EEG Classification for MI-BCI using CSP with Averaging Covariance Matrices: An Experimental Study

Abu Saleh Musa Miah
Department of Computer Science and Engineering
Bangladesh Army University of Science and Technology, Rangpur, Bangladesh
musa@baust.edu.bd

Md Rabiul Islam
Institute of Global Innovation Research
Tokyo University of Agriculture and Technology, Tokyo, Japan
rabiul@sip.tuat.ac.jp

Md.Khademul Islam Molla
Department of Computer Science and Engineering, University of Rajshahi
Rajshahi, Bangladesh
khademul.cse@ru.ac.bd

***Abstract*—To assist disabled people by controlling an external system by using motor imagery (MI) is a common applications of brain computer interface (BCI) field. This paper we focused on an experimental comparison of covariance matrix averaging ways of EEG signal and EEG classification of two types of MI tasks (right-hand*foot and right-hand *left hand). Indeed averaging covariance matrices of EEG signal might be a used in brain computer interfaces (BCI) with common spatial pattern (CSP) method. Structured into trials is a usually paradigms of BCI which we have a tendency to use this structure into account. In addition, covariance matrices with non-Euclidean structure should be consideration likewise. We review much method for averaging covariance matrices in SVM from literature and observe through the experimented result using publicly available four datasets. Our experimental result show that for the case of averaging covariance matrices using Riemannian geometry with small dimension feature issue improve the classification performance. Our result shows the performance increase (2% >performance), but also the limit of this method once the increase feature dimension.**

Keywords*— Brain Computer Interface(BCI),Motor imagery(MI), Riemannian geometry, Electroencephalogram, EEG, support vector machine, Symmetric Positive Definite (SPD) matrices, robust averaging.

I. INTRODUCTION

A modern communication theme neither depends above the brain's tradition pathways output nerve nor the muscles that is the brain computer interface (BCI).For an output device like PC or robot, the BCI machine can translate directly an activities of brain into sequence of control command discussed in [1-2]. An individual researcher or simulators can give action in mind while movement not producing in really by using motor imagery is a cognitive process. Motor imagery (MI) has been broadly utilized with a major methodology in BCI studies discussed in [7, 8]. It's accepted to include comparable cortical areas that are actuated throughout motor preparation and execution discussed in [7].

To seek out reliable representation of brain signal is a very challenging task. To do so different feature extraction and classification method are applied for classify the MI task. To classify the right and left hand motor imagery brain signal first proposed an adaptive autoregressive (AAR) model with linear discriminant analysis (LDA) method [3]. Nevertheless, recent different studies in MI-BCI have shown the improvement of classification accuracy using Tangent space mapping and multi band TSM [4,5,6], but major drawback of using the methods is that the computation time increases dramatically with the number of channels and subbands increase, which can limit the practical use of MI-BCI. For decoding motor imagery an optimal spatial filters with common spatial pattern (CSP) is proposed in [9].In practically CSP reduces the effect of volume conduction effect, electrical signals spatial information go through the skin and skull on the filtered signal.

In this paper we proposed the most well-known classifying EEG signal feature extraction method name is Common Spatial Pattern (CSP) with covariance matrices such as those observed during MI. Discriminating the varies types in BCI protocol of EEG signals based on change in oscillation to compute the suitable spatial filter from the covariance matrices of original signal is a main concept behind this CSP with covariance matrices.

One of the main concept of Common Spatial Pattern (CSP) is simultaneously diagnosable (trials average) with class-related signal of an average covariance matrices. In this manner, utilizing ineffectively assessed or noisy. Main challenge is the poor BCI performance [8] due to Covariance matrices usually lead with poor spatial filters. In this study used improving estimator of covariance matrices should improve Common Spatial Pattern performance.

Also, the symmetric positive definite (SPD) of covariance matrices and Riemannian geometry have appeared effective with such types of data handling [10–12]. Mention the link between Riemannian approaches and CSP have been in [13]. Although, many strategies and theoretically theory are available for average covariance matrices relevant. After all, it's not famed that such methodology is that the most appropriate for EEG signal classification. Subsequently, we reviewed and experimentally compared the advantage and drawback of different methodology for BCI averaging covariance matrices structure.

