ADAPTIVE REORDERING OF OBSERVATION SPACE TO IMPROVE PATTERN RECOGNITION

JULIUSZ L. KULIKOWSKI

Institute of Biocybernetics and Biomedical Engineering,
Polish Academy of Sciences,
4 Ks. Trojdena Str., 02-109 Warsaw, Poland
jlkulik@ibib.waw.pl

(Received received 02 January 2007; revised manuscript received 27 January 2007)

Abstract: The problem of observation space reordering is presented as a novel approach to pattern recognition based on non-parametric, combinatorial statistical tests. It consists in linearly ordering the elements of a discrete multi-dimensional observation space along a curve such that elements belonging to different similarity classes are as close to each other as possible, the similarity classes are mutually separated, and the length of the curve is kept to minimum. The problem is NP-difficult and it is shown how its approximate solution can be reached by a series of transformations improving the initial lexicographic linear order of a discrete observation space. Recommendations are formulated for linear order improvement leading to a pattern recognition algorithm based on serial statistical test.

Keywords: pattern recognition, serial statistical tests, linear ordering, permutations

1. Introduction

The commonly known pattern recognition methods can be roughly divided into two classes: (1) analytical methods based on calculating similarity between the recognized objects (formulated in the terms of linear or angular distance, logical, syntactic or structural consistency, *etc.*) and (2) methods based on concepts of artificial neural networks. In both cases, models of similarity classes of objects in a multi-dimensional observation space separated by hyper-planes are widely used [1–3]. They are particularly useful for examination of non-supervised pattern recognition methods based on Bayesian, distance or correlation approaches, as well as supervised pattern recognition learning (potential functions, *k-nearest neighbors*, *etc.*), providing that well mathematically-defined, homogenous, metric observation spaces are considered. The latter condition is satisfied if, in particular, the objects are represented by matrices or vectors consisting of components to which a common meaning can be assigned and which, as a consequence, can be measured in the same physical or geometrical units. Otherwise, if elements of the observation space represent various object parameters, the problem of scaling arises. For example, it is not clear what is the meaning of an Euclidean distance between two vectors used in medical diagnosis whose components are: x_1 – the patient's age [years], x_2 – the patient's weight [kg], x_3 – systolic blood pressure [hPa], x_4 – diastolic blood pressure [hPa], *etc.* There also arises a question whether it is better to express the blood pressure in hPa or in mmHg, as their relative influence on the measure of distance between the vectors depends significantly on the choice of units. Additionally, it may happen that some Boolean components should be considered, *e.g.* x_5 – *was the patient hospitalized due to cardiac infarct* [YES, NO]. It is evident that in such cases (rather typical in computer-aided medical diagnosis) an observation space $\boldsymbol{X}$ consisting of pseudo-vectors like $\boldsymbol{x} = [x_1, x_2, x_3, x_4, x_5]$ satisfy neither the metric space nor the linear vector space assumptions. Even if the pseudo-vectors' components have been normalized, there remain the problems of choosing the normalization coefficients and of their influence on the quality of pattern recognition. Pattern recognition methods neglecting these problems lead to algorithms which can be used only in local, strictly defined areas, decisions made under different assumptions being incomparable.

Pattern recognition methods based on the concept of linear ordering of observation space enable overcoming the above-mentioned difficulties. This is connected with the fact that a linear order remains invariant with respect to any continuous transformations of observation space components. The methods originated as an attempt to use non-parametric statistical tests for pattern recognition given limited primary information about the similarity classes. Combinatorial serial tests have been found to be well-suited to such situations [4, 5].

The idea can be easily illustrated geometrically. A 2-dimensional observation space and three similarity classes of objects represented by training sets S_1 (whose elements are denoted by ×), S_2 (denoted by •) and S_3 (denoted by +) are shown in Figure 1. The training sets are considered as mutually disjoint subsets of a total reference set, S, unclassified elements being denoted by ∘.

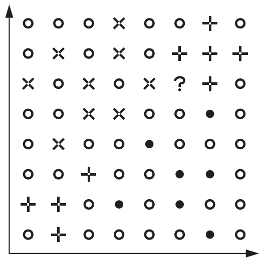

Figure 1. Example of a 2D observation space with elements representing three similarity classes

If a new-observed element, ?, is to be recognized (*i.e.* assigned to one of the above-mentioned classes) then, using the *k-NN* approach and a Manhattan metric, its 8-connective neighborhood should be examined. It contains 3 unclassified elements ∘, 1 element ×, 3 elements + and 1 element •. Therefore, ? will be recognized (for $k < 4$) as an element of S_3 (*i.e.* as +).

