Available online at www.sciencedirect.com
SCIENCE DIRECT®
Sensors and Actuators B 110 (2005) 225–230
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

Odor classification using similarity-based representation

Manuele Bicego*

DEIR, University of Sassari, via Torre Tonda, Sassari 07100, Italy

Received 8 November 2004; accepted 27 January 2005
Available online 26 February 2005

Abstract

In this paper a new approach to odor classification is presented, founded on the similarity-based representation paradigm. The proposed approach builds a new representation space, called similarity space, in which each object is not represented by features, but by its similarities with respect to other objects in the data set. The classification step is performed using support vector machines, a technique introduced in the statistical learning theory context. One of the major drawbacks of the similarity-based representation paradigm is the dimensionality of the similarity space: a method for addressing this problem has been introduced in this paper, based on a notion of the unsupervised classification (clustering) theory, namely the medoid concept. The approach outperforms standard features-based representations on tests regarding data gathered from a chemical sensors array electronic nose.

Keywords: Similarity-based classification; Electronic nose; Odor recognition; Odor classification; Support vector machines; Carbon black-polymer sensors

1. Introduction

The problem of recognition and classification of odors is a challenging multidisciplinary research area, whose importance has impressively grown in the last decade, for both methodological [1,2] and applicative reasons [3,4]. Initial efforts in this field were performed by chemical researchers, in order to obtain adequate and reproducible sensors. In recent years, nevertheless, an increasing interest has been shown by the pattern recognition community, and several state of the art methodologies developed in this context have been applied to odor classification [5–7].

These approaches are mostly related to the classification aspects of the problem, that is, the problem of deciding the category of an odor given a representation of it. Typically, this representation strongly depends on the kind of sensors used (SAW, QCM, optical, polymeric, etc.), nevertheless resulting in a vector of features. In this paper we propose to employ an alternative representation scheme, namely the similarity-based representation [8–14] for odor classification. This paradigm, which has recently been introduced, differs from typical pattern recognition approaches where objects to be classified are represented by sets (vectors) of features. In the similarity-based paradigm, objects are described using pairwise (dis)similarities, i.e. distances from other objects in the data set. In this way, objects are not constrained to be explicitly represented in a feature space, and all that is necessary is a way to compute (dis)similarities between pairs of objects. Then the goal is to learn a classifier only from these relational data. The advantage is that, with this representation, the algorithm could be generic and independent from the actual data representation, allowing the use of non-metric similarities (thereby violating the triangular inequality). Further, this representation makes standard feature-based PR techniques applicable to problems that do not have a natural embedding to a uniform feature space, i.e. problems for which it is not possible to straightforwardly extract features, but it is easier to compute similarities, such as problems concerning images [15] or sequences [14,16].

The literature on similarity-based classification is not vast [8–14] (a brief review is given in Section 2.1). The general idea behind all these approaches is basically the same: given a set of pairwise dissimilarity values, a new representation

* Tel.: +39 079 2017323; fax: +39 079 2019016.
E-mail address: bicego@uniss.it (M. Bicego).

0167-0987/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.snb.2005.01.034

space can be built, in which each object is described by these values.

In this paper we propose to build a similarity[1] space, representing each object by the vector of its similarities with respect to a predetermined set of objects (this can be the whole data set, in the simplest approach), called the *representatives set*; the classification is then performed in this new representation space. In this paper, for the classification stage, we used support vector machines (SVMs) [17,18], a statistical learning tool which has been recently applied in several different contexts, with excellent results; SVM has been recently introduced also in the context of odor detection and classification, showing promising performances [19,20].

One problem to be addressed with the similarity-based approach is the dimensionality of the resulting space, which, in the basic approach, could be equal to the cardinality of the training set. The problem, and the related state of the art, is briefly summarized in Section 2.2. In this paper we proposed a solution to this problem, using the concept of medoid, a notion of the unsupervised classification (clustering) theory.

The proposed approach has been tested on a real case, regarding data gathered from an e-nose [7] composed by an array of carbon black-polymer detectors [21,22]. The system has been thoroughly described in [7], and briefly summarized in Section 4. Classification accuracies on standard features-based representation and on similarity-based representation have been compared, showing that the proposed approach is effective in odor classification. Moreover, the method for reducing the dimensionality of the similarity space has been tested, showing that it is able to make more tractable the space while maintaining reasonable accuracies.

