Knowledge-Based Systems 91 (2016) 252–262

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Fast multi-class recognition of piecewise regular objects based on sequential three-way decisions and granular computing

A.V. Savchenko*

National Research University Higher School of Economics, Laboratory of Algorithms and Technologies for Network Analysis, 136 Rodionova Ulitsa, Nizhny, Novgorod 603093, Russia

ARTICLE INFO

Article history:
Received 21 January 2015
Revised 16 September 2015
Accepted 18 September 2015
Available online 28 September 2015

Keywords:
Statistical pattern recognition
Sequential three-way decisions
Granular computing
Image recognition
PHOG (pyramid histogram of oriented gradients)

ABSTRACT

The paper is focused on an application of sequential three-way decisions and granular computing to the problem of multi-class statistical recognition of the objects, which can be represented as a sequence of independent homogeneous (regular) segments. As the segmentation algorithms usually make it possible to choose the degree of homogeneity of the features in a segment, we propose to associate each object with a set of such piecewise regular representations (granules). The coarse-grained granules stand for a low number of weakly homogeneous segments. On the contrary, a sequence with a large count of high-homogeneous small segments is considered as a fine-grained granule. During recognition, the sequential analysis of each granularity level is performed. The next level with the finer granularity is processed, only if the decision at the current level is unreliable. The conventional Chow's rule is used for a non-commitment option. The decision on each granularity level is proposed to be also sequential. The probabilistic rough set of the distance of objects from different classes at each level is created. If the distance between the query object and the next checked reference object is included in the negative region (i.e., it is less than a fixed threshold), the search procedure is terminated. Experimental results in face recognition with the Essex dataset and the state-of-the-art HOG features are presented. It is demonstrated, that the proposed approach can increase the recognition performance in 2.5–6.5 times, in comparison with the conventional PHOG (pyramid HOG) method.

1. Introduction

Three-way decisions (TWD) [1] is an emergent theory, which is expected to be widely used in many knowledge-based systems. It has grown from the key idea of the rough set theory [2], namely, the division of the universal set of the analyzed objects into positive, negative and boundary regions. In practice, the probabilistic rough sets with variable precision [3] are used to introduce an uncertainty in the definition of these regions. Though the theory of TWD is well-studied [1,4–6], its application in practical tasks is still under research [7].

Pattern recognition [8] and data mining [9] are one of the most crucial research area, where TWD can be obviously used. There are several successful applications of TWD in such data mining tasks, as clustering [10,11], attribute reduction [12,13] and recommender systems [14]. However, the usage of TWD in the classification problem is still under an active research [15]. In the classification task it is required to assign a query object X to one of C classes, specified by the database of the reference (model) objects $\{X_r\}$, $r \in \{1, \dots, R\}$. It is assumed, that the class label $c(r) \in \{1, \dots, C\}$ of the rth model object is known. In fact, three options of acceptance, rejection and non-commitment are well interpreted in the *binary* classification task ($C = 2$). It includes three kinds of decision: positive (accept the first class), negative (reject the first class and accept the second class), and boundary (delay the final decision and do not accept either first or second class). Unlike the traditional two-way decision, the three-way decision incorporates the delay decision as an optional one. It is selected, if the cost of such delay is minimal [1]. For example, in a still-to-still video-based image recognition [16], it may not be possible to classify the video frame, reliably. Hence, the decision may be delayed and the next frame is analyzed.

The idea of TWD can be enhanced to *multi-class* recognition ($C > 2$). In this case, we have $(C + 1)$-way decisions: accept any of C classes or delay the decision process, in case of an unreliable classification result [17]. It is of great importance in practice, besides taking a hard decision, to allow such a reject option ("I do not know"). Statistically optimal rule for such a reject option was introduced by Chow in [18]. Nevertheless, as it is mentioned in [19], this option is often ignored in the statistical literature. On the other hand, in the engineering

Abbreviations: FAR, False-Accept Rate; HOG, Histograms of Oriented Gradients; LBP, Local Binary Patterns; PHOG, Pyramid Histogram of Oriented Gradients; SVM, Support Vector Machine; TWD, Three-Way Decisions.

* Corresponding author. Tel.: +79030434003.
E-mail address: avsavchenko@hse.ru

http://dx.doi.org/10.1016/j.knosys.2015.09.021
0950-7051/

community the reject option is more common and empirically shown to effectively reduce the misclassification rate [20–22].

Though TWD is incorporated in the pattern recognition task as a reject option, its combination with granular computing [23–25] in a hierarchical classification system [26] seems to be very promising [27,28]. If the analyzed object is complex (e.g., the facial image which consists of eyes, a nose and mouth regions), its examination at various granulation levels can be used to significantly improve the recognition quality. It is the key idea of the convolution neural networks, which have been recently proved to reach the quality of human recognition in several image classification tasks [29]. One of the most practically important examples of the potential of granular computing here, is the PHOG (Pyramid Histogram of Oriented Gradients) image descriptor [30]. According to this method, an image is divided at each granularity level into a regular grid of blocks and the HOG (Histogram of Oriented Gradients) is calculated in each block. The pyramid of grids is built, and the HOGs of the query and the model objects are matched at each level of the hierarchy (pyramid). The final distance between objects is calculated as a weighted sum of distances between the HOGs in the pyramid. It was shown [30] that this method achieves 10%-higher accuracy, in comparison with the state-of-the-art HOG method [31]. Unfortunately, the performance of such a hierarchical approach is usually much worse, in comparison with the conventional non-hierarchical algorithms. For instance, the dimensionality of the PHOG descriptor is usually 2–3 times higher, than the dimensionality of the HOG. Hence, despite the lower number of blocks in the coarse levels, the PHOG recognition time is 1.3–2 times higher, than the classification time of the conventional HOG.

