{\mathcal{K}}\mathcal{F}\left(\frac{\partial \operatorname{dist}(S \circ \psi_1, T)}{\partial (S \circ \psi_1)} \cdot \nabla (S \circ \psi_1)\right).$$

(d) Integrate $\nabla_{\widetilde{v}_1} E$ backward via the adjoint equations to get $\nabla_{\widetilde{v}_0} E$:

$$\frac{d\hat{v}}{dt} = -\operatorname{ad}^{\dagger}_{\widetilde{v}} \hat{h}, \qquad \frac{d\hat{h}}{dt} = -\hat{v} - \operatorname{ad}_{\widetilde{v}} \hat{h} + \operatorname{ad}^{\dagger}_{\hat{h}} \widetilde{v}.$$

(e) Update $\widetilde{v}_0 \leftarrow \widetilde{v}_0 - \delta \nabla_{\widetilde{v}_0} E$.

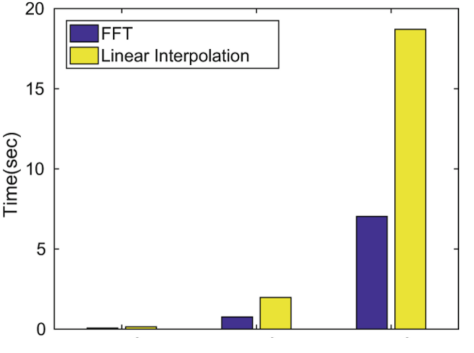
Only step (c) touches the full image grid.

$\hat{v}, \hat{h} \in \widetilde{V}$: adjoint variables; $\operatorname{ad}^{\dagger}$: adjoint operator; δ: step size.

Frequency Diffeomorphisms

Step	FLASH	This paper
(a), (d) shooting and adjoint	$O(Tn^d)$	$O(Tn^d)$
(b) computing the map	$O(TN^d)$	$O(Tn^d)$
(c_1) back to the grid	interpolation $O(N^d)$	FFT $O(N^d \log N)$

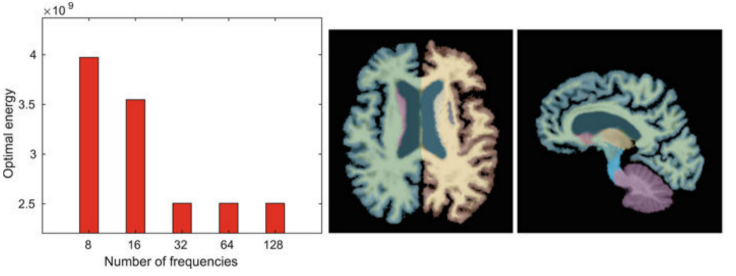
- 256^3 vs. 32^3 values per step: 512 times fewer in 3D.
- In practice the FFT is over twice as fast as linear interpolation (Fig. 2).



N: grid size per axis; n: truncated dimension; T: time steps; d: dimension. Fig. 2 bars: 128^3, 256^3, 512^3.

Results I

36 T1 MRIs, 256^3 at 1 mm; atlas from 20, registered to the other 16. $\alpha = 1.5$, $c = 3$, $\lambda = 10^4$, $T = 10$, NCC, 3-level multi-resolution.

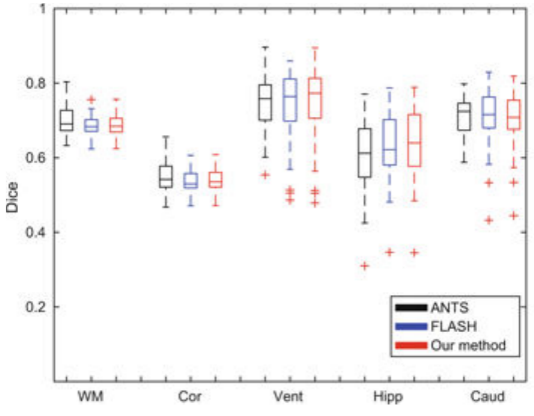


Final energy is flat from $n = 32$ (Fig. 3), so $n = 32$ is used.

Method	Time per image (h)	Memory (MB)
ANTs	2.1	4885
FLASH	0.5	3375
This paper	**0.3**	**3083**

Dice is comparable across methods (paired t-test $p = 0.369$, Fig. 4).

Results II



Model	Time(h)	Memory(MB)
ANTS	2.1	4885
FLASH	0.5	3375
Our method	**0.3**	**3083**

Fig. 4: Dice between atlas-based and manual segmentations for white matter, cortex, ventricles, hippocampus and caudate. Black: ANTs; blue: FLASH; red: this paper.

Results III

10 pregnancies, about 300 volumes each, registered frame to frame; $n = 16$.

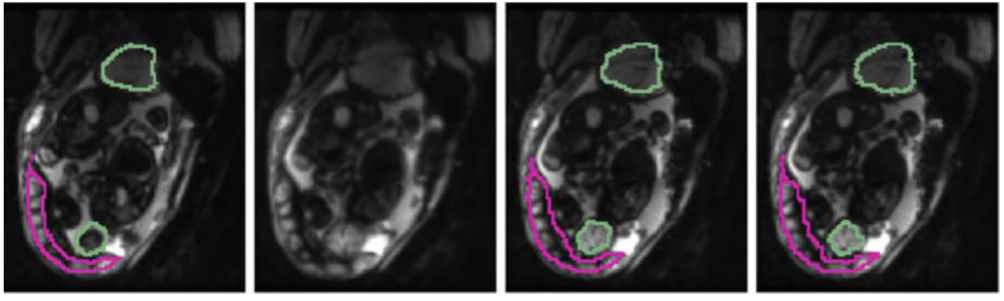
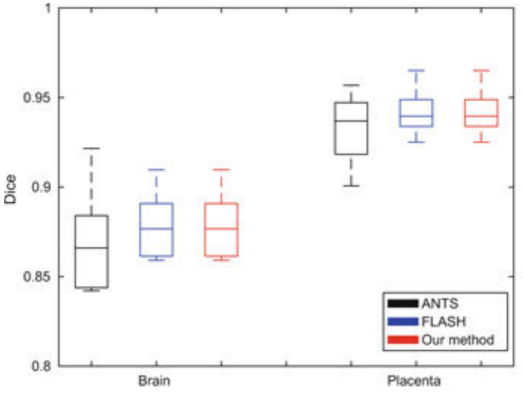


Fig. 5: source with manual segmentation, deformed source, target with manual segmentation, target with propagated segmentation (green: fetal brains; pink: placenta).

Results III



Model	Time(h)	Memory(MB)
ANTS	96.75	253.62
FLASH	37.75	109.52
Our method	**27.00**	**67.32**

Fig. 6: Dice for fetal brains and placenta, and time and memory for 300 registrations. Accuracy is comparable (paired t-test $p = 0.4671$); this paper takes 27.00 h vs. 37.75 h (FLASH) and 96.75 h (ANTs).

Conclusion

What the paper does

- Represent diffeomorphisms in Fourier space.
- Computes the transformation entirely in the bandlimited space, except one FFT per iteration.
- Reports comparable accuracy with lower runtime and memory.
- Applies to other parametrizations, e.g. stationary velocity fields and the original LDDMM.

Limitations

- Band truncation is an approximation checked by experiment, not proven.
- The method assumes a periodic image domain.