Classification of Seismic Volcanic Signals Using Hidden-Markov-Model-Based Generative Embeddings

Manuele Bicego, *Member, IEEE*, Carolina Acosta-Muñoz, and Mauricio Orozco-Alzate, *Member, IEEE*

***Abstract*—The automated classification of seismic volcanic signals has been faced with several different pattern recognition approaches. Among them, hidden Markov models (HMMs) have been advocated as a cost-effective option having the advantages of a straightforward Bayesian interpretation and the capacity of dealing with seismic sequences of different lengths. In the volcano seismology scenario, HMM-based classification schemes were only based on a standard and purely generative scheme, i.e., the Bayes rule: training an HMM per class and classifying an incoming seismic signal according to the class whose model shows the highest likelihood. In this paper, a novel HMM-based classification approach for pretriggered seismic volcanic signals is proposed. The main idea is to enrich the classical HMM scheme with a discriminative step that is able to recover from situations when the classical Bayes classification rule is not sufficient. More in detail, a generative embedding scheme is used, which employs the models to map the signals into a vector space, which is called generative embedding space. In such a space, any discriminative vector-based classifier can be applied. A thorough set of experiments, which is carried out on pretriggered signals recorded at Galeras Volcano in Colombia, shows that the proposed approach typically outperforms standard HMM-based classification schemes, also in some cross-station cases.**

***Index Terms*—Generative embeddings, hidden Markov models (HMMs), pattern recognition, seismic volcanic signals, volcano seismology.**

I. Introduction

In RECENT years, the importance of automatic analysis of seismic volcanic signals has rapidly grown, permitting a better understanding of the volcano activity and its interactions with earthquakes and seismic processes [1], [2]. In this context, a rather interesting and challenging problem is represented by the classification task, which is aimed at differentiation among the different seismic events that can be caused by volcanic processes. In the literature, many approaches have been proposed to face this problem, with each one characterized by different representation strategies, accuracy (ACC) levels, interpretability features, efficiencies, and computational requirements (see [3] for a comprehensive list of references).

Among all these techniques, a relevant class is composed by methods based on hidden Markov models (HMMs) [4]–[13], which are a probabilistic approach whose usefulness has been assessed in many different pattern recognition scenarios. Such models appear to be very useful for the analysis of seismic signals, due to their capability to deal with highly variable sequential data (possibly of different lengths), their clear and elegant Bayesian interpretation, and their fast and effective algorithms for training and testing [14]. Moreover, such models can also be applied, either in their basic or extended form, for simultaneous detection (signal versus noise) and classification of continuous signals. Two appealing approaches, in this sense, have been recently proposed [12], [13]. In [12], Beyreuther and Wassermann demonstrated that the modified versions of HMMs [e.g., the so-called hidden semi-Markov models (HSMM)] can increase the performances, due to the inclusion of more realistic time dependence in the model. In [13], additionally, the HSMM-based approach is refined with respect to the parameter fitting by introducing state clustering to improve the time discretization without increasing the number of states and, thereby, the computational complexity.

In this paper, a novel HMM-based approach for the classification of seismic signals is proposed, which is based on recent advances in the machine learning and pattern recognition fields. In particular, the proposed approach is based on a novel classification paradigm that is alternative to the classical Bayes rule, with this being the standard approach employed in all the aforementioned papers [4]–[13]. The main idea is to follow a hybrid generative–discriminative paradigm [15], [16], namely to take advantage of the best of the generative and discriminative paradigms. The former paradigm is based on probabilistic class models and *a priori* class probabilities, which are learnt from training data and combined via the Bayes law to yield posterior probabilities. This is the standard approach for HMM classification schemes; the latter paradigm is aimed at directly learning class boundaries or posterior class probabilities from data, without relying on generative class models [17], [18]. A clear example here is represented by support vector machines (SVMs).

Manuscript received February 17, 2012; revised July 13, 2012 and September 13, 2012; accepted September 16, 2012. Date of publication December 10, 2012; date of current version May 16, 2013. This work was supported in part by the University of Verona through the CooperInt Program 2010 Edition and in part by Universidad Nacional de Colombia through "Convocatoria de fortalecimiento a programas de pregrado y postgrado de la Facultad de Ingeniería y Arquitectura."

M. Bicego is with the Department of Computer Science, University of Verona, 37134 Verona, Italy.

C. Acosta-Muñoz is with Observatorio Vulcanológico y Sismológico de Manizales, Servicio Geológico Colombiano, Manizales, Colombia.

M. Orozco-Alzate is with Departamento de Informática y Computación, Universidad Nacional de Colombia, Manizales, Colombia.

Digital Object Identifier 10.1109/TGRS.2012.2220370

Following this route, the typical HMM classification scheme, which is purely generative, can be enriched with a discriminative part; this hybrid scheme has shown to be beneficial in different applications and scenarios dealing with HMMs [19]–[23]. Encouraged by these results, in this paper, we explore the use of such hybrid scheme in the seismic signal classification scenario. In particular, here, we employ a generative embedding scheme, where the basic idea is to exploit a generative model to map the objects to be classified into a vector space, where discriminative vector-based techniques (e.g., kernel-based SVMs) can be used. These hybrid generative–discriminative schemes seem to be very suitable for this context. It has been shown in the literature (e.g., [23]) that they may be very advantageous in situations when generative models cannot be properly estimated, due to a bad choice of parameters, insufficient data, or high complexity of the problem, with all these features characterizing at different levels the seismic signal classification problem. Moreover, their capability of being effective even with somehow rough models may open the possibility of employing such schemes for the classification of signals recorded from different stations (e.g., applying a classifier trained on a station to signals gathered from a newly established one, for which no training signals are available). A study concerned to this issue has also been provided in the final part of this paper.

More in detail, the approach that we propose works as follows. First, pretriggered seismic waveforms, i.e., detected by an amplitude-ratio-based [short-term average/long-term average (STA/LTA)] algorithm, are trasformed into the frequency domain (this being a common choice in many approaches since it is widely accepted that differences in the spectral content allow for discriminating different types of volcanic earthquakes [24]). Then, class-related HMMs are learnt from training data, which are subsequently used to project all the signals (both training and testing) into a vector space (the generative embedding space). Different embedding schemes have been tested, ranging from the Fisher score, which is the first and most used one [15], to more recent and HMM-specific methods [23]. Finally, in the generative embedding space, a discriminative classifier has been employed, namely a radial-basis-function (rbf) SVM.

In spite of having restricted ourselves to pretriggered data (pretriggered data are often directly provided by most of the seismic acquisition technologies[1]), our proposed approach can be straightforwardly adapted to continuous seismic data by considering the background noise as an additional class and by using a sliding fixed-length window to define a portion of the signal to be classified (clearly provided that the sampling rate is the same during all the continuous acquisition). Alternatively, HMM-based generative embeddings can also be used to cast the triggering problem into a two-class classification task, i.e., distinguishing noise from seismicity. This alternative has been adopted by authors working on continuous data [13].

The proposed approach has been thoroughly tested with more than 1250 pretriggered signals (divided into four classes) coming from Galeras Volcano in Colombia. A comparison with the classical generative scheme [4]–[11] is provided, showing the superiority of the proposed approach with respect to standard HMM-based classification schemes.[2] An analysis of the impact of the data set dimension on the performances is also provided via learning curves. Finally, some insights on the applicability of the proposed schemes to signals acquired in other stations are given.