II. MOTIVATION

A. Common Spatial Pattern

Today CSP is most the foremost prominent methodology for spatial filtering in motor imagery(MI) experiment which first used as a pioneering work Fukurnaga et al. in 190[6].The idea is to utilize a linear transformation with transfer an EEG multi-channel information toward a low-set dimensional subspace. Here the pointed is transformation maximizes the signals of one class variance and in the meantime minimum variance of the alternate class [14].

978-1-7281-3060-6/19/$31.00 ©2019 IEEE

In mathematically, let $X \in R^{N \times C}$ is the matrix of data set which relate to an imaginary movement trials; Number of channel C and N-Number of observations in a trial.

We like to the linear transformation $X_{CSP} = X.W^T$ wherever them spatial filters WJ ∈ RC, composing the projection matrix $w_j \in R^C$, composing the subsequent Rayleigh quotient:

$$J(\omega) = \frac{\omega^T \Sigma_1 \omega}{\omega^T \Sigma_2 \omega} \quad (1)$$

$\sum_i \in R^{C \times C}$ is the band-pass filtered EEG signals of spatial covariance matrices from class i. From the input matrix X, the covariance estimators is

$$S = \frac{1}{N-1} X^T X. \quad (2)$$

Σ_i ,is the i class spatial covariance, that is typically calculated by averaging of covariance matrices as shown

$$\Sigma_i = \frac{1}{|\varphi_i|} \sum_{j \epsilon \varphi_i} S_j \quad (3)$$

Here $|\phi|$ indicates the cardinality of ϕ and φ_i is the arrangement of trials where belonging to each and every class. Particularly assume the computation mean of the EEG signal from covariance matrices have a zero mean (That is true due to filtered with band-pass processing). The problem in Eq. (1) will be resolved by using Eigen value problem including the matrices Σ_1 and Σ_2 as a generalization and has cause different types of extensions and variants [7]. During this article, we tend to don't focus on the CSP method itself however the Σ estimation from the covariance matrices trials. In fact, each variation of CSP depends on the estimation of covariance and contends that this point is over and over again belittled and that the verifiable decision of an averaging that discussed in Equation (3) ought to be considered as more cautiously. It ought to be declared that the maximum likelihood estimator (MLE) about the covariance matrix as displayed in Equation (2) that will be terribly conscious to outliers. In the case of viewpoint, the covariance matrices trials having powerful estimations also additionally enhance the performances of CSP. We are tending to adopt an attempt perspective, as additionally studied discussed in [15], along with we have tried and found a down-weighting changing trials within an estimation of class-covariance. During this section, we tend that don't try or nor to discard down-weight individual electroencephalogram samples which will be clanging however rather a whole trial e.g., art factual trials. Victimization the trial of covariance matrices options, part of recent research discussed in [9–12] examines the quality of victimization specific geometric with manipulating these sorts of data.

B. SPD matrices for Different geometries

In our case covariance matrices is a Symmetric Positive Definite (SPD) matrices— involve into a Euclidean space (linear space). Such explanation, 2 × 2 symmetric positive definite (SPD) matrix A will be written with

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ With } ac - b^2 > 0, a > 0 \text{ and } c > 0,$$

Then $\mathbb{R}^3$ points can represent symmetric matrices. The requirements could be as a cone inside plotted in whose lie entirely SPD matrices. A direct way to deal with averaging matrices in this space is essentially utilize the Euclidean distanceδ_e:

$$\delta_e(A, B) = \|A - B\|_{\mathcal{F}} \quad (4)$$

Wherever $\|.\|_{\mathcal{F}}$ are stands for the Frobenius norm, Computed distance along straight line according to δ by implies the Euclidean geometry of symmetric matrices. Inherently, covariance matrices averaging dependent on resorts a Euclidean geometry to the technique in Equation. (3). Nonetheless, this is a Euclidean geometry experiences the drawbacks. Firstly, as discussed in [16], The SPD matrices space has outfitted with a Euclidean geometry that non-complete space creates. Two SPD matrices Expolating may prompt to indefinite matrices. Only an interpolation issue for averaging covariance matrices, this is not a certainly bot a noteworthy issue however the alleged swelling impact featured in [17] is increasingly risky. The impact reminds the way that the individual determinant of the two average covariance matrices will be greater than combine of their determinants. An artifact from the geometry has suggested for bending.