The rest of the paper is organized as follows: the similarity-based approach is proposed in Section 3, together with the state of the art. The e-nose apparatus used for gathering data is briefly summarized in Section 4, together with the SVM-based classification strategy. Experimental results and discussions are presented in Section 5, and, finally, in Section 6 conclusions are drawn and future perspectives are envisaged.

2. State of the art

2.1. The similarity-based representation

The literature on similarity-based classification is not vast. The approach seems to have been first introduced by Jain and Zongker [8], who obtained a dissimilarity measure, based on deformable templates, for the handwritten digit recognition problem. A multidimensional scaling approach was then used to project this dissimilarity space onto a low-dimensional space, where a one-nearest-neighbor (1-NN) classifier was employed to classify new objects. In [9], Graepel et al. investigate the problem of learning a classifier based on data represented in terms of their pairwise proximities, using an approach based on Vapnik's structural risk minimization [23]. Jacobs and Weinshall [10] studied the use of distance-based classification with non metric distance functions (i.e. that do not satisfy the triangle inequality). Duin and Pekalska are very active researchers in this area,[2] having recently produced several papers [11–13]. Motivation and basic features of similarity-based methods were first described in [11]: it was shown, by experiments in two real applications, that a Bayesian classifier (the RLNC—regularized linear normal density-based classifier) in the dissimilarity space outperforms the nearest neighbor rule. These aspects were more thoroughly investigated in [13], where other classifiers in the dissimilarity space were studied, namely on digit recognition and bioinformatics problems. Finally, in [12], a generalized kernel approach was introduced, dealing with classification aspects of the dissimilarity kernels.

Recently, the similarity-based representation has been applied to other contexts: images [15], where the paradigm was used for determining feedback in image retrieval by content, and sequences [14,16], where the paradigm is combined with hidden Markov models [24], in order to make both supervised and unsupervised classification of sequential data.

2.2. The dimensionality issue

The main problem of the similarity-based approach, of particular relevance in practical applications, is the high dimensionality of the resulting similarity space. Two types of solutions have been proposed in order to address this problem. The first consists of building the similarity space using all available patterns, and subsequently applying some standard dimensionality reduction technique. One example of this kind of approach is the multidimensional scaling method used in [8]. Another recent example is presented in [25], where a reduction of the dimensionality of the dissimilarity space is obtained by a modified multidimensional scaling scheme, able to reduce the computational burden and allow generalization to new data. The second type of solution works by directly choosing a small set of representatives. An example of this type of solution can be found in [13], where random selection, *most-dissimilar* rule, and the *condensed nearest neighbor* (CNN) rule were employed. Other examples can also be found in [10], where a new type of CNN method is proposed, in [26], where a greedy approach is proposed, able to find prototypes encoding the principal components of the similarity space, or in [14], where two approaches were proposed: a matching pursuit approach [27], used to determine the representative most "useful" for classification, and a "one per class" approach, able to identify one representative for each class of the training set.

[1] Note that we refer indifferently to similarity or dissimilarity.

[2] See http://www.ph.tn.tudelft.nl/Research/neural/index.html.

3. The similarity-based representation paradigm

The idea at the basis of the proposed approach is conceptually simple: to build a new representation space, using the similarity values between objects, and construct a classifier in that space. One of the justifications for this approach lies in the fact that similarity is high for similar objects, i.e. belonging to the same class, and low for objects of different classes, making discrimination possible [12]. Therefore, we can interpret the similarity measure $\mathcal{D}(\mathbf{X}, \mathbf{X}_i)$ between an object $\mathbf{X}$ and another "reference" object $\mathbf{X}_i$ as a "feature" of the object $\mathbf{X}$. This fact suggests the construction of a feature vector for $\mathbf{X}$ by taking the similarities between $\mathbf{X}$ and a set of reference objects $\mathcal{R} = \{\mathbf{X}_k\}$, so that $\mathbf{X}$ is characterized by a *pattern* (i.e. a set of features) $\{\mathcal{D}(\mathbf{X}, \mathbf{X}_k),\ \mathbf{X}_k \in \mathcal{R}\}$. Notice that the fact that two objects, say $\mathbf{X}_i$ and $\mathbf{X}_j$, present similar degrees of similarity to several other objects (e.g., they are both very similar to some objects, and also both very dissimilar to some other objects) enforces the hypothesis that $\mathbf{X}_i$ and $\mathbf{X}_j$ belong to the same class.

