



HAL
open science

Extraction of Nystagmus Patterns from Eye-Tracker Data with Convolutional Sparse Coding

Clément Lalanne, Maxence Rateaux, Laurent Oudre, Matthieu Robert,
Thomas Moreau

► **To cite this version:**

Clément Lalanne, Maxence Rateaux, Laurent Oudre, Matthieu Robert, Thomas Moreau. Extraction of Nystagmus Patterns from Eye-Tracker Data with Convolutional Sparse Coding. EMBC 2020 - 42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society, Jul 2020, Montreal / Virtuel, Canada. pp.928-931, 10.1109/EMBC44109.2020.9175621 . hal-03022547

HAL Id: hal-03022547

<https://hal.science/hal-03022547>

Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extraction of Nystagmus Patterns from Eye-Tracker Data with Convolutional Sparse Coding.

Clément Lalanne^a, Maxence Rateaux^b, Laurent Oudre^c, Matthieu P. Robert^b, Thomas Moreau^d

Abstract—The analysis of the Nystagmus waveforms from eye-tracking records is crucial for the clinical interpretation of this pathological movement. A major issue to automatize this analysis is the presence of natural eye movements and eye blink artefacts that are mixed with the signal of interest. We propose a method based on Convolutional Dictionary Learning that is able to automatically highlight the Nystagmus waveforms, separating the natural motion from the pathological movements. We show on simulated signals that our method can indeed improve the pattern recovery rate and provide clinical examples to illustrate how this algorithm performs.

I. INTRODUCTION

Nystagmus is a variety of involuntary eye movement, acquired in infancy or later in life, that may be associated with reduced or limited vision. It is characterized by a periodic movement of one or both eyes. The many types of Nystagmus can be roughly classified in two large classes of motions. The first ones, referred to as *pendular* Nystagmus, manifest as slow periodic oscillations. The second ones, referred to as *jerk* Nystagmus, are characterized by the alternation of slow phases and quick phases corresponding to eye saccades. In total, Hertle et al. [1] reported 49 different types of Nystagmus and 17 types of saccadic intrusions and oscillations; classified in 8 main categories according to the Classification of Eye Movement Abnormalities and Strabismus (CEMAS)[2]. Several etiological conditions may explain the presence of Nystagmus, such as the presence of an infantile strabismus, a developmental disturbance of the afferent visual pathways or of the oculomotor system. However, there is currently no medical consensus on the diseases that are linked to a specific type of Nystagmus. It is therefore an active research question: the automatic classification and characterization of this phenomenon could improve the diagnosis, prevention and follow-up of patients suffering from this condition.

Traditionally, the detection and monitoring of Nystagmus on young infants was performed by specialized ophthalmologists who study the eye movements by visual inspection. Recently, some dedicated sensors were developed to track the position of both eyes with respect to time (see Figure 1).

^a Clément Lalanne is with Centre Borelli, CNRS, ENS Paris Saclay, Université de Paris, SSA.

^b Maxence Rateaux and Matthieu P. Robert are with Hôpital Necker-Enfants Malades, AP-HP, and Centre Borelli, CNRS, ENS Paris Saclay, Université de Paris, SSA.

^c Laurent Oudre is with Université Sorbonne Paris Nord, L2TI, UR 4443, F-93430, Villetaneuse, France

^d Thomas Moreau is with University Paris-Saclay, Inria, CEA thomas.moreau@inria.fr.



Fig. 1: EyeFant Tracker

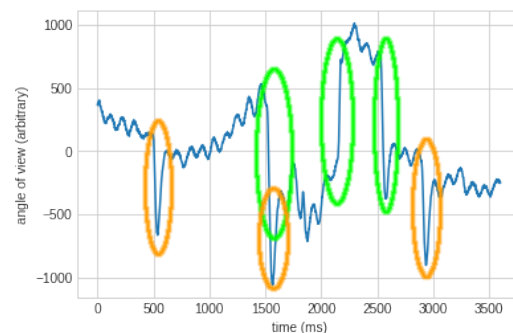


Fig. 2: Example of eye tracker recording of a patient with a pendular Nystagmus (left-eye, horizontal). The signal includes saccades (*green*) as well as eye blinks artefacts (*orange*) hindering the automatic extraction of Nystagmus waveforms.