To maintain this drawback, we've utilized a progressively regular measurement to think about SPD matrices, in particular, the Log Euclidean distance δ_l:

$$\delta_l(A, B) = \|log(A) - log(B)\|_{\mathcal{T}} \quad (5)$$

Where log (·) represents the logarithmic matrix. The LogEuclidean metric (is kernel and derived distance) has been utilized in the writing in [6, 14].

From discussed [18], utilizing a distance between 2 symmetric positive definite (SPD) matrices is a Riemannian metrics such as A and B can be calculated with geodesic. From this Riemannian geometry, the SPD matric turn into a whole space and therefore the distance calculation between two SPD matrices of A and B can be written as:

$$\delta_r(A, B) = \left\| log\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) \right\|_{\mathcal{T}} \quad (6)$$

This distance is invulnerable to the swelling impact. It may be a decent contender for the Covariance Matrices averaging.

III. Covariance Matrices Averages

Such noted [19] with dissemination tensor imaging, tackled averaging covariance matrices from a few different ways, contingent upon the picked geometry. The issue with a set of items an averaging in a metric space is communicated as Equation (7) and sums up it with the notable least squares standard in Euclidean spaces.

$$\min_{\Sigma} \sum_i \delta^2(S_i, \Sigma) \quad (7)$$

Utilizing the Euclidean distance δ_e (in Eq. (4)), the Euclidean normal Σ_E is got with the solution of closed-form during the Eq. (3). Other side, as appeared discussed [17], once exploitation the LogEuclidean distance δ_l (according to the equation in Eq. (5)), we've the subsequent shut shape solution:

$\Sigma_L = exp(\sum_i \log(S_i))$ (8)

Here exp(.) means the exponential of matrices. Notwithstanding, once the Remaining distance δ_r is utilized, there have no closed form of answer for calculating Karcher means Σ_R and optimization techniques discussed in [20, 21] are used. Practically, it is numerically discovered steady and quick to meet, we used the proposed algorithm discussed in [21].

Even once used a minimum distance in the Equation (7), this is often within the formula makes the outlier sensitive problem. To cure this issue, the square is discarded from the equation. As that point, computing a median is done by comprehending

$$\min_{\Sigma} \sum_i \delta(S_i,\ \Sigma) \quad (9)$$

Attributable to the curvature and space flow properties, the existence and uniqueness of the median Σ_m for δ_r have been examined in [23] and has been proposed an iterative algorithm. Utilizing the data geometry [24], to deal with structured objects such as covariance matrices can be also used the divergences [15] has been used a trial approach observed a set of covariance matrices S_i.From a collection of observed covariance matrices S_i assumptions that the Wishart distribution with μ degrees of freedom followed by the matrices. Using β-divergence with observations, it is found that Σ_d is a powerful estimator for matrix minimizing.

As examined in [15] by an interactive procedure this estimator can be computed. Certainly, this methodology compared with Wishart dispersion and it finds a distribution —during the practice, just only with its parameter— that is the nearest with the distributions that in all probability produce the observations.

In this paper, we've compared these 5 standard averaging approach such Arithmetic mean, Harmonic mean, Resolved Mean, Log Euclidean mean P_{Mean} (Eq. (3)), the Log Euclidean mean P_{Med}, the karcher mean(Riemannian Geometric Mean) discussed in[22] minimizing, the Riemannian median.

IV. NUMERICAL EXPERIMENTS

A. Data Description

To check the algorithms of covariance, we tend to used data of from 4 publicly available datasets and from 15 subjects from the BCI competitions; discussed [7].These four datasets contains MI-BCI EEG data.