The remainder of this paper is organized as follows. In Section II, we first review the theory of HMMs, mainly to fix the notation; subsequently, the generative embedding procedure is summarized. The proposed approach is described in Section III, whereas experimental results and discussion are detailed in Section IV. Finally, Section V concludes this paper. A list of the abbreviations used throughout this paper is provided in the Appendix.

II. Background

A. HMMs

A discrete-time first-order HMM [14] is a probabilistic model that describes a stochastic sequence[3] $\boldsymbol{O} = (O_1, O_2, \ldots, O_T)$ as being an indirect observation of a hidden Markovian random sequence of states $\boldsymbol{Q} = (Q_1, Q_2, \ldots, Q_T)$, where, for $t = 1, \ldots, T$, $Q_t \in \{1, 2, \ldots, N\}$ (the set of states). Each state has an associated probability function that specifies the probability of observing each possible symbol, given the state. An HMM is thus fully specified by a set of parameters $\boldsymbol{\lambda} = \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}\}$, where $\boldsymbol{A} = (a_{ij})$ is the transition matrix, i.e., $a_{ij} = P(Q_t = j \,|\, Q_{t-1} = i)$; $\boldsymbol{\pi} = (\pi_i)$ is the initial state probability distribution, i.e., $\pi_i = P(Q_1 = i)$, and $\boldsymbol{B} = (\boldsymbol{b}_i)$ is the set of emission probability functions. If the observations are continuous, each $\boldsymbol{b}_i$ is a probability density function, e.g., a Gaussian or a mixture of Gaussians. If the observations belong to a finite set $\{v_1, v_2 \ldots, v_S\}$, each $\boldsymbol{b}_i = (b_i(v_1), b_i(v_2), \ldots, b_i(v_S))$ is a probability mass function with $b_i(v_s) = P(O_t = v_s \,|\, Q_t = i)$ being the probability of emitting symbol v_s in state i.

The training of the model, given a set of sequences $\{\boldsymbol{o}^{(i)}\}$, is usually performed using the standard Baum–Welch reestimation technique [14], which determines parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi})$ by maximizing probability $P(\{\boldsymbol{o}^{(i)}\}|\boldsymbol{\lambda})$. The evaluation step, i.e., the computation of the log probability $\log P(\boldsymbol{o}|\boldsymbol{\lambda})$, given model $\boldsymbol{\lambda}$ and sequence $\boldsymbol{o}$ to be evaluated, is performed using the *forward–backward procedure* [14].

B. Standard HMM-Based Classification Scheme

Given a set of sequences $\{\boldsymbol{o}^{(i)}\}$, relative to a C-class problem, the standard classification scheme (i.e. the Bayes rule) is

[1] As an example, consider sensor technologies by Güralp Systems, whose digitizers are able to simultaneously run a STA/LTA event-triggering algorithm in parallel with the continuous acquisition. This way, the system can record continuously at a relatively low sample rate and record at a much higher sample rate during short periods of interest.

[2] Other experiments—not shown here—show that the proposed scheme outperforms also standard feature-based approaches, such as those based on neural networks [25].

[3] We adopt the common convention of writing stochastic variables with uppercase and realizations thereof in lowercase.

realized in the following way.

1) **Training**: For every class c, HMM $\boldsymbol{\lambda}_c$ is trained, using only the training sequences belonging to such class. At the end of the training process, a set of HMMs $\{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_C\}$ is available.
2) **Testing**: An unknown sequence $\boldsymbol{o} = (o_1, \ldots, o_T)$ is assigned to the class whose model shows the highest likelihood (assuming that every class has the same prior probability), i.e., label $\ell(\boldsymbol{o})$ is determined as

$$\ell(\boldsymbol{o}) = \arg\max_c \log P(\boldsymbol{o}|\boldsymbol{\lambda}_c). \tag{1}$$

C. Generative Embedding Approach

Even if the Bayes rule represents the theoretical optimal decision rule (i.e., leading to the minimum probability of error [26]), in practice, generative HMMs may suffer from poor discriminative capabilities. This is likely to occur in one of the following scenarios:

- poorly estimated class models, e.g., due to insufficient learning examples;
- improper models, e.g., due to bad model definition or conditional dependence of the states;
- possible class overlap, as may occur, for instance, in medical problems where patient diagnoses may not be consistent between different medical doctors.

To face this issue, several efforts have been recently made to enrich the generative paradigm with discriminative information. Among others, the so-called generative embedding methods (or generative score spaces) have gained remarkable importance. In this class of schemes, the basic idea is to use the HMM (or, in general, the generative model) to map the objects to be classified into a vector space, where discriminative techniques, possibly kernel based, can be used.

More formally, in our scenario, the generative embedding is defined as function Φ that maps, through a set of HMMs $\{\boldsymbol{\lambda}_c\}$, an observed sequence $\boldsymbol{o} = (o_1, \ldots, o_T)$ into a vector. Different approaches have been proposed to determine the set of models used to build the embedding [27]. Here, we adopt the following method: Given a C-ary classification problem, we train one HMM per class, performing one embedding for every model (resulting in C different embeddings) and, finally, concatenating the vectors, i.e.,

$$\Phi(\boldsymbol{o}) = \left[\phi(\boldsymbol{o}, \boldsymbol{\lambda}_1), \ldots, \phi(\boldsymbol{o}, \boldsymbol{\lambda}_C)\right]. \tag{2}$$

where $\phi(\boldsymbol{o}, \boldsymbol{\lambda}_c)$ is the embedding of the object $\boldsymbol{o}$ through model $\boldsymbol{\lambda}_c$.

In the following, we describe how $\phi(\boldsymbol{o}, \boldsymbol{\lambda}_c)$ is defined in the four cases considered in this paper. All the quantities needed to compute the different embeddings can be easily obtained using the forward–backward procedure [14].

1) FSE: In the Fisher score embedding (FSE), each sequence is represented by a vector containing derivatives of the log likelihood of the generative model with respect to each of its parameters, evaluated in such sequence. Formally, we have

$$\phi^{\mathrm{FSE}}(\boldsymbol{o}, \boldsymbol{\lambda}) = \left[\frac{\partial \log P(\boldsymbol{O}=\boldsymbol{o}|\boldsymbol{\lambda})}{\partial \lambda_1}, \ldots, \frac{\partial \log P(\boldsymbol{O}=\boldsymbol{o}|\boldsymbol{\lambda})}{\partial \lambda_L}\right]^\top \in \mathbb{R}^L \tag{3}$$

where λ_i represents one of the L parameters of the model $\boldsymbol{\lambda}$ (elements of the transition matrices, emission, and initial probabilities). For more details see [28].

2) LLE: The log-likelihood embedding (LLE) is a very simple generative embedding, which is first introduced in [22]. With this rule, every sequence is just mapped on its log-likelihood, given the model. Formally, we have

$$\phi^{\mathrm{LLE}}(\boldsymbol{o}, \boldsymbol{\lambda}) = \left[\log P(\boldsymbol{o}|\boldsymbol{\lambda})\right] \in \mathbb{R}. \tag{4}$$

Even if very simple, this approach has shown to be very effective in different cases, particularly when the HMM training is adequately good [22], [23].