The EyeFant Tracker, designed by Ober Consulting¹, has been specifically conceived for eye movement recording in infants. The sensors – located in the shadow of the nose – localize the minimum of infra-red cornea reflexion of the eye and deduce the position of the pupil from it. The signal are sampled at 1kHz with a spatial resolution of 0.1° and a linear range of $\pm 35^\circ$ horizontally and $\pm 20^\circ$ vertically. They output left- and right-eye trajectories in the 2D plane which can be visualized with 4 graph: vertical and horizontal components for the each eye.

Figure 2 displays a typical eye-trajectory in one axis. This record comes from one eye of a patient with a pendular Nystagmus. It shows the horizontal projection of the angle that makes the left eye with a normal vector to the face plan. Three simultaneous phenomena are actually recorded by the eye tracker: the low-frequency component is due to

¹<http://www.ober-consulting.com/>

the natural eye movement, the abrupt changes correspond to saccades and eye blinks artefacts (in green and orange) and the small oscillations are due to the Nystagmus (here periodic). The aim of this article is to automatically extract from such signals the components linked to the Nystagmus as patterns associated with activations. This task is especially tricky since it requires to be robust to the other phenomena, i.e. saccades, artefacts and eye movements. To that aim, we propose to use a machine learning tool known as convolutional sparse coding [3] and to adapt it to deal with signals corrupted by a trend (eye movements) and ruptures (saccades and artefacts).

II. BACKGROUND

Convolutional sparse coding aims at representing an input signal as the convolution between patterns (that capture the phenomenon present in the signals) and activations (that specify when these patterns appear in the signal).

More specifically, given a input signal \mathbf{x} of length T , the aim is to learn a library of K patterns $\mathbf{D} = (\mathbf{d}_k)_{1 \leq k \leq K}$ of length W (with $W \ll T$) and their associated activations $\mathbf{Z} = (\mathbf{z}_k)_{1 \leq k \leq K}$ of length $L = T - W + 1$. The general model writes as

$$\mathbf{x} = \sum_{k=1}^K \mathbf{d}_k * \mathbf{z}_k + \mathbf{e} = \mathbf{D} * \mathbf{Z} + \mathbf{e}, \quad (1)$$

where $*$ is the discrete convolution product and \mathbf{e} a noise term [3].

With matrix notations, the Convolutional Sparse Coding (CSC) problem rewrites as

$$\min_{\mathbf{D}, \mathbf{Z}} \frac{1}{2} \|\mathbf{D} * \mathbf{Z} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{Z}\|_1, \quad (2)$$

$\forall k, \|D_k\|_2^2 \leq 1$

where λ allows to control the level of sparsity put on the activations.

This problem is usually solved through alternate minimization, where matrix \mathbf{Z} is updated while \mathbf{D} is fixed (SPARSECODING), and reciprocally (DICTUPDATE). Popular solvers for this problem include AlphaCSC [4] and Sporco [5]. It has been used in several contexts related to healthcare such as blood cell detection in holographic imaging [6] or to measure respiration and heartbeat rates in EEG signals [7].

III. PROPOSED METHOD

In our context, the input signal \mathbf{x} corresponds to an eye trajectory in horizontal or vertical direction, and we aim to learn from this signal a series of recurring patterns stored in \mathbf{D} that would correspond to Nystagmus waveforms and their temporal activations stored in \mathbf{Z} . However, as seen on [Figure 2](#), two types of eye movements limit the use of CSC in this context. Indeed, the input signal \mathbf{x} does not only include the Nystagmus phenomenon, but also the voluntary eye movements, that manifests as a smooth motion and a few eye saccades. Hence, the input signal presents abrupt ruptures that are not taken into account in the model (1). The aim of the proposed model is to modify the standard CSC problem (1) so as to deal with smooth trends and abrupt ruptures.