- Dataset JK-HH 1(four-class) is a personnel dataset contains during left hand, both feet, and right hand MIs EEE signals, as well as the idle state. This dataset were recorded from five subject sc, se, sd, sb, sa from each subject using 29 (twenty nine) channels at position of the international 10-20 system. Recorded each class in 100 trials.
- BCI competition IV- IIa Dataset [25] are EEG signals from subject containing tongue (class-4), foot (class-3), right hand (class-2) and left hand (class-1) these contains 72 trails.
- BCI competition III-IIIa dataset [26] is recorded EEG signal from 3 subjects with 60 electrodes contains left and right hand, foot and tongue of MI (class-1, class-2, class-3, class-4 respectively). Each signal has 72 trials.
- BCI competition III-Iva dataset [27] contains EEG signals that recorded from 5 subjects and each subject from 118 electrodes, are contained right hand and foot MI. This signal has 272 trials.

Here used bandpass filter in 8-30 Hz from All data sets EEG signals with a fifth order Butterworth filter. Feature extraction from the time segment located between 0.5 and 2.5s.

B. Classifier Setup

We set here $\mu = 20$ and $\beta = 0.001$ as the recommended for estimator from [12] the minimum divergence, here we extracted feature from each trial consists of the variance of log projected EEG signals on the CSP filters selection (that examined [14]). After all features are employed to a linear support vector machine (SVM) by using a cross-validation procedure [28] (for 30 iterations with 80%−20% splits).

C. Results and Ddiscussion

1) Table I Describes JKHH Dataset (2-class):

Table I shows the experimental results of classification with obtained various SCM average. Here Reference proposed method P_{MEAN} and geometric median P_{MED} shows the little improvement from P_{Mean}, and the Riemannian geometry (Riemannian median, mean and Log Euclidean mean, Median). It thinks to have an advantage in an average accuracy (varies on subject) ranging between 68.88% to 73.63% against 67%.

2) Table II Describes IV-IIa Dataset (2-class):

Table II shows the experimental results of classification with obtained various SCM average. Here Reference proposed method P_{MEAN} and geometric median P_{MED} shows the little improvement from P_{Mean}, and the Riemannian geometry (Riemannian median, mean and Log Eucliedean mean, Median). It appears to have a unmistakable advantage in an average accuracy (varies on subject) ranging between 78% to 82.50 % against 76.31%.

3) Table III Describes III-IIIa Dataset (2-class):

Table III shows the experimental results of classification with obtained various SCM average. Here Reference proposed method P_{MEAN} and geometric median P_{MED} shows the little improvement from P_{Mean}, and the Riemannian geometry (Riemannian median, mean and Log Euclidean mean, Median). It appears to have an unmistakable advantage in an average accuracy (varies on subject) ranging from 77% to 83.51%.

4) Table IV Describes III-IVa Dataset (2-class):

Table IV shows the experimental results of classification with obtained various SCM average. Here Reference proposed method P_{MEAN} and geometric median P_{MED} shows the little improvement from P_{Mean}, and the Riemannian geometry (Riemannian median, mean and Log Eucliedean mean, Median). It appears to have a unmistakable advantage in an average accuracy (varies on subject) ranging between 85.18% to 88.03% against 82%.

V. Conclusion

In this experimental study, empirically observation compared some methodology for averaging trials with adopted a perspective trial on EEG data. The averaging covariance is used with Riemannian geometry for small dimensional problems. The Euclidean geometry more suited when numerical problem appear and dimensionality grows. During this experiment, we left aside the complex issue with a covariance matrices trial evaluation.

On the other hand, Once the insufficient number of time samples and if the correlated time samples, almost indefinite matrices, shrinkage a type of regularization-ought to be utilized as in[28,29]. Likewise, the interrelation between these estimators and the averaging stays vague and ought to be experimentally explored.

For a future research, some progressively numerical experiments ought to be done by ever-changing the rate of outlier sand plotting of the CSP design. We've tried to assess what different geometries may convey to CSP technique in a promisingly way to extract smaller dimensional feature covariance matrices. All things considered, approaches like [30] may bridge the gap that presently separate CSP-based methodology and Riemannian technique in BCI.