Using a serial statistical test requires linear ordering to be introduced to the observation space. This can be achieved in a number of ways, the simplest one based on lexicographical ordering of elements, 1st in rows, 2nd in columns. Starting from the lowest left element, we can observe that classified (non-$\circ$) elements occur in the following order:

$$+ \bullet + + \bullet \bullet + \bullet \bullet \times \bullet \times \times \bullet \times \times \times \; ? \; + \times \times + + + \times +$$

where the position of the unknown element, **?**, has also been indicated. It should be recognized as an element minimizing the number of homogenous sub-series in the given sequence of elements. As long as the **?** element is neglected, the number of sub-series is 16. Including **?** into the sequence leads to the following possible solutions (the corresponding numbers of sub-series are given in brackets):

$$\begin{array}{l} + \bullet + + \bullet \bullet + \bullet \bullet \times \bullet \times \times \bullet \times \times \times \boxed{\times} + \times \times + + + \times + \;(16)\\ + \bullet + + \bullet \bullet + \bullet \bullet \times \bullet \times \times \bullet \times \times \times \boxed{\bullet} + \times \times + + + \times + \;(17)\\ + \bullet + + \bullet \bullet + \bullet \bullet \times \bullet \times \times \bullet \times \times \times \boxed{+} + \times \times + + + \times + \;(16) \end{array}$$

Therefore, it can be concluded that in the given case **?** can be recognized as either $\times$ or $+$, the number of series in both cases being the same, 16. Let it be stressed that:

1. the method can be applied to any non-homogenous multi-dimensional observation space (*i.e.* representing data of varying formal nature),
2. it is independent of observation space scaling,
3. suitable for any finite number of similarity classes (easily merged into higher-order classes or split into similarity subclasses) and
4. easily implemented on computers, while
5. storage of the learning sets' elements in computer memory is not required (the rules of the observation space's linear ordering are stored instead).

However, there arises a problem of linear ordering of the observation space adequate to the given pattern recognition problem, which will be considered in detail below. The present paper is an extension of a poster presented at the XXII Autumn Meeting of the Polish Information Processing Society in Wisla, 2006 [6]. It is organized as follows. The impact of linear ordering of an observation space on the pattern recognition efficiency is discussed in Section 2. The problem of optimization of the linear order of an observation space based on a given family of reference subsets is analyzed in Section 3. Section 4 is a presentation of methods of step-wise linear order improvement (instead of its optimization) based on reversion, shifting and segmental permutation of sub-segments of linearly-ordered segments of elements, including several theorems and corollaries justifying these methods. Finally, conclusions are presented in Section 5.

2. Impact of linear ordering on pattern recognition efficiency

In this section, the influence of linear ordering introduced into an observation space on the effectiveness of pattern recognition will be illustrated. For this purpose the 2D observation space and training sets shown in Figure 1 will be considered.

Examples of linear ordering are shown in Figure 2: (a) lexicographical, (b) reversible sequential, (c) diagonal and (d) spiral. The (a), (b) and (c) orderings can be transformed into alternative orderings, based on different direction of observation space scanning by symmetrical reflection with respect to a vertical axis and rotation by the $-90°$ angle; they are respectively denoted by (a'), (b') and (c'). Corresponding sequences of training elements and the numbers of homogenous sub-series can be calculated for the formerly given training sets S_1, S_2 and S_3 and the defined linear orderings:

(a) + • + + • • + • • × • × × • × × × + × × + + + × + (16)
(a') + × + + × × + × × • × × × • × • • + • • • + + + + (14)
(b) + × × × + + + × × × × × • • × + • • • • • + + + + (9)
(b') + • • • + + + • • • × × × • + × × × × × + + + + × (10)
(c) • • • • • • + + • + + + + + × × + + × × × × × × × (7)
(c') + + + × + • × × • • • • × × × • × × × • + + + + + (11)
(d) + × × + + • + + × × × + + + • • • • + × × × • × • (13)

In the given case, ordering (c), leading to the lowest number of sub-series (7) is the best; ordering (b) (9 sub-series) is also satisfactory. However, neither (c) nor (b) is an optimal ordering, their minimum possible number of sub-series being 3 – the number of distinguished similarity classes. This minimum should be taken into account as the aim of optimization of linear ordering of the observation space.