3.1. Formal definition

Formally, the proposed strategy is defined as follows. Consider a classification problem with C classes; for each class $k \in \{1, 2, \ldots, C\}$, we have a set of N_k training objects $\mathcal{T}_k = \{\mathbf{X}_1^{(k)} \cdots \mathbf{X}_{N_k}^{(k)}\}$; thus $N = \sum_k N_k$ is the total size of the training set $\mathcal{T} = \bigcup_{k=1}^{C} \mathcal{T}_k$.

Let $\mathcal{R} = \{\mathbf{P}_1, \ldots, \mathbf{P}_R\}$ be a set of R "reference" or "representative" objects; these objects may belong to the set of training objects ($\mathcal{R} \subseteq \mathcal{T}$) or may be otherwise defined. Now let $\mathcal{D}_{\mathcal{R}}(\mathbf{X})$ be a function that returns the vector of similarities between an arbitrary sequence $\mathbf{X}$ and all the sequences in $\mathcal{R}$, that is

$$\mathcal{D}_{\mathcal{R}}(\mathbf{X}) = \begin{bmatrix} \mathcal{D}(\mathbf{X}, \mathbf{P}_1) \\ \vdots \\ \mathcal{D}(\mathbf{X}, \mathbf{P}_R) \end{bmatrix} \in \mathbb{R}^R \tag{1}$$

We will designate the space $\mathbb{R}^R$ in which the dissimilarity vector exists as the "similarity space" and denote it as $\mathcal{S}_{\mathcal{R}}$, where the subscript $\mathcal{R}$ is used to emphasize the dependence of the similarity space on the set $\mathcal{R}$. Once this similarity space is defined, any standard classifier can, in principle, be used.

3.2. The dimensionality issue

Regarding the choice of $\mathcal{R}$, different approaches can be adopted; the basic one is to choose $\mathcal{R} = \mathcal{T}$, the whole training set. With this choice, the dimensionality of $\mathcal{S}_{\mathcal{R}} = \mathcal{S}_{\mathcal{T}}$ is equal to N, the cardinality of the training set $\mathcal{T}$. This is obviously a problem, because it makes the proposed method inapplicable in most cases; nevertheless it is interesting to investigate the discrimination ability of this space.

As summarized in previous section, the problem could be addressed in two ways: the first is to create the whole similarity space, subsequently reducing it using some dimensionality reduction technique (as principal component analysis, for example). The second is to directly choose the representative objects. The approach proposed in this paper belongs to the latter class, and proposes to choose one or more representative for each category, trying to determine the "most" descriptive for each class. The concept of "most" descriptive is derived from the PAM (partitioning around medoid) method [28], an algorithm used for clustering data: the descriptor of each cluster is the element nearest to the centroid of the group, called medoid. Nevertheless, in our representation we cannot directly compute the medoid, since we cannot compute the centroid: we only have distances, not features. The problem is solved by defining the medoid as the most "central" pattern of the class, i.e. the object with the minimum distance to all the other patterns. A similar concept could be found in the DPAM (distance partition around medoid) algorithm, proposed in [29] to perform HMM-based clustering of sequences. Please note that this definition is in some sense similar to the definition of the median, which is the central element in a list of ordered numbers. More formally, we define the medoid $md^0(k)$ of the class C_k as:

$$md^0(k) = \min_{\mathbf{X} \in C_k} \left(\sum_{\mathbf{X}_j \in C_k} \mathcal{D}(\mathbf{X}, \mathbf{X}_j) \right) \tag{2}$$

The representative set becomes:

$$\mathcal{R} = MD^0 = \{md^0(0), md^0(1), \ldots, md^0(C)\} \tag{3}$$

With this set, one element for each class was chosen, reducing the dimensionality of the similarity space from the number of pattern to the number of classes. This method could be generalized, introducing in the $\mathcal{R}$ set not only the medoid, but also the "second" medoid (the second element most central), or the third. We could have different representative sets MD^M, of increasing dimensionality, which include the $M + 1$ most central elements in the class. More formally, we could inductively define these sets as:

$$MD^M = MD^{M-1} \bigcup \{md^M(0), md^M(1), \ldots, md^M(C)\}$$

where

$$md^M(k) = \min_{\mathbf{X} \in C_k \mathbf{X} \notin \{md^0(k), \ldots, md^{M-1}(k)\}} \left(\sum_{\mathbf{X}_j \in C_k} \mathcal{D}(\mathbf{X}, \mathbf{X}_j) \right)$$

We call this method the "medoid" approach, where M represents the "order" of the medoid. Using as representative set MD^M, we obtain a similarity space of dimensionality equal to $(M + 1)C$, reducing the dimensionality from the number of elements in the training set to a factor of the number of classes.