3) SE: The state embedding (SE) is a recently introduced generative embedding [23], in which the ith component of the generative embedding vector, for an observed sequence $\boldsymbol{o}$, measures the sum (over time) of the probabilities that the HMM $\boldsymbol{\lambda}$ is in state i while observing $\boldsymbol{o}$. Formally, we have

$$\phi^{\mathrm{SE}}(\boldsymbol{o}, \boldsymbol{\lambda}) = \left[\sum_{t=1}^{T} P(Q_t=1|\boldsymbol{o}, \boldsymbol{\lambda}), \ldots, \sum_{t=1}^{T} P(Q_t=N|\boldsymbol{o}, \boldsymbol{\lambda})\right]^\top \in \mathbb{R}^N. \tag{5}$$

Every component of the generative embedding vector can be interpreted as the expected number of transitions from the corresponding state, given the observed sequence [14].

4) TE: Transition embedding (TE) is similar to the SE, but it considers probabilities of transitions rather than states [23]. Naturally, it is defined as

$$\phi^{\mathrm{TE}}(\boldsymbol{o}, \boldsymbol{\lambda}) = \begin{bmatrix} \sum_{t=1}^{T-1} P(Q_t = 1, Q_{t+1} = 1|\boldsymbol{o}, \boldsymbol{\lambda}) \\ \sum_{t=1}^{T-1} P(Q_t = 1, Q_{t+1} = 2|\boldsymbol{o}, \boldsymbol{\lambda}) \\ \vdots \\ \vdots \\ \sum_{t=1}^{T-1} P(Q_t = N, Q_{t+1} = 1|\boldsymbol{o}, \boldsymbol{\lambda}) \\ \vdots \\ \sum_{t=1}^{T-1} P(Q_t = N, Q_{t+1} = N|\boldsymbol{o}, \boldsymbol{\lambda}) \end{bmatrix} \in \mathbb{R}^{N^2}. \tag{6}$$

Each of the N^2 components of the vector can be interpreted as the expected number of transitions from a given state to another state, given the observed sequence [14].

III. Proposed Approach

The proposed approach is summarized in the following steps.

1) **Preprocessing**. Seismic events in the continuous records are detected by applying the so-called STA/LTA trigger [3]. Afterward, spectrograms are computed by using a 128-point fast Fourier transform (FFT) and a Hanning window of 128 points with an overlapping of 64 points. The magnitude in decibels is computed as $20\log_{10}|X|$, where X is a matrix containing the short-time Fourier transform of the input signal. Spectrograms were selected

for representation because spectral analysis, either in the frequency domain or the time–frequency domain, is widely used for both visual [24] and computer-based [1] inspections of seismic phenomena; other largely applied representations are based on descriptive spectral parameters [25] and Mel-frequency cepstral coefficients [29]. FFT length and window overlapping parameters were chosen by taking into account a reasonable tradeoff between resolution and computational load (size of X) for the subsequent step.

2) **HMM training.** Given a C-class classification problem, a set of HMMs $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_C$ is trained, i.e., one per class. Only training sequences (spectrograms) have been used for learning the models. Gaussian HMMs have been used, i.e., HMMs where the emission probability is a Gaussian.[4]
3) **Generative Embedding**. Within this step, all the objects involved in the problem (namely training and testing sequences) are projected through the set of learnt models, to the generative embedding vector space. The different embeddings employed in this paper have been summarized in the previous section.
4) **Discriminative Classification**. Now, the problem of classifying sequences has been cast to a more standard problem of classifying points, for which any vector-based classifier can be used. As in many generative embedding applications, in our approach, we employed SVMs with an rbf kernel.

IV. Experimental Evaluation

In this section, the experimental evaluation is proposed. In particular, in Section IV-A, the seismic volcanic signals employed in this paper are presented. Experimental details are given in Section IV-B, whereas in the remaining sections, three different experiments, with the corresponding results and discussions, are presented.

A. Data: Acquisition, Types of Earthquakes and Preprocessing

The experiments were performed using signals coming from Galeras Volcano in Colombia. This volcano is a stratovolcano located at latitude $1°13'43.8''$ N and longitude $77° \ 21' \ 33''$ W in the Colombian Andes mountains. The active crater is situated about 7-km W from the city of Pasto, Colombia, and the summit reaches an elevation of 4270 m above sea level (m.a.s.l.) After more than four decades of dormancy, Galeras reawaked in 1988 and has had several major eruptions since then: in May 1989, in July 1992, several ones during the first semester of 1993, the second semester of 2004, the end of 2005, in January 2008, during the whole 2009, and in January and August 2010.

—**Recording system.** Earthquakes at Galeras Volcano are registered using a seismic network deployed by the Volcanological and Seismological Observatory at Pasto (OVSP). The network is composed by seven short-period seismic stations; namely Anganoy (ANGV), Cráter-2 (CR2R), Urcunina (URC), Cobanegra, Cóndor, Nariño-2, and Calabozo. In addition, there are two broadband stations called Cufiño and Obonuco. The ANGV station is the closest one to the active crater (0.8 km) and the highest one (4227 m.a.s.l). Signals are telemetered by radio from the station locations to the OVSP headquarters. The data acquisition system includes a 12-bit analog-to-digital converter with a sampling rate of 100.16 samples/s, an automatic detection/segmentation stage based on the STA/LTA algorithm, and a series of servers where segmented events are recorded according to the Seismic Unified Data System protocol.

—**Types of volcanic earthquakes.** Volcano-tectonic (VT) earthquakes, long-period (LP) events, tremors (TR), and hybrid (HB) events are the most important volcano-induced earthquakes. An example of each one is shown in Fig. 1, together with their corresponding spectrograms on the bottom part of each waveform. A detailed description of the mechanisms leading to these types of volcanic earthquakes is out of the scope of this paper. The interested reader is referred to comprehensive reviews such as [31]–[33].

B. Experimental Details

In our experimental evaluation, we performed three different sets of experiments (largely detailed in the following), in two different classification problems. The first classification problem involves three classes: VT earthquakes, LP earthquakes, and TR events. This is a well established and interesting classification task with many other examples in the literature. The second classification problem is a more challenging one, including four classes: the three considered above and the HB class. Adding this class of events makes the problem more complex since it is usually difficult to distinguish between LP and HB earthquakes [7], [34]. However, from a practical point of view, accurately identifying this class of earthquakes is one of the most challenging tasks for staff members at the observatory.

In all the experiments, the earthquake signals were characterized by the spectrograms, as described at the beginning of Section III. HMM training has been performed using the Baum–Welch procedure, stopping it after likelihood convergence. Initialization has been carried out, as in many applications, with a clustering based on Gaussian mixture models. The Gaussians in the emission probabilities were assumed to have a diagonal covariance matrix (due to the high dimensionality of the signal, full covariance matrices would have been poorly estimated). In the generative embedding space, SVMs were employed, using an rbf kernel. Parameters C and σ were estimated with a cross-validation (CV) procedure on the training set, as implemented in the PRTools MATLAB toolbox.[5] In all cases, classification ACC values were computed using averaged holdout CV, with results averaged over 20 repetitions.