Algorithm 1 Locally Greedy Coordinate Descent

Require: $x, \mathbf{D}^{(0)}$, parameter $\lambda > 0, \lambda_{TV} > 0$, and $\epsilon > 0$.

- 1: $y^{(0)} = \text{PROXTV}(x, \lambda_{TV})$
 - 2: **repeat**
 - 3: $Z^{(q+1)} = \text{SPARSECODING}(X - y^{(q)}, D^{(q)}, \lambda)$
 - 4: $y^{(q+1)} = \text{PROXTV}(x - Z^{(q+1)} * D^{(q)}, \lambda_{TV})$
 - 5: $D^{(q+1)} = \text{DICTUPDATE}(x - y^{(q)}, Z^{(q)}, \lambda)$
 - 6: **until** $\|Z^{(q+1)} - Z^{(q)}\|_\infty < \epsilon$
-

A. Convolutional Sparse Coding with Piecewise Constant Trend

The Convolutional Sparse Coding with Piecewise Constant Trend (CSC-PCT) model introduced in this article consists in adding an additional term into the standard CSC model:

$$\mathbf{x} = \mathbf{D} * \mathbf{Z} + \mathbf{y} + \mathbf{e}, \quad (3)$$

where \mathbf{y} corresponds to a piecewise constant trend term. The associated optimization problem writes as

$$\min_{\mathbf{D}, \mathbf{Z}, \mathbf{y}} \frac{1}{2} \|\mathbf{D} * \mathbf{Z} + \mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{Z}\|_1 + \lambda_{TV} \|\nabla \mathbf{y}\|_1, \quad (4)$$

$\forall k, \|D_k\|_2^2 \leq 1$

where $\|\nabla \mathbf{y}\|_1$ is a Total Variation (TVL1) constraint defined as

$$\|\nabla \mathbf{y}\|_1 = \sum_t |y_{t+1} - y_t|. \quad (5)$$

The additional constraint imposes a piecewise constant structure to \mathbf{y} , which allows the modelling of the abrupt ruptures as well as the changes in baseline values due to low-frequency phenomena.

B. Resolution

In order to solve (4) with respect to $\mathbf{D}, \mathbf{Z}, \mathbf{y}$, the problem is divided into 3 sub-problems that each updates one term while keeping the remaining terms fixed. These steps are summarized in [Algorithm 1](#).

Note that the \mathbf{D} - and \mathbf{Z} -updates are actually the same than in classical CSC, by replacing \mathbf{x} in (2) by $\mathbf{x} - \mathbf{y}^{(q)}$, where $\mathbf{y}^{(q)}$ is the trend estimation at the q^{th} iteration. Therefore, classical solvers can be used. We relied on AlphaCSC solver [4] which uses a coordinate descent approach for SPARSECODING and FISTA for DICTUPDATE.

For the \mathbf{y} -update, the aim is to compute a piecewise estimate $\mathbf{y}^{(q)}$ of the current residual $\mathbf{x} - \mathbf{D}^{(q)} * \mathbf{Z}^{(q)}$. This corresponds to applying the proximal operator of TVL1 to the residual, for which several off-the-shelf solvers are available. In this article, we used PROXTV [8] which computes the update through a dual framework.

C. Scalability to other setups

The problem (4) uses a classical method in optimization called penalization. Namely, it includes the constraints (smooth trend with saccades) in the function to optimize by measuring "how far" the result is from the constraints set. Here, this is done with the term $\|\nabla \mathbf{y}\|_1$. As a consequence,

our algorithm can be extended to any setup where the motion to get rid of can be expressed by a proximal penalization term.

IV. DATA

A. Synthetic dataset

To evaluate our framework, we used the generative model Equation 3 to simulate look-a-like signals for our data. The voluntary gaze movements are created as the sum of a low frequency pattern and saccades that are sigmoidal ruptures and hence allow abrupt yet smooth mode shifting. The localisation and amplitude of the saccades are chosen randomly with a Poisson law of parameter 1s and a mean amplitude of 20° . The Nystagmus waveforms are generated according to the rough existing Nystagmus classification. The pendular Nystagmus are modelled by a sinusoidal signal and the jerks are modelled by quadratic slow phases and linear quick phases. The frequency and amplitude of the Nystagmus are also chosen randomly with mean 5Hz and 3° . The resulting signal x is obtained by summing these two components. The generated signals have a sampling rate of 1000Hz and length of 10s and a white noise of 0.5° is added to match visually the signals obtained with our recording device.