4. The e-nose

In this section the electronic nose is briefly introduced: firstly, the sensors are described, and the instrumental apparatus is briefly sketched. Secondly, the classification strategy based on support vector machines is introduced.

4.1. Carbon black-polymer sensors based electronic nose

The data used in this paper were obtained using the electronic nose described in [7], and here summarized. The sensor device is a chemical sensors array, composed by eight different types of sensors. Each sensor is a carbon black-polymer detector [21,22], a particular kind of conducting polymer sensors [30,31]. Briefly, the individual sensor elements were constructed from films consisting of carbon black particles dispersed into insulating organic polymers. The carbon black endows electrical conductivity to the films (chemical diversity among elements in the array was obtained using different organic polymers for each sensor). Swelling of polymer during solvent exposure increases film resistance; by this way we could simply and efficiently monitor the presence of vapor of interest. The lack of reproducibility of those sensors imposes that the subsequent analysis should be carefully performed, using flexible and sophisticated techniques. These sensors were organized in an array, connected to a gathering PC. A typical experiment consisted of a three step process, beginning with a 5 minutes of air flow (in order to determine the sensor baseline resistance), followed by a variable amount of time of odor exposure (up to 10 minutes) and by 10 minutes of air flow, in order to recover the baseline resistance value. For each experiment, the maximum relative variation of each sensor resistance was used as feature, composing a feature vector of length 8. Moreover, in order to test the minimum time needed by the e-nose for resolving vapor tasks, we compute the maximum relative variation of the resistance after 1, 2, ... seconds from the vapor exposure. These values, gathered together, forms the so called "photo sets", where the name recalls the way they are obtained, that is making a "photo" of the sensors situation after 1, 2, ... seconds. For further information please refer to [7].

4.2. The classification strategy

Data coming from the e-nose have been classified using support vector machines, a classification tool derived from the statistical learning theory [17,23,18]. SVMs have been successfully employed in a wide range of applications in the recent years, with successful performances. They have been chosen due to their high generalization capability, and to their major ability to deal with high dimensionality space, such as that resulting from the proposed approach. SVMs are not fully described here, and an exhaustive general introduction can be found in [18], while its use in the e-nose context has been investigated in [19,20].

The basic SVM scheme relies to binary classification and in order to deal with multi-class problems a generalized scheme should be introduced. In this paper we have employed the method called *SVM 1-vs-1 Max Win*, proposed in [32]. This scheme trains one SVM for each pair of classes. Given an unknown pattern, all SVMs are evaluated, counting for each class the number of wins. The pattern is assigned to the class with the maximum number of wins.

The SVM parameters have been determined using a cross-validation averaged holdout procedure [33]: the data set has been randomly split in two mutually exclusive parts, one used for training and one for testing. This process has been repeated several times (50 repetitions), and classification accuracies have been computed for different parameter configurations. The parameters leading to the best classification accuracies have been chosen. These experiments, not reported here, have shown that the best kernel is the radial basis function kernel in every case, result confirmed by the literature [34–36], while the corresponding optimal σ and C vary depending on the classification tasks.

5. Experimental results and discussion

This section describes experimental evaluation of the proposed approach. The aim is to compare standard features-based representation with the proposed similarity-based representation using SVM classification. The classification accuracies have been determined using the averaged holdout procedure described in previous section, comparing the two approaches on the same sets. The 2 approaches has been tested on 2 sets, both composed by 102 elements and 3 classes, with 34 elements each, deriving from the same experiment. The e-nose has been exposed to 2-propanol (with a concentration of 5.8 ppth[3]), acetone (22.62 ppth) and ethanol (5.7 ppth). The difference between the first and the second set regards the time exposure: in the former case sensor answers were gathered after 10 minutes of exposure, while in the latter case they were gathered after 1 second of exposure, making a "photo" of the sensor situations. We called the first set "Whole" set and the second "Photo" set. Clearly the second task is more difficult than the first, but more challenging: recognizing an odor after few seconds of exposure could have great practical implications.