[4]We are aware that the standard choice in the seismic community is to use a mixture of Gaussians; nonetheless, it has been formally proved in [30] that an HMM with a mixture of Gaussians in every state is equivalent to an HMM with more states having only one Gaussian per state. This removes one of the free parameters to be set, namely the number of Gaussians for the mixture.

[5]See http://www.prtools.org/.

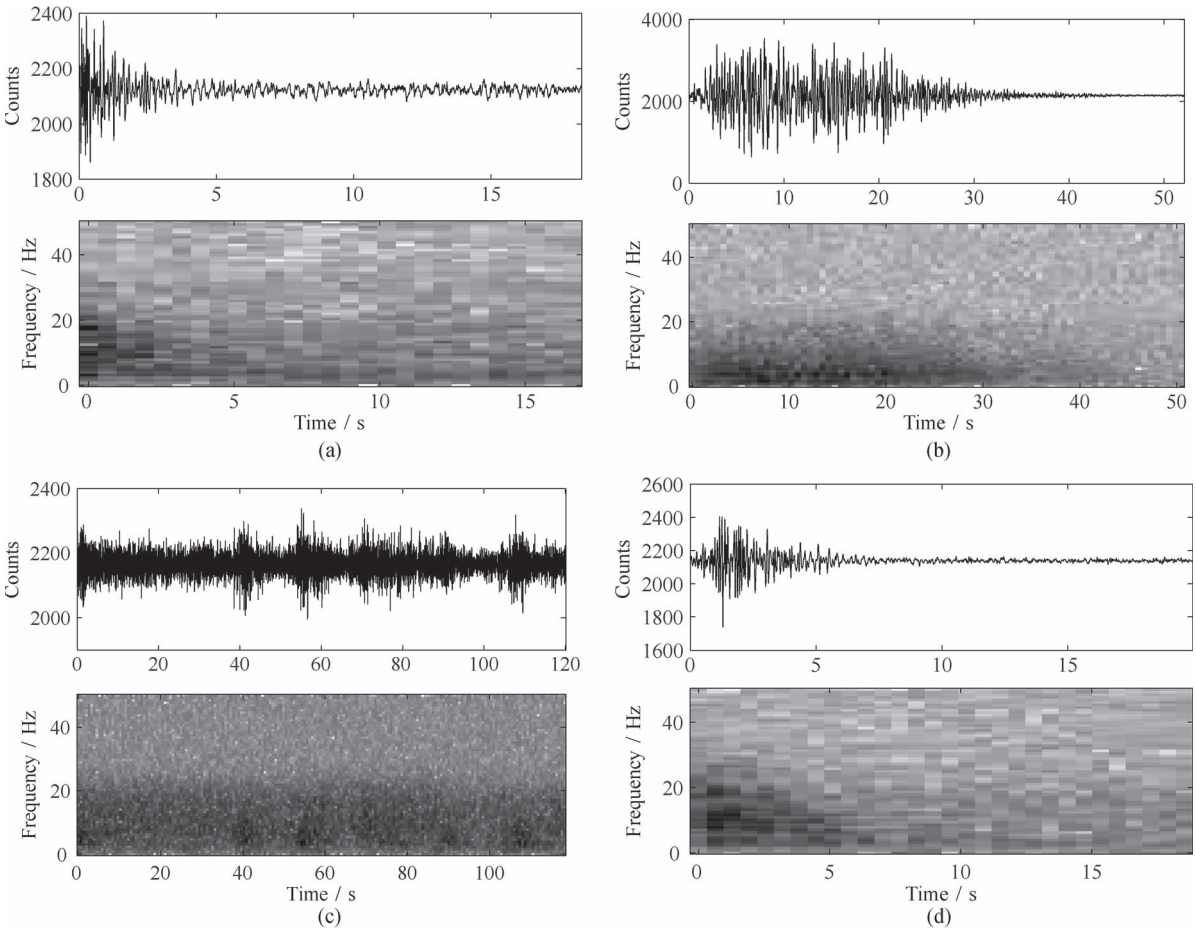

Fig. 1. Examples of volcanic earthquakes and their corresponding spectrograms. Signals were recorded at the ANGV station in Galeras Volcano. (a) VT event. (b) LP event. (c) TR event. (d) HB event.

TABLE I
AVERAGED ACC VALUES FOR THE BAYES RULE AND FOR THE PROPOSED APPROACH WHILE VARYING THE NUMBER OF STATES N AND THE EMBEDDINGS (SEE SECTION II FOR THE MEANING OF THE ACRONYMS): (a) THREE-CLASS PROBLEM. (b) FOUR-CLASS PROBLEM

N	Bayes	LLE	SE	TE	FSE
2	0.8723 (0.0011)	0.9460 (0.0009)	0.8540 (0.0012)	0.9230 (0.0015)	**0.9600** (0.0008)
3	0.8880 (0.0011)	0.9550 (0.0009)	0.9157 (0.0016)	0.9200 (0.0014)	0.9500 (0.0011)
4	0.9113 (0.0011)	0.9587 (0.0009)	0.9340 (0.0015)	0.9427 (0.0012)	0.9503 (0.0011)
5	**0.9150** (0.0011)	0.9553 (0.0009)	0.9347 (0.0010)	0.9417 (0.0012)	0.9323 (0.0016)
6	0.8740 (0.0075)	0.9510 (0.0010)	0.9390 (0.0011)	0.9297 (0.0013)*	0.9327 (0.0014)*
7	0.8263 (0.0103)	0.9473 (0.0012)	0.9390 (0.0011)	0.9187 (0.0010)	0.9297 (0.0014)
8	0.8313 (0.0104)	0.9463 (0.0012)	0.9443 (0.0010)	0.9033 (0.0016)*	0.9250 (0.0014)
9	0.7860 (0.0121)	0.9413 (0.0015)	0.9367 (0.0011)	0.9033 (0.0012)	0.9130 (0.0015)
10	0.5617 (0.0151)	0.9233 (0.0045)	0.9360 (0.0015)	0.8860 (0.0018)	0.9033 (0.0021)
15	0.3600 (0.0069)	0.8940 (0.0034)	0.9313 (0.0011)	0.8820 (0.0012)	0.8910 (0.0016)

(a)

N	Bayes	LLE	SE	TE	FSE
2	0.7807 (0.0014)	0.8297 (0.0010)	0.7265 (0.0010)	0.7678 (0.0014)*	**0.8520** (0.0011)
3	0.8000 (0.0011)	0.8313 (0.0008)	0.7718 (0.0013)	0.7675 (0.0013)	0.8495 (0.0012)
4	**0.8348** (0.0012)	0.8448 (0.0015)*	0.7980 (0.0009)	0.7735 (0.0011)	0.8437 (0.0009)*
5	0.8347 (0.0012)	0.8417 (0.0013)*	0.7833 (0.0014)	0.7662 (0.0015)	0.8212 (0.0012)*
6	0.8225 (0.0040)	0.8298 (0.0024)*	0.7812 (0.0012)	0.7685 (0.0014)	0.8242 (0.0014)*
7	0.7993 (0.0048)	0.8355 (0.0014)*	0.7820 (0.0012)*	0.7600 (0.0014)*	0.8063 (0.0015)*
8	0.6860 (0.0117)	0.8043 (0.0021)	0.7888 (0.0011)*	0.7575 (0.0014)*	0.7987 (0.0015)
9	0.5825 (0.0143)	0.8008 (0.0019)	0.7720 (0.0013)	0.7402 (0.0015)	0.7782 (0.0017)
10	0.5685 (0.0132)	0.8088 (0.0023)	0.7800 (0.0014)	0.7403 (0.0011)	0.7680 (0.0014)
15	0.2790 (0.0063)	0.7365 (0.0024)	0.7750 (0.0013)	0.7212 (0.0012)	0.7630 (0.0015)

(b)

In the remaining part of this section, the three experiments are detailed. In particular, in the first one, a comprehensive comparison between the proposed approach and the standard Bayes rule is provided by varying the number of HMM states and the different embeddings. An analysis of the confusion matrices for the best cases is also provided. In the second

experiment, we investigate the impact of the dimension of the data set for the two different approaches, via learning curves. Finally, in the third experiment, we compare the two approaches with respect to their generalization capability when changing the recording station.