B. Real dataset

Our clinical evaluation was performed on two different eye movements recordings from two different patients. The first one was realized on an 6-years-old girl with a Down syndrome. The patient exhibited a *Spasmus Nutans type-Nystagmus* (SNTN). There was no ocular deviation and the refraction was: $-6.25 (+1.75 80^\circ)$ for the right eye (RE) and $-5.00 (+1.25 110^\circ)$ for the left eye (LE). The Nystagmus was monocular (only in the RE) and purely pendular. There was no head nodding and no head turn. The second eye movement recording was performed on a 36-years-old women with an oculocutaneous albinism (AOC) and an infantile Nystagmus syndrome (INS). The best corrected visual acuity was $+0.7 \log\text{MAR}$ (RE and LE) and the refraction was: $-7.25 (+4.00 80^\circ) / -8.75 (+3.25 100^\circ)$. Ophthalmological examination revealed an iris transillumination, a peripheral hypopigmentation of the fundus and a foveal hypoplasia. Consents were collected in both cases, according to the French regulation about personal data.

V. RESULTS

All experiments are performed in Python and the code is available online²

A. Pattern recovery on synthetic signals

To evaluate the benefit of joint detrending with CSC on pattern discovery, we used 160 signals generated according to the procedure described in subsection IV-A with random parameters. For each signal x_i , we learn $K = 1$ atom d_i using either our method (JOINT) or the vanilla CSC method

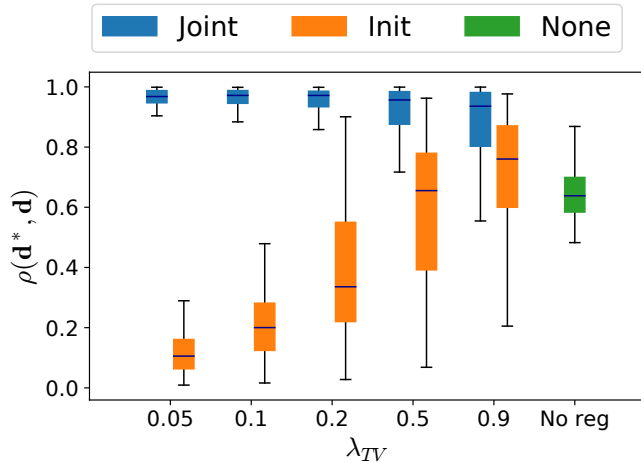


Fig. 3: Pearson correlation coefficient between D and the patterns used to generate the signals for different value of λ_{TV} and $\lambda = 0.5\lambda_{\max}$. The atoms learned without detrending (*green*) are better than the one learned with the addition of a initial detrending (*orange*) while the ones learned with our proposed methods (*blue*) are closer to the ground truth.

with either no pre-processing (NONE) or an initial detrending step (INIT). We then evaluate the recovery of the pattern \mathbf{d}_i^* used to generate the Nystagmus part. Due to the periodicity of the Nystagmus signal, the learned atoms might have a different phase compared to the original pattern and capture multiple periods. To make sure we correctly capture the pattern with the learned atoms, we used the following metric

$$\rho(\mathbf{d}^*, \mathbf{d}) = \max_{\substack{t=1..L \\ l=1..W-L}} \frac{\langle \tilde{\mathbf{d}}^*[t:t+L], \mathbf{d}[l:l+L] \rangle}{\|\mathbf{d}^*\|_2 \|\mathbf{d}\|_2}$$

where $\tilde{\mathbf{d}}$ is the periodized version of \mathbf{d} and $\mathbf{d}[t_1:t_2]$ denotes the coefficient from t_1 to t_2 in \mathbf{d} . This metric extends the Pearson coefficient to be invariant to the phase and the repetition of \mathbf{d}^* in the learned atom \mathbf{d} .