The methods of linear ordering shown in Figure 2, as defined by simple geometrical rules, can be called regular. They can easily be extended to any discrete finite-dimensional observation spaces. Other, slightly more sophisticated, regular ordering methods, preferring local observation space scanning and based on a Hilbert general concept of curves filling compact geometrical areas, have been presented in [5].

However, the numbers of series generated by simple regular orderings for given reference sets are usually far from the minimum and thus arises the problem of linear ordering optimization. A heuristic solution of this problem based on the hyper-cube permutation approach has been proposed in [7]; in this paper a more general approach to the problem is presented. In particular, choosing the best linear ordering is considered as an optimization problem which should take into account both the low implementation complexity costs and the high pattern recognition effectiveness requirements.

3. The linear ordering optimization problem

A discrete multi-dimensional observation space will be considered, defined as a Cartesian product of a finite number of discrete sub-spaces:

$$\boldsymbol{D} = D_1 \times D_2 \times \ldots \times D_n. \tag{1}$$

Sub-spaces D_ν, $\nu = 1, 2, \ldots, n$, representing various features of the objects under observation, may have different formal natures and are assumed to be linearly ordered independently of each other. Taking into account computer implementation of pattern recognition systems, it is additionally assumed that each feature represented by

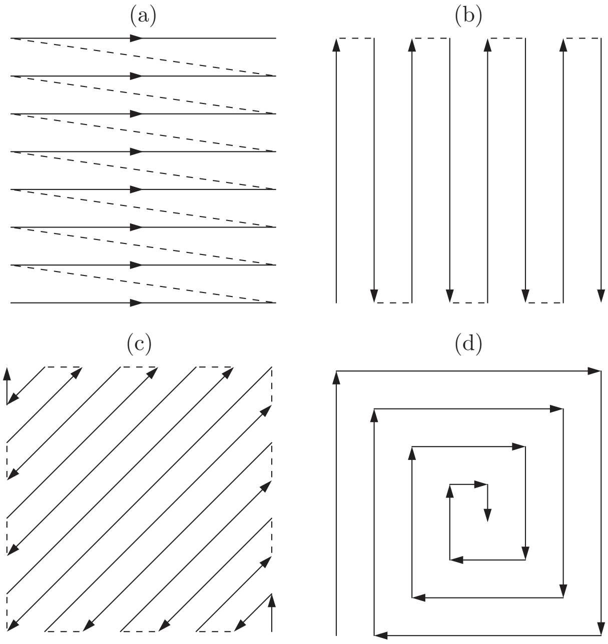

Figure 2. Selected types of discrete 2D observation space linear orderings: (a) lexicographical (with respect to columns and rows), (b) reversible sequential, (c) diagonal, (d) spiral

a discrete scale D_ν assumes values from a finite interval. With a simple transformation, this interval can be represented in a standard form of a sequence of integers $\Delta_\nu = [1, 2, \ldots, m_\nu]$. Therefore, a real observation space $\boldsymbol{D}$ can be reduced in practice to a discrete hyper-cube:

$$\boldsymbol{\Delta} = \Delta_1 \times \Delta_2 \times \ldots \times \Delta_n \tag{2}$$