Thus, the insufficient classification speed of such hierarchical systems is a crucial problem. An obvious way to improve the performance is to use a sequential analysis [32], and, in particular, sequential TWD [6,33]. Namely, an object is analyzed at many granularity levels. If the decision at the higher granularity level is reliable enough, i.e., the reject option has not been chosen, the classification process is terminated. Only if it is impossible to obtain a reliable solution at the current level, an object is analyzed in a more detailed way at the next level. Though this idea is quite obvious, it is necessary to address at least three questions:

- How to define granularity levels in a general way, in practically important pattern recognition tasks, so that high granularity levels are processed faster, than the lower one? For instance, the paper [15] introduces sequential TWD for face recognition problem. The facial image at each granularity level is represented with an eigenface [34], i.e. a variable granulation technique is used [23]. Though, this approach definitely allows to increase the recognition accuracy by the fusion of decisions at all levels, the classification performance at every level is the same.
- In contrast with conventional sequential TWD [33], the most reliable decision in pattern recognition is not necessary obtained at the low granularity level in all cases. How to define the way to make a final decision if the reject option is chosen at the last level?
- If the number of classes C is high, then the recognition speed will be slow even at the high granularity level [35]. Is it possible to apply sequential TWD to deal with this issue?

In this paper we suggest our answers to all these questions. First, the objects of interest are detected with any known object detection algorithm [36]. We assume, that the analyzed objects contain several independent homogeneous (regular, "stationary") parts. We will call such objects *piecewise regular*. Hence, each object can be divided into a sequence of homogeneous parts by an arbitrary segmentation method [37]. Among the known segmentation algorithms, a researcher can choose the region growing, contour detection algorithms or a simple object's division into non-overlapping frames of the fixed size. Each segment is represented as a sample of independent identically distributed feature vectors [38]. In this case, the granularity level is defined by the parameters of the segmentation algorithm. As such algorithms usually make it possible to choose the degree of homogeneity of the features inside a segment, we associate each object with a set of such piecewise regular representations (granules). Fine-grained granules stand for a sequence with a large count of high-homogeneous small segments. On the contrary, sets of a low number of weakly homogeneous large samples are considered as a coarse-grained granule. The recognition procedure at each level is implemented with the statistical approach [8] by matching the distributions of each segment. In this case, it is known that the Bayesian decision is equivalent to the nearest neighbor method with the Kullback–Leibler divergence [39]. The Chow's rule is used to reject an unreliable solution and move to the finest granularity level [18].

Second, if decisions at all levels are unreliable, we propose to implement the classifier fusion with the maximal posterior probability method [8]. In fact, the least unreliable decision is chosen.

Third, we perform the sequential analysis at each hierarchical level. Namely, the probabilistic rough set is created for the distance of objects from different classes. If the distance between the query object and the next model is included in the negative region (i.e., it is less than a fixed threshold), the search procedure is terminated and, as a result, this model is returned. In fact, this approach is integrated in various approximate nearest-neighbor methods in the range queries [40,41].

The rest of the paper is organized in the following way: Section 2.1 briefly presents the application of granular computing to statistical recognition of piecewise regular objects. In Section 2.2, sequential TWD is explored. Section 2.3 deals with the problem of insufficient recognition speed in case of the large number of classes. In Section 3, we present the experimental study of the proposed approach in face recognition task with the Essex dataset. Finally, concluding comments are given in Section 4.

2. Materials and methods

2.1. Recognition of piecewise regular objects with granular computing

To simplify the further discussion, we assume that an object of interest X is detected by any known algorithm. For instance, Viola–Jones cascade classifier [36] can be used in image processing. For clarity, let us define the query object X as M-dimensional vector $[x_1, \ldots, x_M]$. Suppose $\mathbf{X} \subset \mathbb{R}^M$ is the finite nonempty set of the analyzed objects.

According to the theory of granular computing [33,42], we analyze L levels of granularity $\{1, 2, \ldots, L\}$, with 1 representing the ground level and L the coarsest granularity. In this article we apply the segmentation procedure to the object X at every level, i.e., the query object X at the lth level is described by $K^{(l)} \geq 1$ homogeneous (regular) segments $X^{(l)} = \{X^{(l)}(k) | k \in \{1, \ldots, K^{(l)}\}\}$. Each kth segment is defined by its boundaries $(m_1^{(l)}(k), m_2^{(l)}(k))$, i.e., the kth segment is represented as a vector $[x_{m_1^{(l)}(k)}, \ldots, x_{m_2^{(l)}(k)}]$. We assume, that the segments are non-overlapped and it is possible, that several points of the original vector are not included in any segment, i.e. $1 \leq m_1^{(l)}(1) < m_2^{(l)}(1) < m_1^{(l)}(2) < \cdots < m_1^{(l)}(k) < m_2^{(l)}(k) < \cdots < m_2^{(l)}(K^{(l)}) \leq M$. The most simple way to divide X into a set of regular segments: $m_1^{(l)}(k) = 1 + \lfloor (M \cdot (k-1))/K^{(l)} \rfloor$, $m_2^{(l)}(k) = \lfloor (M \cdot k)/K^{(l)} \rfloor$, where $\lfloor \cdot \rfloor$ is the floor round function. In fact, more complex segmentation algorithms can be applied. For example, region-growing, clustering, histogram-based methods, etc. are widely used in image processing [37].

We assume, that each granularity level differs in the parameters of segmentation algorithm. The more homogeneous the resulted