Figure 3 summarizes the results for different values of λ_{TV} and $\lambda = 0.5\lambda_{\max}$. Here, λ_{\max} is the smallest value of λ for which 0 is solution of (2). NONE (*green*) is able to recover the generating patterns with a median score of 0.64 in this configuration. When adding the detrending as a preprocessing step, INIT (*orange*) is able to better recover the patterns, with a median score of 0.76 if $\lambda_{TV} = 0.9$ is large enough. However, the variance of the recovery is largely increased, indicating that it can either recover well the patterns or fail badly. The proposed method JOINT (*blue*) has a median score above 0.95 for $\lambda_{TV} < 0.9$, with a low variance. Note that the behavior of JOINT and INIT are reverse with the variation of λ_{TV} . When λ_{TV} is small, the initial guess for the trend is very close to the full signal and thus vanilla CSC fails to learn the generating atoms as it is only working on the residual. The JOINT method is able to modify the trend estimate to increase the residual where it matters for pattern recovery. As λ_{TV} increases,

²https://github.com/tommoraldetrending_cs_oculo

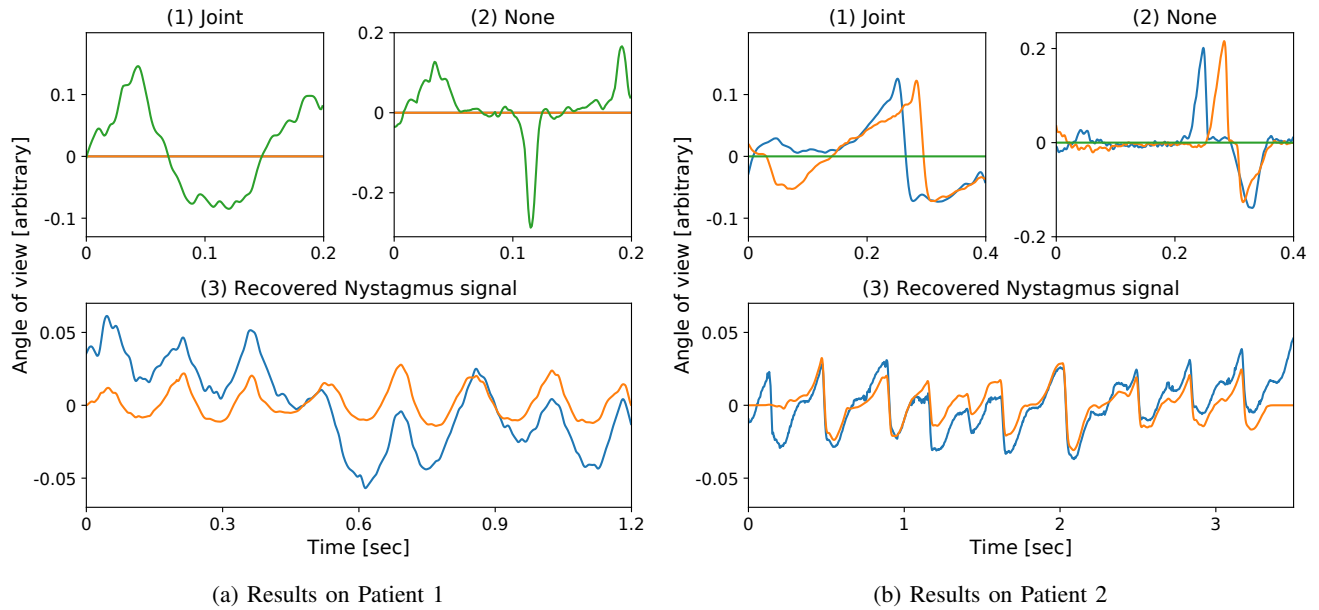


Fig. 4: Results for recordings on two patients presenting (a) a SN syndrome with a pendular pattern and (b) a INS syndrome with characteristic jerky pattern. The patterns learned with JOINT (1) are better at capturing the characteristic waveform of the Nystagmus than the one learned with NONE (2). The bottom part (3) displays a selected part of the original signals (*blue*) as well as the estimated nystagmic component (*orange*). The signal parts are selected to exclude artefacts and saccades, as it is easier to see the Nystagmus patterns in this configuration.