The first analysis was performed on the "Whole" set: for the similarity-based approach, the distances have been computed using the Euclidean metric. The representative set $\mathcal{R}$ was equal to the whole training set $\mathcal{T}$. Classification accuracies are shown in Table 1: for each SVM, the used parameters have been also reported. "Standard" stands for the standard feature based approach, while "$\mathcal{S}_\mathcal{T}$" stands for the similarity-based method, using as representative set the whole training set $\mathcal{T}$. From the table it is evident that both approaches' accuracy perform about perfectly on this set. Since there is no

[3] Parts per thousand.

Table 1
Classification accuracies on the "Whole" set

Representation space	Parameters		Classification accuracy (%)
	σ	C	
Standard	0.02	100	99.02
$\mathcal{S}_{\mathcal{T}}$	0.02	120	99.02

Table 2
Classification accuracies on the "Photo" set, using Euclidean metric

Representation space	Parameters		Classification accuracy (%)
	σ	C	
Standard	0.02	200	78.76
$\mathcal{S}_{\mathcal{T}}$	0.04	140	80.52

chance of improving results (it is difficult to capture differences between the two approaches if the accuracy is about perfect), we concentrated our analysis on the "Photo" set, which is more difficult and challenging, and for which a thorough analysis has been carried out.

The first analysis was again using the Euclidean metric for computing distance in the similarity-based approach, while the representative set $\mathcal{R}$ was equal to the whole training set $\mathcal{T}$. Results are proposed in Table 2, following the same notations of Table 1. Looking to this table we could notice that the proposed approaches is more effective than the standard approach, resulting in an improvement of about 2%. We could also notice that this classification task is harder, as expected, since the classification accuracies are reduced.

The second analysis was about the metric used to compute the similarity space: our definition does not rely to a particular metric, and could be used starting from any pairwise similarity matrix. In particular, non metric similarity functions could also be used, as for example probabilistic values [14,16]. In this analysis, we repeated the previous experiment, performed with the Euclidean metric, using the Manhattan metric, which, given two vectors $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_p)$, is defined as

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{p} |x_i - y_i|$$

Results are shown in Table 3: we could note that also in this case the similarity-representation could be quite useful for improving the classification accuracies.

One of the most severe problem of the similarity-based representation paradigm is the dimensionality of the resulting similarity space, which, in the basic approach, is equal to the cardinality of the training set: this could result in a severe curse of dimensionality problem [37]. We performed

Table 3
Classification accuracies on the "Photo" set, using Manhattan metric

Representation space	Parameters		Classification accuracy (%)
	σ	C	
Standard	0.02	200	80.00
$\mathcal{S}_{\mathcal{T}}$	0.10	140	81.76

Table 4
Classification accuracies on the "Photo" set with the "medoid" approach, using Euclidean metric

Representation space	Parameters		Classification accuracy (%)
	σ	C	
$\mathcal{S}_{\mathcal{T}}$	0.04	140	80.52
$\mathcal{S}_{MD^0}$	0.02	180	74.18
$\mathcal{S}_{MD^1}$	0.02	200	78.56
$\mathcal{S}_{MD^2}$	0.02	200	79.22
$\mathcal{S}_{MD^3}$	0.02	120	79.48
$\mathcal{S}_{MD^4}$	0.02	140	79.80

Table 5
Classification accuracies on the "Photo" set with the "medoid" approach, using Manhattan metric

Representation space	Parameters		Classification accuracy (%)
	σ	C	
$\mathcal{S}_{\mathcal{T}}$	0.10	140	81.76
$\mathcal{S}_{MD^0}$	0.02	120	74.18
$\mathcal{S}_{MD^1}$	0.04	200	78.04
$\mathcal{S}_{MD^2}$	0.04	140	80.00
$\mathcal{S}_{MD^3}$	0.04	200	80.85
$\mathcal{S}_{MD^4}$	0.08	160	81.24

some experiments in order to test the "medoid" system introduced in Section 3.2. Results are shown in Tables 4 and 5, for Euclidean and Manhattan metrics, respectively. The "medoid" approach was applied with an order increasing from 0 to 4: following the notation introduced in the paper, the corresponding similarity space are denoted as $\mathcal{S}_{MD^0}, \mathcal{S}_{MD^1}, \ldots, \mathcal{S}_{MD^4}$.