TABLE I: Classification Accuracy [%] of TSM and MbTSM for Dataset JK-HH1 (two classes)

Method	**Subject**				
	Sa	Sb	Sd	Se	Ave
Arithmetic Mean P_{MEAN}	78.00	56.50	53.50	88.00	68.88
Riemannian Geometric Mean P_{MEAN}	79.00	60.50	59.50	86.50	**71.38**
Log Euclidean Mean P_{MEAN}	**76.50**	58.00	59.00	88.00	70.38
Harmonic Mean P_{MEAN}	75.50	62.50	54.50	87.00	69.88
Resolvent Mean P_{MEAN}	78.50	**62.50**	**60.50**	**89.00**	**72.63**
Euclidean geometric median P_{MED}	76.50	64.00	53.00	87.50	70.25
Riemannian geometric median P_{MED}	78.50	56.50	60.00	86.00	70.25
Log-Euclidean geometric median P_{MED}	73.5	61.00	59.00	88.00	70.38

TABLE II: Classification Accuracy [%] of TSM and MbTSM for Dataset IV-IIa (two classes)

Method	**Subject**					
	A01T	A03T	A07T	A08T	A09T	Ave
Arithmetic Mean P_{MEAN}	73.61	84.02	81.94	88.19	72.91	80.13
Riemannian Geometric Mean P_{MEAN}	76.38	83.33	81.94	94.44	71.52	81.52
Log Euclidean Mean P_{MEAN}	**80.55**	83.33	83.33	93.05	72.22	**82.50**
Harmonic Mean P_{MEAN}	77.08	79.86	7986	93.75	72.22	80.55
Resolvent Mean P_{MEAN}	80.55	**81.94**	79.86	**94.44**	**72.91**	**81.94**
Euclidean geometric median P_{MED}	75.69	84.02	80.55	86.11	68.05	**78.88**
Riemannian geometric median P_{MED}	77.77	83.33	81.25	94.44	71.52	81.66
Log-Euclidean geometric median P_{MED}	73.61	81.94	81.25	87.50	71.52	79.16

TABLE III: Classification Accuracy [%] of TSM and MbTSM for Dataset III-IIIa (two classes)

Method	**Subjects**			
	K3b	K6b	l1b	Ave
Arithmetic Mean P_{MEAN}	92.22	90.00	63.33	81.85
Riemannian Geometric Mean P_{MEAN}	74.44	93.33	63.33	77.03
Log Euclidean Mean P_{MEAN}	**78.88**	95.00	63.33	79.07
Harmonic Mean P_{MEAN}	85.55	93.33	60.00	79.63
Resolvent Mean P_{MEAN}	85.55	**96.66**	68.33	83.51
Euclidean geometric median P_{MED}	86.66	95.00	63.33	81.66
Riemannian geometric median P_{MED}	74.44	93.33	65.00	77.59
Log-Euclidean geometric median P_{MED}	80.00	73.33	65.00	72.78

TABLE IV: Classification Accuracy [%] of TSM and MbTSM for Dataset III-IVa (two classes)

Method	Subject		
	aa	al	Ave
Arithmetic Mean P_{MEAN}	75.71	94.64	85.18
Riemannian Geometric Mean P_{MEAN}	80.00	94.28	87.14
Log Euclidean Mean P_{MEAN}	81.42	94.64	88.03
Harmonic Mean P_{MEAN}	77.50	95.35	86.43
Resolvent Mean P_{MEAN}	78.21	95.35	86.78
Euclidean geometric median P_{MED}	81.42	94.28	87.85
Riemannian geometric median P_{MED}	80.00	94.28	87.14
Log-Euclidean geometric median P_{MED}	80.35	94.28	87.32

REFERENCES

[1] G. Dornhege, J. del R. Mill´an, T. Hinterberger, D. J. McFarland, and K.-R. M¨uller, Toward Brain-Computer Interfacing, The MIT Press, 2007.

[2] K. Fukunaga and W. Koontz, “Application of the Karhunen-Lo`eve expansion to feature selection and ordering,” IEEE Trans. Comp., 1970.

[3] Pfurtscheller G, Pregenzer M and Neuper C 1994 Visualization

of sensorimotor areas involved in preparation for hand movement based on classification of μ and central βrhythms in single EEG trials in man *Neurosci. Lett.***181** 43–6

[4] Barachant A, Bonnet S, Congedo M and Jutten C 2012 Multiclass brain-computer interface classification by Riemannian geometry IEEE Transactions on Biomedical Engineering 59 920–928

[5] Islam, M. R., Tanaka, T. & Molla, M. K. I. Multiband tangent space mapping and feature selection for classification of EEG during motor imagery. Journal of Neural Engineering 15, 046021 (2018).