C. Experiment 1: Comparison

In this first experiment, we compare the proposed approach with the standard HMM-based Bayes rule on the two classification problems. We randomly selected 100 events per class, resulting in problems with 300 and 400 events, respectively. Following the averaged hold-out CV protocol, these data sets have been then randomly divided into two parts: one used for training and one for testing (with the process repeated 20 times). We employed the four generative embeddings described in Section II. In order to assure a fair comparison, the same HMMs were used for the Bayes rule and for the embeddings. Averaged classification ACC values (together with the standard errors of the mean, i.e., between brackets) for the two classification problems are shown in Table I, for different number of HMM states N.

In order to get an idea of the statistical significance of the differences in the table, we performed a standard t-test comparing the CV results, with a significance level of 5%. In the table, we put an asterisk $(*)$ in those entries where the hypothesis "the averaged ACC of the corresponding generative embedding and that of the Bayes rule are equal" cannot be rejected with a confidence level of 5% (in other and simpler words, an asterisk in a generative embedding column indicates that there is no statistically significant difference within that generative embedding and the Bayes result).

From the tables, some comments can be evinced:

- As a general comment, we can notice the beneficial impact of the generative embedding scheme in the HMM classification. This is more evident in the three-class problem, where almost all the classification ACC values obtained with the generative embedding schemes are statistically significantly better than those obtained with the Bayes rule. In the four-class problem, this trend is maintained, although in a less pronounced way.
- The behavior of the results when varying the number of states confirms the robustness of the generative embedding scheme with respect to poor models. The Bayes rule suffers from a bad choice of the parameter N (for ten states, the performances are around 56%; with 15 states, the performances are even worse), whereas the generative embedding scheme does not suffer too much. Even in case of nonproper models, the discriminative part of the scheme succeeds in finding reasonable boundaries, possibly due to the still descriptive information distilled by the generative part.
- By looking at the behavior of the different embeddings, we can notice that the best results are obtained with LLE and with FSE. In particular, FSE seems to be better with small models, whereas LLE always maintains a reasonably good behavior. In fact, the best overall results in both problems were obtained with FSE and the smallest models. This is somehow expected since the FSE space has a dimension linked to the numbed of parameters (whereas LLE has a dimension linked to the number of classes in the problem). Therefore, by increasing too much the number of states, it is likely that the highly dimensional FSE space induces the well-known curse of the dimensionality problem.

TABLE II
AVERAGED CONFUSION MATRICES AND PERFORMANCE MEASURES (TP, FP, AND ACC, GIVEN IN DETECTION RATE) OF THE PROPOSED APPROACH AND OF THE BAYES RULE IN THE DIFFERENT CLASSIFICATION PROBLEMS. (a) BAYES RULE (THREE-CLASS PROBLEM). (b) FSE (THREE-CLASS PROBLEM). (c) BAYES RULE (FOUR-CLASS PROBLEM). (d) FSE (FOUR-CLASS PROBLEM)

(a)

		Predicted			TP	FP	ACC
		VT	LP	TR			
Actual	VT	0.924	0.025	0.000	0.974	0.037	0.974
	LP	0.076	0.925	0.104	0.837	0.040	0.837
	TR	0.000	0.050	0.896	0.947	0.051	0.947

Overall accuracy = 0.9150

(b)

		Predicted			TP	FP	ACC
		VT	LP	TR			
Actual	VT	1.000	0.027	0.015	0.960	0.000	0.960
	LP	0.000	0.922	0.027	0.972	0.038	0.972
	TR	0.000	0.051	0.958	0.949	0.021	0.949

Overall accuracy = 0.9600

(c)

		Predicted				TP	FP	ACC
		VT	LP	TR	HB			
Actual	VT	0.784	0.008	0.000	0.165	0.819	0.071	0.819
	LP	0.025	0.881	0.101	0.060	0.826	0.041	0.826
	TR	0.000	0.052	0.899	0.000	0.945	0.033	0.945
	HB	0.191	0.059	0.000	0.775	0.756	0.076	0.756

Overall accuracy = 0.8348

(d)

		Predicted				TP	FP	ACC
		VT	LP	TR	HB			
Actual	VT	0.767	0.013	0.002	0.200	0.781	0.077	0.781
	LP	0.003	0.883	0.013	0.016	0.965	0.038	0.965
	TR	0.000	0.052	0.977	0.003	0.947	0.008	0.947
	HB	0.230	0.052	0.008	0.781	0.729	0.075	0.729

Overall accuracy = 0.8520

It is interesting, from an application point of view, to better understand the behavior of the proposed approach with respect to the different classes. In order to do that, we reported the confusion matrices for both the generative and Bayes approaches. For the sake of interpretation, we decided to show only the confusion matrices relative to the best generative embedding result (FSE with a two-state model, for both problems) and for Bayes (five- and four-state models, for the first and the second problems, respectively). The obtained confusion matrices, averaged over the 20 runs of the CV experiment, are displayed in Table II.

In the three-class problem, the most remarkable result is the 13% difference, i.e., in true positive rate (TP) and ACC,

