Parametric Dictionary Learning in Diffusion MRI

Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, Rachid Deriche

Parametric Dictionary Learning in Diffusion MRI

Sylvain Merlet*, Emmanuel Caruyer*, Aurobrata Ghosh* and Rachid Deriche*
* Athena Project-Team, Inria Sophia-Antipolis – Méditerranée

I. Introduction

In this work, we propose an approach to exploit the ability of compressive sensing to recover diffusion MRI signal and its characteristics from a limited number of samples. Our approach is threefold. First, we learn and design a parametric dictionary from a set of training diffusion data. This provides a highly sparse representation of the diffusion signal. The use of a parametric method presents several advantages: we design a continuous representation of the signal, from which we can analytically recover some features such as the ODF; besides, the dictionary we train is acquisition-independant. Next, we use this sparse representation to reconstruct the signal of interest, using cross-validation to assess the optimal regularization parameter for each signal reconstruction. The use of cross-validation is critical in the ℓ_1 minimization problem, as the choice of the parameter is sensitive to the noise level, the number of samples, and the data sparsity. Third, we use a polynomial approach to accurately extract ODF maxima. In the last section, we motivate and describe the choice of experimental parameters for the HARDI contest.

II. Parametric Dictionary Learning

A. Sparse reconstruction

We reconstruct the signal in an over-complete dictionary of continuous functions. Given a signal $\mathbf{y} \in \mathbb{R}^N$, this is expressed from the coefficient vector $\mathbf{x} \in \mathbb{R}^M$ as

$$\mathbf{y} = \mathbf{HDx} + \epsilon. \tag{1}$$

The matrix $\mathbf{H}$ is the standard $N \times R$ design matrix of the underlying family of continuous functions. The dictionary $\mathbf{D} \in \mathbb{R}^{R \times M}$, is a linear sparsifying transform, such that the signal $\mathbf{y}$ can be reconstructed from a limited number of atoms. This matrix does not depend on the choice of a specific acquisition protocol. The reconstruction is done by minimizing $||\mathbf{y} - \mathbf{HDx}||_2 + \lambda||\mathbf{x}||_1$. Here we propose to learn the dictionary $\mathbf{D}$ from a training dataset.

B. Parametric dictionary learning

We can rewrite Eq. 1 as $\mathbf{y} = \mathbf{Hc} + \epsilon$, and $\mathbf{c} = \mathbf{Dx}$. Given a training set of signals $\mathbf{Y}$ of size $N \times S$, we first estimate the coefficients $\mathbf{C} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{Y}$. Then we look for a dictionary $\mathbf{D}$ of size $R \times M$ and a code $\mathbf{X}$ of size $M \times S$ verifying

$$\arg\min_{\mathbf{D},\mathbf{X}} ||\mathbf{C} - \mathbf{DX}||_2^2 + \alpha||\mathbf{X}||_1, \text{ s.t. } \forall m \leq M, ||\mathbf{d}_m||_2 = 1 \tag{2}$$

with $\mathbf{d}_m$ the m^{th} column of $\mathbf{D}$. We use the python library scikit-learn [5] to solve this problem.

III. Cross-Validation

We use cross validation to assess λ [6]. In particular, we use a K-fold Cross Validation, which consists in splitting the entire data set in K subsets. $K-1$ subsets are used to reconstruct the signal whereas the K^{th} subset enables a tight estimation of regularization parameter λ_K via the evaluation of a cross validation distance. This operation is repeated K times by considering another subset. Then, we keep an average value, $\lambda = \frac{1}{K}\sum_{i=1}^{K}\lambda_i$. To avoid a drastic increase of the computational effort, we split our data set in 5 partitions, i.e. $K = 5$. Our previous experiments showed that it is sufficient to obtain a close approximation of the optimal λ.

IV. A Polynomial Approach for Maxima Extraction

The HARDI contest is based on fiber orientation estimation via the extraction of ODF maxima. After computing the ODF, we extract its maxima using a polynomial approach that can analytically bracket and numerically refine with high precision *all* maxima, ensuring that none are missed [4]. First, leveraging the linear bijection between the spherical harmonic basis and the homogeneous polynomial basis of the same order and degree, we rewrite the ODF as a polynomial, $\mathrm{P}(\mathbf{x} = [x_1, x_2, x_3])$, with $\|\mathbf{x}\|_2 = 1$. Second we formulate the optimization problem for computing the ODF extrema:

$$\partial F/\partial x_1 = \partial F/\partial x_2 = \partial F/\partial x_3 = \|\mathbf{x}\|_2^2 - 1 = 0, \tag{3}$$

where using Lagrange multipliers $F(\mathbf{x}, \Lambda) = \mathrm{P}(\mathbf{x}) - \Lambda\left(\|\mathbf{x}\|_2^2 - 1\right)$. However, instead of optimizing, which is inherently a local approach dependent on initializations, we solve the polynomial system Eq. 3 using a polynomial system solver that can analytically bracket all roots – missing none and refine them numerically to high precision. Finally, we identify the maxima from the extrema using the Bordered Hessian, which is the generalization of the Hessian.

V. Experimental parameters

For the HARDI contest, we choose $\mathbf{H}$ to be constituted of the continuous SPF basis functions [1], [2] and constrain our dictionary $\mathbf{D}$ to represent a combination of the SPF functions. We choose the SPF basis because it facilitates the ODF estimation [3]. We learn our dictionary on the training data set given for the HARDI contest. Then we probe the testing data set on 15 measurements uniformly spread on a single shell at b value $b = 2000\mathrm{s} \cdot \mathrm{mm}^{-2}$.

References

[1] H.E. Assemlal, D. Tschumperlé, and L. Brun. Efficient and robust computation of pdf features from diffusion mr signal. *Medical Image Analysis*, 13(5):715–729, 2009.
[2] Emmanuel Caruyer and Rachid Deriche. Optimal Regularization for MR Diffusion Signal Reconstruction. In *ISBI - 9th IEEE International Symposium on Biomedical Imaging*, Barcelona, Spain, May 2012.
[3] J. Cheng, A. Ghosh, R. Deriche, and T. Jiang. Model-free, regularized, fast, and robust analytical orientation distribution function estimation. In *MICCAI*, volume 6361 of *Lecture Notes in Computer Science*, pages 648–656. Springer, 2010.
[4] A. Ghosh, D. Wassermann, and R. Deriche. A polynomial approach for maxima extraction and its application to tractography in hardi. In Gábor Székely and Horst K. Hahn, editor, *IPMI*, volume 6801 of *Lecture Notes in Computer Science*, pages 723–734. Springer, jul 2011.
[5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Duchesnay E. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
[6] R. Ward. Compressed sensing with cross validation. *Information Theory, IEEE Transactions on*, 55(12):5773 –5782, 2009.