From these tables we could observe that the medoid system represents a quite effective system for reducing the dimensionality of the similarity space, except for the order equal to 0 or 1. In that last case the similarity space is a three (or six)-dimensional space, which probably is a too reduced version of the original space. Nevertheless, for orders greater than 2, the proposed approach reaches a performance which is inferior than the one in the original similarity space, but still outperforming standard feature-based representation approach.

As a last consideration, we could say that the similarity-based representation seems to be suitable for odor classification, improving the classification accuracy. To be really employable, nevertheless, one has to employ classification strategies which do not suffer too much from the curse of dimensionality problem, as support vector machines, or to develop techniques able to reduce the dimensionality of the similarity space, as the medoid system proposed in this paper.

6. Conclusions

In this paper a new approach to odor classification has been presented, founded on the similarity-based representation: the method proposes to build a new representation space, in which each object is represented by the vector of similarities to other objects in the data set. The classification step

is then performed using support vector machines. The proposed approach has been tested on real data gathered from an electronic nose, showing performances that outperform standard approaches. A method for reducing the dimensionality of the similarity space has been also proposed in this paper, able to reduce that dimensionality without affecting the classification accuracies.

Acknowledgements

The author is very grateful to Dr. G.E. Felis for useful discussions and final reading of the paper.

References

[1] J.W. Gardner, P.N. Bartlett, Electronic Noses: Principles and Applications, Oxford University Press, 1999.
[2] C.D. Natale, A. D'amico, F. Davide, Redundancy in sensors arrays, Sens. Actuators A 37–38 (1993) 612–617.
[3] B. Borrounet, T. Talou, A. Gaset, Application of multigas sensor device in the meat industry for boar-taint detection, Sens. Actuators B 26–27 (1995) 250–254.
[4] M. Schweizer-Berberich, S. Vaihinger, W. Göpel, Characterization of food freshness with sensor arrays, Sens. Actuators B (18–19) (1994) 282–290.
[5] P.C. Jurs, G.A. Bakken, H.E. McClelland, Computational methods for the analysis of chemical sensor array data from volatile analytes, Chem. Rev. 100 (2000) 2649–2678.
[6] R. Shaffer, S. Rose-Pehrsson, R. McGill, A comparison study of chemical sensor array pattern recognition algorithms, Anal. Chim. Acta 384 (1999) 305–317.
[7] M. Bicego, G. Tessari, G. Tecchiolli, M. Bettinelli, A comparative analysis of basic pattern recognition techniques for the development of small size electronic nose, Sens. Actuators B 85 (1–2) (2002) 137–144.
[8] A. Jain, D. Zongker, Representation and recognition of handwritten digits using deformable templates, IEEE Trans. Pattern Anal. Mach. Intell. 19 (12) (1997) 1386–1391.
[9] T. Graepel, R. Herbrich, P. Bollmann-Sdorra, K. Obermayer, Classification on pairwise proximity data, in: D.C.M. Kearns, S. Solla (Eds.), Advances in Neural Information Processing, vol. 11, MIT Press, 1999.
[10] D. Jacobs, D. Weinshall, Classification with nonmetric distances: image retrieval and class representation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (6) (2000) 583–600.
[11] E. Pekalska, R. Duin, Automatic pattern recognition by similarity representations, Electron. Lett. 37 (3) (2001) 159–160.
[12] E. Pekalska, P. Paclik, R. Duin, A generalized kernel approach to dissimilarity-based classification, J. Mach. Learning Res. 2 (2) (2002) 175–211.
[13] E. Pekalska, R. Duin, Dissimilarity representations allow for building good classifiers, Pattern Recog. Lett. 23 (8) (2002) 943–956.
[14] M. Bicego, V. Murino, M. Figueiredo, Similarity-based classification of sequences using hidden Markov models, Pattern Recog. 37 (12) (2004) 2281–2291.
[15] G. Giacinto, F. Roli, Dissimilarity representation of images for relevance feedback in content-based image retrieval, in: P. Perner, A. Rosenfeld (Eds.), Machine Learning and Data Mining in Pattern Recognition, vol. LNAI2734, Springer, 2003, pp. 202–214.