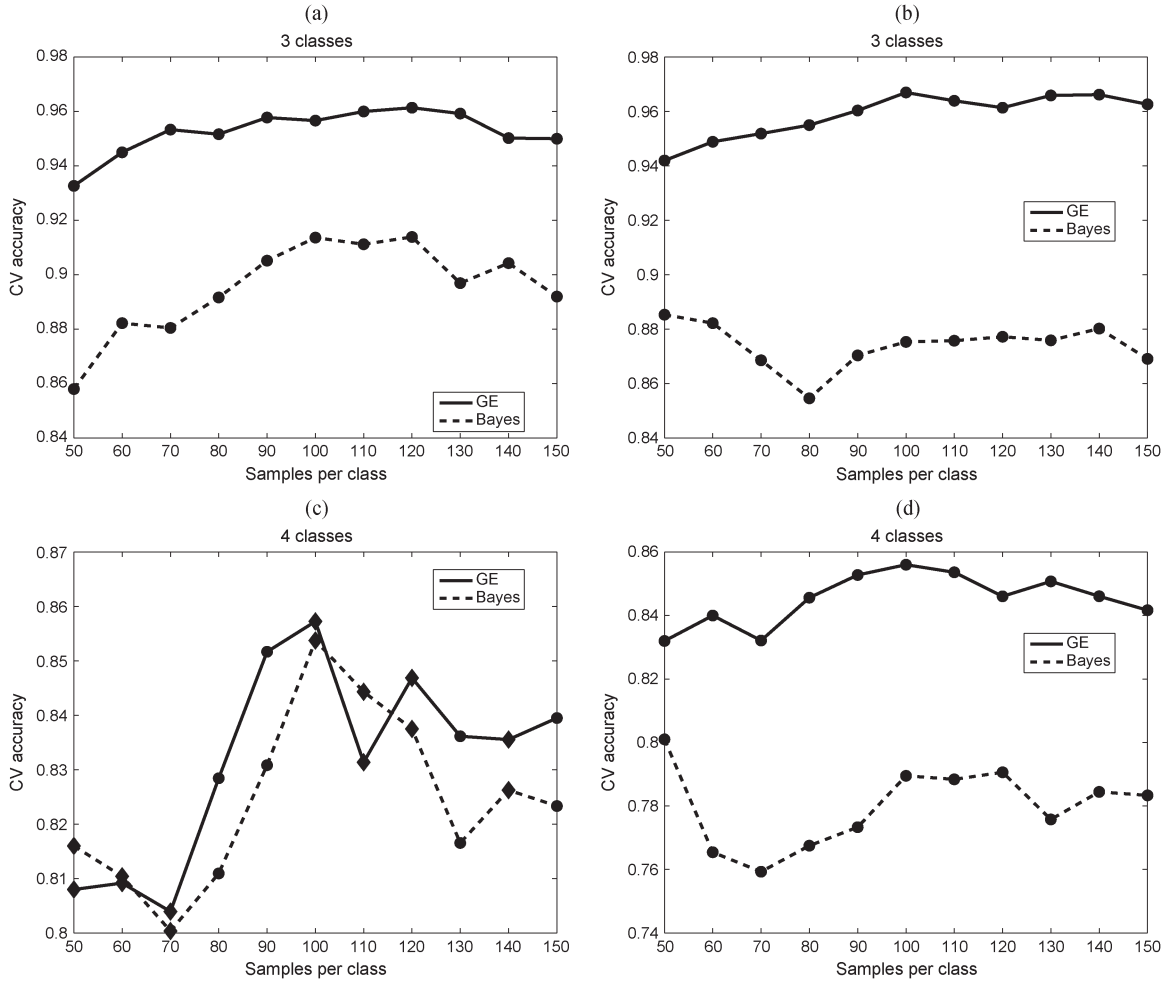

Fig. 2. Learning curves for the proposed approach and the Bayes rule, in two different situations and for the different problems. (a) Three-class problem: Best parameters for the Bayes rule. (b) Three-class problem: Best parameters for the generative embedding approach. (c) Four-class problem: Best parameters for the Bayes rule. (d) Four-class problem: Best parameters for the generative embedding approach.

between the FSE and the Bayes rule for the LP class. Such a result reveals a great effect of the proposed approach on the ability to distinguish LP signals, as confirmed by the reduction of the number of LP events wrongly identified as VT or TR ones. In detail, confusions of LP events as VT events were reduced from 7.6% to 0%, and similarly, confusions of LP events as TR events decreased from 10.4% to 2.7%. As a result of this improvement, the false positive (FP) detections of VT and TR events decreased in 4% and 3%, respectively.

Recognition ACC of LP signals also significantly increased with the FSE embedding in the four-class problem (from 83% to 97%), which is consistent with an equal percentage increase in TP for that class. In both approaches, HB signals are frequently confused with VT ones. That confusion is slightly lower in the opposite relation, i.e., VT events wrongly recognized as HB ones. In contrast with the increase in ACC for the LP class, a 4% performance deterioration for the VT class is observed when applying the FSE. In spite of that, an improvement of the overall ACC is achieved by the proposed approach.

D. Experiment 2: Learning Curves

In this experiment, we tested the proposed approach with respect to different sizes of the data set. In order to do that, we repeated the aforementioned experiments while changing the number of events per class from 50 to 150. In this set of experiments, we select the best generative embedding from the previous analysis (FSE). In order to be fair, we considered two situations. We set the number of states as the best for the Bayes rule (five- and four-state models, for the first and second problems, respectively) and as the best for the proposed approach (two-state model). Learning curves are shown in Fig. 2. In order to get an idea of the statistical significance, we performed again the t-test explained in the previous section (with a confidence level of 5%). In the plot, a diamond indicates that the corresponding values are not statistically significantly different; they are present only in the third plot.

From the plot, it is evident that the proposed approach outperforms the Bayes rule, except for two particular cases of the four-class problem with the approaches tuned with the best parameters for the Bayes rule [see Fig. 2(c)]: 1) at the beginning; and 2) in about the middle of the studied range for the training set size.

E. Experiment 3: Cross-Station Experiment

In this experiment, the generalization capability of the proposed approach is tested in a very challenging scenario, namely

TABLE III
AVERAGED ACC VALUES OF THE BAYES RULE AND OF THE PROPOSED APPROACH IN THE DIFFERENT PROBLEMS OF THE CROSS-STATION EXPERIMENT. (a) THREE-CLASS PROBLEM, TWO STATES. (b) THREE-CLASS PROBLEM, FIVE STATES. (c) FOUR-CLASS PROBLEM, TWO STATES. (d) FOUR-CLASS PROBLEM, FOUR STATES

(a)

		Test					
		Bayes			Generative Embedding		
		ANGV	URC	CR2R	ANGV	URC	CR2R
Training	ANGV	0.8833	0.6407	0.3960	0.9460	0.5480	0.4840
	URC	0.4120	0.8993	0.5527	0.7953	0.9573	0.6200
	CR2R	0.4187	0.3320	0.9480	0.6127	0.5893	0.9813

(b)

		Test					
		Bayes			Generative Embedding		
		ANGV	URC	CR2R	ANGV	URC	CR2R
Training	ANGV	0.8733	0.6847	0.3720	0.9020	0.6087	0.4353
	URC	0.4580	0.9080	0.3340	0.5467	0.8833	0.5333
	CR2R	0.4393	0.3213	0.8007	0.5860	0.6160	0.9593

(c)

		Test					
		Bayes			Generative Embedding		
		ANGV	URC	CR2R	ANGV	URC	CR2R
Training	ANGV	0.7745	0.4250	0.2290	0.8560	0.5195	0.4045
	URC	0.3670	0.7750	0.5305	0.6695	0.8155	0.5285
	CR2R	0.2925	0.2365	0.8300	0.5120	0.5340	0.8870

(d)

		Test					
		Bayes			Generative Embedding		
		ANGV	URC	CR2R	ANGV	URC	CR2R
Training	ANGV	0.8070	0.5110	0.2245	0.8415	0.4855	0.3625
	URC	0.4345	0.7965	0.3840	0.5710	0.8065	0.5915
	CR2R	0.3510	0.2730	0.7195	0.5195	0.5540	0.8825

using data coming from different stations. It is reasonable that several stations are positioned on a volcano, observing and registering the same seismic events. Clearly, registered signals are different due to the so-called source, path, and local site effects that introduce time delays (distance from the source) and amplifications or attenuations of signal components at certain frequencies (local geology acting as a filter) [3]. They may also differ due to differences between sensor and preprocessing–transmission technologies. The usefulness of multiple stations in the volcano signal classification task has been recently studied [35], particularly from the multiple-classifier-system perspective.