The elements of each hyper-cube of this type can be ordered linearly in $M!$ ways, where $M = m_1 \cdot m_2 \cdot \ldots \cdot m_n$ is the number of elements of $\boldsymbol{\Delta}$. A lexicographical ordering of $\boldsymbol{\Delta}$ based on a fixed order of its components will be distinguished as one that can be easily realized technically. However, permutation of the components of $\boldsymbol{\Delta}$ results in changing the lexical order, excepting a situation when the order is merely reversed. Therefore, on the basis of the given family, $F = \{\Delta_\nu\}$, of n discrete sets, the number, $c = \frac{1}{2} \cdot n!$, of lexicographical orders can be established, significantly different from the pattern recognition point of view. For example, a family, $F = \{\Delta_1, \Delta_2, \Delta_3\}$, of three linearly ordered sets generates $3! = 6$ hyper-cubes: $\Delta_1 \times \Delta_2 \times \Delta_3$, $\Delta_1 \times \Delta_3 \times \Delta_2$, $\Delta_2 \times \Delta_1 \times \Delta_3$, $\Delta_2 \times \Delta_3 \times \Delta_1$, $\Delta_3 \times \Delta_1 \times \Delta_2$ and $\Delta_3 \times \Delta_2 \times \Delta_1$, only three of which (say, $\Delta_1 \times \Delta_2 \times \Delta_3$, $\Delta_1 \times \Delta_3 \times \Delta_2$ and $\Delta_2 \times \Delta_1 \times \Delta_3$) can be taken into account while generating substantially different lexicographical orders. We shall denote by Q_M a line representing the order of the elements of $\boldsymbol{\Delta}$, as illustrated in Figure 2; it can also be interpreted as a sequence (finite linearly ordered set) of M elements.

It is assumed that a pattern recognition task relies on assigning to any newly-observed element $\boldsymbol{x}$ in $\boldsymbol{\Delta}$ a pattern-index k, $k = 1, 2, \ldots, K$, K being a natural number > 1, so as to minimize the probability of false recognition in a long series of experiments. It is assumed in a supervised learning pattern recognition system that pattern-indices have been assigned to the elements of a reference sub-set $S \subset \boldsymbol{\Delta}$

according to verified results of earlier experiments. It is also assumed that each similarity class is represented in S by at least one element. Therefore, indices from an extended set $[0, 1, 2, \ldots, K]$ can be assigned to the elements of Q_M, 0 being assigned to elements not belonging to S. Taking into account the linear order in Q_M, the indices assigned to Q_M can also be presented by the following sequence:

$$V(Q_M) = [v_1, v_2, v_3, \ldots, v_M]. \tag{3}$$

At the same time, $V(Q_M)$ can be represented as a sequence of sub-series of homogenous elements:

$$V(Q_M) = [\sigma_1, \sigma_2, \ldots, \sigma_r], \tag{4}$$

where σ_ϱ, $\varrho = 1, 2, \ldots, r$, consists of elements of a fixed value. Any sub-series is delimited on the left and on the right by different-value elements or by the sequence's ends. The r integer (number of sub-series) plays a substantial role in the pattern recognition algorithm.

Theorem 1. If $M \geq K$ then:

(a) the minimum value of r is $r_{\min} = K$;

(b) the maximum value of r is

$$r_{\max} = N_1 \cdot (K+1) + (N_2 - N_1) \cdot K + \ldots + (N_K - N_{K-1}) \cdot 2 + 1, \tag{5}$$

where $N_1, N_2, \ldots, N_{K+1}$ are the numbers of elements of $V(Q_M)$ of a given value, taken in a non-decreasing order, $N_1 \leq N_2 \leq \ldots \leq N_{K+1}$.

Proof. Part (a) is evident, as the elements of $V(Q_M)$ can be linearly ordered so that the elements tagged by "0" are taken first (except when $S \equiv \boldsymbol{\Delta}$, *i.e.* when all elements of the observation space have been classified *a priori*), followed by those tagged by "1" and so on, up to the elements of the last reference subset.

In order to prove part (b) it will be initially assumed that a series of strong inequalities, $N_1 < N_2 < \ldots < N_{K+1}$, is satisfied. Then, at the 1^{st} step, we can form N_1 sub-sequences consisting of elements differently tagged by $K+1$ indices. Each sub-sequence of this type thus represents $K+1$ one-element sub-series. This justifies the first term of the right side of Equation (5). After this operation, there remain only elements tagged by K indices and the less numerous, uniformly tagged subset contains only $N_2 - N_1$ elements. The earlier reasoning can thus be re-applied to the reduced sequence of elements, leading to the second term of Equation (5). Such operations can be repeated until only elements tagged by 2 different indices remain; their number $(N_K - N_{K-1}) + (N_{K+1} - N_K)$ enables us able to form $(N_K - N_{K-1})$ two-element sub-sequences of differently tagged elements. At last, there remain $(N_{K+1} - N_K)$ uniformly tagged elements which allow us to form only one series represented by the last term of Equation (5).