[16] M. Bicego, V. Murino, M. Figueiredo, Similarity-based clustering of sequences using hidden Markov models, in: P. Perner, A. Rosenfeld (Eds.), Machine Learning and Data Mining in Pattern Recognition, vol. LNAI2734, Springer, 2003, pp. 86–95.
[17] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, 1995.
[18] C. Burges, A tutorial on support vector machine for pattern recognition, Data Min. Know. Disc. 2 (1998) 121–167.
[19] D. DeCoste, M. Burl, Support vector machines and kernel fisher discriminants: A case study using electronic nose data, in: Proceedings of the Fourth Workshop on Mining Scientific Datasets (KDD-2001), 2001.
[20] C. Distante, N. Ancona, P. Siciliano, Support vector machines for olfactory signals recognition, Sens. Actuators B 88 (1) (2003) 30–39.
[21] M. Lonergan, E. Severin, B. Doleman, S. Beaber, R. Grubbs, N. Lewis, Array-based vapor sensing using chemically sensitive carbon black-polymer resistors, Chem. Mater. 8 (1996) 1198–1312.
[22] B. Doleman, M. Lonergan, E. Severin, T. Vaid, N. Lewis, Quantitative study of the resolving power of arrays of carbon black-polymer composites in various vapor-sensing tasks, Anal. Chem. 70 (1998) 4177–4190.
[23] V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[24] L. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE 77 (2) (1989) 257–286.
[25] E. Pekalska, R. Duin, Spatial representation of dissimilarity data via lower-complexity linear and non linear mappings, in: T. Caelli, A. Amin, R. Duin, M. Kamel, D. de Ridder (Eds.), Structural, Syntactic and Statistical Pattern Recognition, vol. LNCS 2396, Springer, 2002, pp. 488–497.
[26] E. Pekalska, R. Duin, Prototype selection for finding efficient representations of dissimilarity data, in: R. Kasturi, D. Laurendeau, C. Suen (Eds.), Proceedings of the International Conference on Pattern Recognition, vol. 3, 2002, pp. 37–40.
[27] S. Mallat, Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (1993) 3397–3415.
[28] A. Jain, R. Dubes, Algorithms for Clustering Data, Prentice-Hall, 1988.
[29] A. Panuccio, M. Bicego, V. Murino, A Hidden Markov Model-based approach to sequential data clustering, in: T. Caelli, A. Amin, R. Duin, M. Kamel, D. de Ridder (Eds.), Structural, Syntactic and Statistical Pattern Recognition, vol. LNCS 2396, Springer, 2002, pp. 734–742.
[30] J. Hatfield, P. Neaves, P. Hicks, K. Persaud, P. Travers, Towards an integrated electronic nose using conducting polymer sensors, Sens. Actuators B 18 (1994) 221–228.
[31] M. Freund, N. Lewis, A chemically diverse conducting polymer-based electronic nose, Proc. Natl. Acad. Sci. USA-Chem. 92 (1995) 2652–2656.
[32] J. Friedman, Another approach to polychotomous classification, Technical Report, Stanford University, 1996.
[33] S. Theodoridis, K. Koutroumbas, Pattern Recognition, Academic Press, 1999.
[34] M. Pontil, A. Verri, Support vector machines for 3-D object recognition, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1998) 637–646.
[35] M. Bicego, G. Iacono, V. Murino, Face recognition with multilevel B-splines and support vector machines, in: Proceedings of the ACM SIGMM Multimedia Biometrics Methods and Applications Workshop, 2003, pp. 17–24.
[36] V. Murino, M. Bicego, I. Rossi, Statistical classification of raw textile defects, in: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR04), vol. 4, 2004, pp. 311–314.
[37] R. Duda, P. Hart, D. Stork, Pattern Classification, 2nd ed., Wiley, 2001.

Biography

Manuele Bicego received his Laurea degree and Ph.D. degree in Computer Science from University of Verona in 1999 and 2003, respectively. From 2000 to 2003 he worked in VIPS (Vision, Image Processing and Sound) laboratory of the Computer Science Department of University of Verona (Italy). Since 2004 he is a PostDoc fellow at University of Sassari (Italy). His research interests include statistical pattern recognition, artificial vision, electronic noses, neural networks, hidden Markov models, video analysis and biometrics. Manuele Bicego is member of the IAPR-IC society and of the IEEE Systems, Man, and Cybernetics society.