The idea here is to train the classifier system using signals coming from one station and testing it with signals from a different one (hereafter cross-station operating condition). Such condition may be of great practical impact, e.g., in cases when a new station is set (and no training data are available) or when a station is out of service, e.g., due to preventive maintenance, power outage caused by ash/snow accumulation on the solar panels, damage caused by lightning or vandalism, and either a nearby or a portable recording station is used as a temporary replacement.

In our experiment, we used signals recorded at three different stations, i.e., ANGV, URC, and CR2R. The results, for the three-class problem and the four-class problem are reported in Table III (again, setting the number of states as the best for the Bayes rule and as the best for the proposed approach). A summary is presented in Table IV, where all off-diagonal elements were averaged to give an immediate idea.

TABLE IV
SUMMARY OF AVERAGED ACC VALUES OF THE BAYES RULE AND OF THE PROPOSED APPROACH IN THE DIFFERENT PROBLEMS OF THE CROSS-STATION EXPERIMENT: ALL OFF-DIAGONAL ELEMENTS HAVE BEEN AVERAGED

Problem	Bayes Accuracy	GE Accuracy
Three-class (2 states)	0.4587	0.6082
Three-class (5 states)	0.4349	0.5543
Four-class (2 states)	0.3467	0.5280
Four-class (4 states)	0.3630	0.5140

It is evident from Table IV that the proposed approach is more robust to the change of the station than the Bayes rule, thus confirming the suggestions given in the literature (see, e.g., [23]) that hybrid generative–discriminative schemes are more robust with respect to poor models (as those estimated in a different station can be).

By carefully looking in Table III, we can observe that even if the ANGV station is the one used by seismologists at OVSP as the main reference to classify the events, the best results for systems trained and tested in the same station are generally obtained for signals from CR2R. Another point is that the best cross-station operating condition result (training with URC signals and testing with ANGV signals) is 0.7953, obtained with the proposed approach. This is remarkably high, considering the challenging nature of the performed experiment and opening the possibility of the realization of a classification system that can be independent from the actual station from which the signals were gathered; in this sense, HMMs are seen as feature extractors more than being simple classifiers. When used as classifiers, HMMs do not go over 0.6847 of ACC.

V. Conclusion

In this paper, a novel HMM-based system to classify seismic volcanic signals has been presented, based on a generative embedding scheme. We have shown, through an extensive set of experiments, that the proposed approach significantly improves the performances of the standard HMM-based schemes that rely on the Bayes decision rule. Among the four different generative embeddings that were applied, FSE with the smallest models (two states) always yielded the best results. Most of the performance improvement can be attributed to the ability of the proposed approach to distinguish LP seismic signals. We also proved that the HMM-based generative embedding approach is a preferable alternative to the standard HMM-based Bayes rule when training the models with signals recorded in one station but testing it, if needed due to technical contingencies, on signals from a different one.

Appendix
Abbreviations

ACC	accuracy.
ANGV	Anganoy station.
CR2R	Cráter-2 station.
CV	cross-validation.
FFT	fast Fourier transform.
FP	false positive rate.
FSE	Fisher score embedding.
HB	hybrid.
HMM	hidden Markov model.
LLE	log-likelihood embedding.
LP	long period.
LTA	long-term average.
OVSP	acronym in Spanish for Volcanological and Seismological Observatory at Pasto.
rbf	radial basis function.
SE	state embedding.
STA	short-term average.
SVM	support vector machine.
TE	transition embedding.
TP	true positive rate.
TR	tremor.
URC	Urcunina station.
VT	volcano–tectonic.

Acknowledgment

The authors would like to thank the staff members from the OVSP for providing the data set.

References

[1] P. Lesage, "Interactive Matlab software for the analysis of seismic volcanic signals," *Comput. Geosci.*, vol. 35, no. 10, pp. 2137–2144, Oct. 2009.

[2] L. Ottemöller, P. Voss, and J. Havskov, "SEISAN: Earthquake analysis software for Windows, SOLARIS, Linux and MacOSX," Dept. Earth Sci., Univ. Bergen, Bergen, Norway, May 2011, version 9.0.1.

[3] M. Orozco-Alzate, C. Acosta-Muñoz, and J. M. Londoño-Bonilla, "The automated identification of volcanic earthquakes: Concepts, applications and challenges," in *Earthquake Research and Analysis—Seismology, Seismotectonic and Earthquake Geology*, S. D'Amico, Ed. Rijeka, Croatia: InTech, Feb. 2012, ch. 19, pp. 345–370.

[4] M. Ohrnberger, "Continuous Automatic Classification of Seismic Signals of Volcanic Origin at Mt. Merapi, Java, Indonesia," Ph.D. dissertation, Univ. Potsdam, Postdam, Germany, Apr. 2001.

[5] P. Alasonati, J. Wassermann, and M. Ohrnberger, "Signal classification by wavelet-based hidden Markov models: Application to seismic signals of volcanic origin," in *Statistics in Volcanology*, H. Mader, C. Connor, and S. Coles, Eds. Trowbridge, U.K.: Geological Society of London, 2006, ser. Special Publications of IAVCEI, no. 1, ch. 13, pp. 161–174.

[6] L. Gutiérrez, J. Ramírez, C. Benítez, J. Ibañez, J. Almendros, and A. García-Yeguas, "HMM-based classification of seismic events recorded at Stromboli and Etna volcanoes," in *Proc. IEEE IGARSS*, 2006, pp. 2765–2768.

[7] M. C. Benítez, J. Ramírez, J. C. Segura, J. M. Ibáñez, J. Almendros, A. García-Yeguas, and G. Cortés, "Continuous HMM-based seismic-event classification at Deception Island, Antarctica," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 138–146, Jan. 2007.

[8] M. Beyreuther, R. Carniel, and J. Wassermann, "Continuous hidden Markov models: Application to automatic earthquake detection and classification at Las Cañadas caldera, Tenerife," *J. Volcanol. Geotherm. Res.*, vol. 176, no. 4, pp. 513–518, Oct. 2008.

[9] M. Beyreuther and J. Wassermann, "Continuous earthquake detection and classification using discrete hidden Markov models," *Geophys. J. Int.*, vol. 175, no. 3, pp. 1055–1066, Dec. 2008.

[10] L. Gutiérrez, J. Ibañez, G. Cortés, J. Ramírez, C. Benítez, V. Tenorio, and I. Álvarez, "Volcano-seismic signal detection and classification processing using hidden Markov models. Application to San Cristóbal volcano, Nicaragua," in *Proc. IEEE IGARSS*, Jul. 2009, vol. 4, pp. IV–522–IV–525.

[11] J. M. Ibáñez, C. Benítez, L. A. Gutiérrez, G. Cortés, A. García-Yeguas, and G. Alguacil, "The classification of seismo-volcanic signals using Hidden Markov Models as applied to the Stromboli and Etna volcanoes," *J. Volcanol. Geothermal Res.*, vol. 187, no. 3/4, pp. 218–226, Nov. 2009.

[12] M. Beyreuther and J. Wassermann, "Hidden semi-Markov model based earthquake classification system using weighted finite-state transducers," *Nonlin. Process. Geophys.*, vol. 18, no. 1, pp. 81–89, 2011.