In order to complete the proof, let us assume that some weak inequalities occur among the series of inequalities; let it for example be $N_1 = N_2$. Then, $N_3 - N_2$ sub-sequences consisting of $K-1$ differently tagged elements can be formed in the next step after forming N_1 sub-sequences consisting of elements differently tagged by $K+1$ indices. This means that the second term on the right side of Equation (5) disappears and, as a consequence, the $r_{\max}$ number is reduced. ■

For effective pattern recognition, the observation space $\boldsymbol{\Delta}$ should be linearly ordered so as to minimize the number, r, of series constituting the $V(Q_M)$ sequence. This can be achieved by a series of transformations of the initial lexicographical order. Each transformation is in fact a permutation of the elements of $V(Q_M)$. Let us denote by $\Pi_{\boldsymbol{\Delta}}$ the set of all permutations that can be applied to the elements of $\boldsymbol{\Delta}$, $\pi_0 \in \Pi_{\boldsymbol{\Delta}}$ being an initial permutation defined by lexicographical ordering of the elements of $\boldsymbol{\Delta}$. It is well known that each multi-element permutation can be realized by a sequence of simple permutations of pairs of the sequence's elements. Thus, a cost, C_s, defined as a minimum number of pair-wise permutations transforming π_0 into π_s may be assigned to any permutation $\pi_s \in \Pi_{\boldsymbol{\Delta}}$, $s = 1, 2, \ldots, M! - 1$, other then π_0. Then, the following problem can be formulated:

For given observation space $\boldsymbol{\Delta}$, number of classes K, reference subset $S \subset \boldsymbol{\Delta}$ and initial linear order in $\boldsymbol{\Delta}$ described by a permutation π_0, find a permutation $\pi_s \in \Pi_{\boldsymbol{\Delta}}$ reducing to K the number, r, of sub-series corresponding to $V(Q_M)$ and minimizing the permutation cost, C_s.

However, the solution of the above-formulated problem is, in general, an NP-difficult numerical task that could be extremely high time-consuming in practice. This is why we are interested in looking for sub-optimal solutions of the problem.

4. Linear order improvement

For linear order improvement, it is practical to consider aggregated transformations consisting of permutations performed on larger segments instead of single elements of a sequence Q_M. The aim of transformations remains minimization of the number of series in Q_M. For this purpose we shall use a notion t for a transformation (a single one or a composition of permutations) of the elements of Q_M (as well as of $V(Q_M)$). The set of transformations of this type will be denoted by T. A sequence $t_a\, t_b \ldots t_f\{\ \}$ will denote the result of a consecutive application to Q_M of transformations $t_f, \ldots, t_b$ and t_a.

The following types of transformations will be considered:

1. *reversion*: $t^r(p,q)$, $1 \le p < q \le M$, which consists in taking from Q_M its segment (a compact sub-sequence) starting from the p^{th} element and ending with the q^{th} element, reversing the order of its elements and inserting them into the same place in Q_M;
2. *shifting*: $t^s(p,q,y)$, $1 \le p,\ q,\ z \le M$, $p < q$, $p \ne y$, which consists in taking from Q_M its segment starting from the p^{th} element and ending with the q^{th} element and shifting it within Q_M to the position starting from y;
3. *segmental permutation*: $t^p(p,q,y,z)$, $1 \le p,\ q,\ y,\ z \le M$, $p < q$, $y < z$, $p \ne y$, which consists in taking from Q_M its segment starting from the p^{th} element and ending with the q^{th} element and a segment starting from the r^{th} element and ending with the s^{th} element and exchanging their positions in Q_M.

It will be shown that the above-described transformations can be used to improve the initial linear order of $\boldsymbol{\Delta}$.

THEOREM 2. Let r be the number of series in $V(Q_M)$. Then the number, r', of series in $t^r(p,q)\,\{V(Q_M)\}$ satisfies the following inequality:

$$r-2 \le r' \le r+2. \tag{6}$$

PROOF. Let us consider a part of the $V(Q_M)$ sequence containing the $[p,q]$ interval. We shall denote by a, b, c, d the values of this sequence delimiting the interval:

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & b & \ldots & c & d & \ldots \end{array}$$

where $a, b, c, d \in [0, 1, 2, \ldots, K]$. The following situations will be taken into account:

(a) $b=c$ meaning

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & b & \ldots & b & d & \ldots \end{array}$$