the residual contains more and more of the original signal and the performances of INIT are improved toward the ones of NONE. It becomes better when the trend of a signal is well estimated to not deform the original patterns. But this depends critically on the choice of λ_{TV} , which cannot be set globally to reach good results. Also, the benefit of JOINT decreases, as the trend is better estimated initially. Thus, the joint estimation of the trend and the patterns allows for an aggressive initial estimate of the trend that will be refined by the joint estimation. We repeated this experiment for different values of λ but it did not change the finding, except by degrading the performances of INIT.

B. Results on real data

We applied Algorithm 1 to learn 3 patterns from data collected on the two patients presented in subsection IV-B. As suggested by Figure 3 we tuned the values $\lambda_{TV} = 0.05$ and $\lambda = 0.9\lambda_{\max}$. The patterns learned with JOINT and NONE are presented in Figure 4. For both patients, the result with vanilla CSC are not highlighting the specific waveform of the Nystagmus while our method is able to capture it. In the first case, the algorithm extracts a single symmetric pattern, and confirms the pendular nature of the Nystagmus identified in the clinical examination. For the second patient, two similar waveforms are extracted. They are both composed of a slow phase followed by a fast movement, which are closed to text book Nystagmus description [1]. The presence of foveation periods in the blue pattern – stabilization of the movement in the slow phase – is specific of infantile Nystagmus syndrome. Their presence allows to confirm the INS diagnosis in dubious cases.

VI. CONCLUSION

Analysing the waveforms of Nystagmus is crucial to the clinician, since the causes – and therefore the workup to be performed – depend on the type of Nystagmus. In many cases, the sole clinical analysis does not allow for a robust classification of the Nystagmus. Sometimes, parents report a transient Nystagmus, which is not visible during the clinic, while infraclinical oscillations are identifiable through oculomotor recording. This can lead to overlooking potentially serious underlying diagnoses, such as retinal dystrophies or brain tumours.

REFERENCES

- [1] R. W. Hertle and L. F. Dell’Osso, “Differential diagnosis of nystagmus in infancy and childhood,” in *Nystagmus in infancy and childhood: current concepts in mechanisms, diagnoses, and management*, 2013, pp. 135–184.
- [2] R. W. Hertle and National Eye Institute Sponsored Classification of Eye Movement Abnormalities and Strabismus Working Group, “A next step in naming and classification of eye movement disorders and strabismus,” *Journal of AAPOS*, vol. 6, no. 4, pp. 201–202, Aug. 2002.
- [3] R. Grosse, R. Raina, H. Kwong, and A. Y. Ng, “Shift-Invariant Sparse Coding for Audio Classification,” *Cortex*, vol. 8, p. 9, 2007.
- [4] T. D. La Tour, T. Moreau, M. Jas, and A. Gramfort, “Multivariate convolutional sparse coding for electromagnetic brain signals,” in *NeurIPS*, 2018, pp. 3292–3302.
- [5] B. Wohlberg, “Efficient Algorithms for Convolutional Sparse Representations,” *IEEE Trans. on Im. Proc.*, vol. 25, no. 1, 2016.
- [6] F. Yellin, B. D. Haeffele, and R. Vidal, “Blood cell detection and counting in holographic lens-free imaging by convolutional sparse dictionary learning and coding,” in *ISBI*, 2017.
- [7] P. Wang, M. Liu, H. Zhu, F. Liang, H. Lv, Z. Li, and J. Wang, “Respiration and Heartbeat Rates Measurement Based on Convolutional Sparse Coding,” in *IMBioC*. IEEE, May 2019.
- [8] A. Barbero and S. Sra, “Modular proximal optimization for multidimensional total-variation regularization,” *JMLR*, vol. 19, no. 56, 2018.