[13] M. Beyreuther, C. Hammer, J. Wassermann, M. Ohrnberger, and T. Megies, "Constructing a hidden Markov model based earthquake detector: Application to induced seismicity," *Geophys. J. Int.*, vol. 189, no. 1, pp. 602–610, Apr. 2012.

[14] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, May 1989.

[15] T. S. Jaakkola and D. Haussler, "Exploiting generative models in discriminative classifiers," in *Proc. 12th Annu. Conf. NIPS II 1998*, M. S. Kearns, S. A. Solla, and D. A. Cohn, Eds., 1999, pp. 487–493, Ser. Neural Information Processing.

[16] J. A. Lasserre, C. M. Bishop, and T. P. Minka, "Principled hybrids of generative and discriminative models," in *Proc. IEEE Comput. Soc. Conf. CVPR*, A. Fitzgibbon, C. J. Taylor, and Y. LeCun, Eds., Jun. 2006, vol. 1, pp. 87–94, IEEE Comput. Soc., Alamitos, CA.

[17] A. Y. Ng and M. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes," in *Proc. Adv. NIPS 14—15th Annu. Conf. NIPS*, vol. II, *Neural Information Processing Systems (NIPS) Foundation*, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds., 2002, pp. 841–848, Ser. Neural Information Processing.

[18] Y. D. Rubinstein and T. Hastie, "Discriminative vs informative learning," in *Proc. 3rd Int. Conf. KDD*, D. Heckerman, H. Mannila, D. Pregibon, and R. Uthurusamy, Eds., 1997, pp. 49–53, Association for the Advancement of Artificial Intelligence.

[19] K. Tsuda, M. Kawanabe, G. Rätsch, S. Sonnenburg, and K.-R. Müller, "A new discriminative kernel from probabilistic models," *Neural Comput.*, vol. 14, no. 10, pp. 2397–2414, Oct. 2002.

[20] A. Bosch, A. Zisserman, and X. Muñoz, "Scene classification via pLSA," in *Proc. 9th ECCV*, A. Leonardis, H. Bischof, and A. Pinz, Eds., May 2006, vol. 3954, pp. 517–530, Springer-Verlag, Berlin, Germany.

[21] A. Perina, M. Cristani, U. Castellani, V. Murino, and N. Jojic, "A hybrid generative/discriminative classification framework based on free-energy terms," in *Proc. 12th IEEE Int. Conf. Comput. Vis.*, 2009, pp. 2058–2065.

[22] M. Bicego, V. Murino, and M. A. T. Figueiredo, "Similarity-based classification of sequences using hidden Markov models," *Pattern Recognit.*, vol. 37, no. 12, pp. 2281–2291, Dec. 2004.

[23] M. Bicego, E. Pekalska, D. M. J. Tax, and R. P. W. Duin, "Component-based discriminative classification for hidden Markov models," *Pattern Recognit.*, vol. 42, no. 11, pp. 2637–2648, Nov. 2009.

[24] V. M. Zobin, *Introduction to Volcanic Seismology*. Amsterdam, The Netherlands: Elsevier, 2003, Ser. Developments in Volcanology.
[25] M. I.-V. Seht, "Detection and identification of seismic signals recorded at Krakatau volcano (Indonesia) using artificial neural networks," *J. Volcanol. Geothermal Res.*, vol. 176, no. 4, pp. 448–456, Oct. 2008.
[26] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.
[27] M. Bicego, M. Cristani, V. Murino, E. Pekalska, and R. P. W. Duin, "Clustering-based construction of hidden Markov models for generative kernels," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, D. Cremers, Y. Boykov, A. Blake, and F. Schmidt, Eds. Berlin, Germany: Springer-Verlag, 2009, ser. Lecture Notes in Computer Science, pp. 466–479.
[28] L. Chen, H. Man, and A. V. Nefian, "Face recognition based on multi-class mapping of Fisher scores," *Pattern Recognit.*, vol. 38, no. 6, pp. 799–811, Jun. 2005.
[29] I. Álvarez, L. García, G. Cortés, C. Benítez, and Á. De la Torre, "Discriminative feature selection for automatic classification of volcano-seismic signals," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 151–155, Mar. 2012.
[30] M. Bicego, V. Murino, and M. A. T. Figueiredo, "A sequential pruning strategy for the selection of the number of states in hidden Markov models," *Pattern Recognit. Lett.*, vol. 24, no. 9/10, pp. 1395–1407, Jun. 2003.
[31] B. A. Chouet, "Volcano seismology," *Pure Appl. Geophys.*, vol. 160, no. 3/4, pp. 739–788, 2003.
[32] S. R. McNutt, "Volcanic seismology," *Annu. Rev. Earth Planet. Sci.*, vol. 33, no. 1, pp. 461–491, 2005.
[33] H. Kawakatsu and M. Yamamoto, "Volcano seismology," in *Treatise on Geophysics*, G. Schubert, Ed. Amsterdam, The Netherlands: Elsevier, 2007, ch. 4.13, pp. 389–420.
[34] R. B. Trombley, *The Forecasting of Volcanic Eruptions*. Bloomington, IN: iUniverse, Sep. 2006.
[35] R. P. W. Duin, M. Orozco-Alzate, and J. M. Londoño-Bonilla, "Classification of volcano events observed by multiple seismic stations," in *Proc. 20th ICPR*, Aug. 2010, pp. 1052–1055.

Manuele Bicego (S'01–A'03–M'04) received the Laurea and Ph.D. degrees in computer science from the University of Verona, Verona, Italy, in 1999 and 2003, respectively.

From 2000 to 2003, he was with the University of Verona. From 2004 to 2008, he was with the University of Sassari, Sassari, Italy. From June 2009 to February 2011, he was a Team Leader and a member with the Pattern Analysis, Learning and Image Understanding Systems Laboratory, Italian Institute of Technology, Genoa, Italy. He is currently a Researcher with the University of Verona and a member of the Vision Image Processing and Sound Laboratory, University of Verona. His main research interests include probabilistic models for classification and clustering, biometrics, and video analysis.

Carolina Acosta-Muñoz received the degree in electronic engineering and the M.Eng. degree in industrial automation from Universidad Nacional de Colombia, Manizales, Colombia, in 2008 and 2010, respectively.

She is currently an Electronics Engineer with Observatorio Vulcanológico y Sismologico de Manizales (OVSM), Servicio Geológico Colombiano, where she is a member of the working groups on electronics and seismology. Her research interests include statistical techniques, multivariate data analysis, and pattern recognition.

Mauricio Orozco-Alzate (M'10) received the degree in electronic engineering, the M.Eng. degree in industrial automation, and the Dr.Eng. degree in automatics from Universidad Nacional de Colombia, Manizales, Colombia, in 2003, 2005, and 2008, respectively.

In 2007, he served as a Research Fellow with the Pattern Recognition Laboratory, Delft University of Technology, Delft, The Netherlands. Since August 2008, he has been with Departamento de Informática y Computación, Universidad Nacional de Colombia, Manizales, Colombia. His main research interests include pattern recognition, digital signal processing, and their applications to analysis and classification of seismic, bioacoustic, and hydrometeorological signals.