Then a reversion of the $[p,q]$ segment, independently on the a and d values, does not change the number of series, $r'=r$;

(b) $a=d$, which, for similar reasons, leads to $r'=r$;

(c) a, b, c, d are all different – then reversion of the $[p,q]$ segment also does not change the number of series, *i.e.* $r'=r$;

(d) $b\ne c$, $a\ne d$, $a=b$, $c\ne d$ (or $a\ne b$, $c=d$), meaning that the situation

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & a & \ldots & c & d & \ldots \end{array}$$

will be changed to

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & c & \ldots & a & d & \ldots \end{array}$$

and the number of series will thus be increased, $r'=r+1$;

(e) $b\ne c$, $a\ne d$, $a=c$, $b\ne d$, (or $a\ne c$, $b=d$), being an exact reversion of (d), thus the number of series will decrease, $r'=r-1$;

(f) $b\ne c$, $a\ne d$, $a=b$, $c=d$, *i.e.*:

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & a & \ldots & d & d & \ldots \end{array}$$

in which case reversion of the $[p,q]$ segment destroys two series and, as a consequence, the number of series increases, $r'=r+2$;

(g) $b\ne c$, $a\ne d$, $a=c$, $b=d$, *i.e.*:

$$\begin{array}{lcccccccc} \text{position:} & \ldots & p-1 & p & \ldots & q & q+1 & \ldots \\ \text{value:} & \ldots & a & d & \ldots & a & d & \ldots \end{array}$$

an exact reversion of (f), so the number of series will decrease, $r'=r-2$;

When $p=1$ or $q=M$ the impact of transformation $t^r(p,q)$ on the number of series is no greater than in the above-analyzed cases. The above-described situations thus complete the proof. ■

COROLLARY 1. The situation described in (g) of the Proof of Theorem 1 suggests the most effective reversion as a transformation improving the linear order in $\boldsymbol{\Delta}$.

THEOREM 3. Let r be the number of series in $V(Q_M)$. Then the number, r', of series in $t^s(p,q,y)\{V(Q_M)\}$ satisfies the following inequality:

$$r-3 \leq r' \leq r+3 \tag{7}$$

PROOF. Like in the Proof of Theorem 2, a part of the $V(Q_M)$ sequence containing the $[p,q]$ segment and the y element and their close environments will be analyzed:

position:	$\ldots$	$p-1$	p	$\ldots$ q	$q+1$	$\ldots$	$y-1$	y
value:	$\ldots$	a	b	$\ldots$ c	d	$\ldots$	e	f

where a, b, c, d, e, $f \in [0, 1, 2, \ldots, K]$. Shifting the $[p,q]$ segment leads to the following situation:

position:	$\ldots$	$p-1$	p	$\ldots$ q	$q+1$	$\ldots$	$y-1$	y
value:	$\ldots$	a	d	$\ldots$ e	b	$\ldots$	c	f

An analysis similar to that given in the Proof of Theorem 2 will lead us to the conclusion that in the following situations:

(a) all values a, b, c, d, e, f are different,

(b) all values a, b, c, d, e, f are equal,

(c) $a=e$ and $d=f$, the number of series after the shifting transformation remains unchanged, $r'=r$, as shifting the $[p,q]$ segment does not modify its close environment.

In other cases, the relationships within pairs of values (a,b), (c,d), (e,f), (a,d), (e,b) and (c,f), having a direct influence on the structure of series before and after the transformation, are critical for changing their number. From this point of view the following extreme situations can be distinguished:

(d) if $a=b$, $c=d$, $e=f$, $a \neq d$, $e \neq b$ and $c \neq f$, three series are split by insertion of segments of elements of different value and we obtain $r'=r+3$ as a consequence;

(e) if $a \neq b$, $c \neq d$, $e \neq f$, $a=d$, $e=b$ and $c=f$, three pairs of series are merged and we obtain $r'=r-3$.

When $p=1$ or $y=M$, the impact of the $t^s(p,q,y)$ transformation on the number of series is no greater than in the above-analyzed cases. Thus, the above-described situations complete the proof. ■

COROLLARY 2. The stuation described in (e) of the Proof of Theorem 2 suggests the most effective shifting as a transformation improving the linear order in $\boldsymbol{\Delta}$.