[6] Islam, M. R., Molla, M. K. I., Nakanishi, M. & Tanaka, T. Unsupervised frequency- recognition method of SSVEPs using a filter bank implementation of binary subband cca. Journal of Neural Engineering 14, 026007 (2017).

[7] F. Lotte and C. Guan, “Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms,” IEEE Trans. Biomed. Eng., 2011.

[8] B. Reuderink and M. Poel, “Robustness of the common spatial patterns algorithm in the BCI-pipeline,” Tech. Rep., University of Twente, 2008.

[9] H. Ramoser, J.Muller-Gerking, and Pfurtschelle, “Optimal Spatial filtering of single trialeeg during imagined hand movement,” IEEE Transaction on Biomedical Engineering, vol.59, no. 4,pp, 441-446, Dec 2000.

[10] Barachant, S. Bonnet, M. Congedo, and C. Jutten, “Riemannian geometry applied to BCI classification,”in LVA/ICA, 2010.

[11] F. Yger, “A review of kernels on covariance matrices for BCI applications,” in Proc. IEEE MLSP, 2013.

[12] F. Yger and M. Sugiyama, “Supervised LogEuclidean metric learning for symmetric positive definite matrices,” preprint arXiv:1502.03505, 2015.

[13] Barachant, S. Bonnet, M. Congedo, and C. Jutten,“Common spatial pattern revisited by Riemannian geometry,”in Proc. IEEE MMSP, 2010.

[14] Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R.M¨uller, “Optimizing spatial filters for robust EEG single-trial analysis,” IEEE Sig. Proc. Mag., 2008.

[15] W. Samek and M. Kawanabe, “Robust common spatial patterns by minimum divergence covariance estimator,” in Proc. ICASSP, 2014.

[16] T. Fletcher and S. Joshi, “Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors,” in Proc. CVAMIA and MMBIA. 2004.

[17] Arsigny, P. Fillard, X. Pennec, and N. Ayache, “Geometric means in a novel vector space structure on symmetric positive-definite matrices,” SIAM Journal on Matrix Analysis and Applications, 2007.

[18] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2009.

[19] Dryden, A. Koloydenko, and D. Zhou, “Non- Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging,” The Annals of Applied Statistics, 2009.

[20] M. Moakher, “A differential geometric approach to the geometric mean of symmetric positive-definite matrices,” SIAM J. Mat. An. App., 2005.

[21] B. Jeuris, R. Vandebril, and B. Vandereycken, “A survey and comparison of contemporary algorithms for computing the matrix geometric mean,” Elec. Trans. Num. An., 2012.

[22] D. Bini and B. Iannazzo, “Computing the Karcher mean of symmetric positive definite matrices,” Linear Algebra and its Applications, 2013.

[23] T. Fletcher, S. Venkatasubramanian, and S. Joshi, “The geometric median on Riemannian manifolds with application to robust atlas estimation,” NeuroImage, 2009.

[24] S.-I. Amari and A. Cichocki, “Information geometry of divergence functions,” Bulletin of the Polish Academy of Sciences: Technical Sciences, 2010.

[25] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller, “Seperability of four-class motor imagery data using independent components analysis,” J. Neur. Eng., 2006.

[26] Schl¨ogl, F. Lee, H. Bischof, and G. Pfurtscheller, “Characterization of four-class motor imagery EEG data for the BCI-competition 2005,” J. Neur. Eng., 2005.

[27] Blankertz, K. Muller, D. Krusienski, G. Schalk, J. Wolpaw, A. Schlogl, G. Pfurtscheller, J. Millan,M Schroder, and N. Birbaumer, “The BCI competition III: Validating alternative approaches to actual BCI problems,” IEEE Trans. Neur. Syst. Rehab., 2006.

[28] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods, Cambridge University Press, 2000.

[29] P. Rousseeuw and K. Van Driessen, “A fast algorithm for the minimum covariance determinant estimator,”Technometrics, 1999.

[30] D. Olive, “Why the Rousseeuw Yohai paradigm is one of the largest and longest running scientific hoaxes in history,” Tech. Rep.,