THEOREM 4. Let r be the number of series in $V(Q_M)$. Then, the number, r', of series in $t^p(p,q,y,z)\{V(Q_M)\}$ satisfies the following inequality:

$$r-4 \leq r' \leq r+4 \tag{8}$$

PROOF. A part of the $V(Q_M)$ sequence containing the $[p,q]$ and $[y,z]$ segments and their close environments will be analyzed:

position:	$\ldots$	$p-1$	p	$\ldots$ q	$q+1$	$\ldots$	$y-1$	y	$\ldots$ z	$z+1$	$\ldots$
value:	$\ldots$	a	b	$\ldots$ c	d	$\ldots$	e	f	$\ldots$ g	h	$\ldots$

where a, b, c, d, e, f, g, $h \in [0, 1, 2, \ldots, K]$. Mutual permutation of segments $[p,q]$ and $[y,z]$ leads to the following situation:

position:	$\ldots$	$p-1$	p	$\ldots$	$p+z-y$	$p+z-y+1$ $\ldots$
	$\ldots$	$p+z-q+1$	$p+z-q+2$	$\ldots$	z	$z+1$ $\ldots$
value:	$\ldots$	a	f	$\ldots$	g	d $\ldots$
	$\ldots$	e	b	$\ldots$	c	h $\ldots$

In the following situations:

(a) all values a, b, c, d, e, f, g, h are different,
(b) all values a, b, c, d, e, f, g, h are equal,
(c) $a=e$ and $d=h$,

the number of series after transformation remains unchanged, $r'=r$, as mutual permutation of segments $[p,q]$ and $[y,z]$ does not modify their close environments.

In other cases, the relationships within pairs of values (a,b), (c,d), (e,f), (g,h), (a,f), (g,d), (e,b) and (c,h), having a direct influence on the structure of series before and after the transformation, are critical for changing their number. From this point of view, the following extreme situations can be distinguished:

(d) if $a=b$, $c=d$, $e=f$, $g=h$, $a\neq f$, $g\neq d$, $e\neq b$ and $c\neq h$, four series are split by insertion of segments of elements of different value and we obtain $r'=r+4$ as a consequence;
(e) if $a\neq b$, $c\neq d$, $e\neq f$, $g\neq h$, $a=f$, $g=d$, $e=b$ and $c=h$, four pairs of series are merged and we obtain $r'=r-4$.

When $p=1$ or $z=M$, the impact of the $t^p(p,q,y,z)$ transformation on the number of series is no greater than in the above-analyzed cases. The above-described situations thus complete the proof. ■

Corollary 3. The situation described in (e) of the Proof of Theorem 3 suggests the most effective segmental permutation as a transformation improving the linear order in $\boldsymbol{\Delta}$.

Corollaries 1–3 contain recommendations on choosing sequences of transformations of an initial linear order in an observation space. However, the criterion of transformations' effectiveness has only been applied on the number of series, r. Therefore, they have not taken into account the cost of calculations connected with using serial statistical tests based on the given linear order. This cost increases with the distance between the initial lexicographical order and the linear order obtained as a result of a series of transformations. The distance between two linear orders of sequences consisting of the same elements is understood here as the minimum number of pair-wise permutations transforming one order into the other. At the same time, no recommendations follow from Theorems 1–3 about choosing the position and lengths of reversed, shifted or permuted segments. However, it seems reasonable to operate, first of all, on the segments delimited by the initial lexicographical order.

Ther remains a problem of selecting the segments of Q_M for order changing. The method of assessing segments' ordering based on numerical parameters described in [5] can be used for this purpose.

5. Conclusions

Linear ordering of an observation space is a new paradigm of pattern recognition based on non-parametric, serial statistical tests. The effectiveness of of this type of tests depends on choosing a linear order appropriate for the geometrical form of similarity classes, which are unknown *a priori*. Therefore, there arises a problem of linear order optimization according to the available learning data-subsets. It has been shown in this paper how to improve the initial linear order (usually lexicographical) by a sequence of transformations: reversing the order at selected intervals and shifting the intervals or permutation of selected pairs of intervals on the sequence of linearly ordered data. Although our considerations are theoretical in character, they indicate ways to construct the corresponding algorithms applicable in pattern recognition systems based on serial statistical tests.

References

[1] Abe S 2005 *Support Vector Machines for Pattern Classification*, Springer, London

[2] Duda R and Hart P 1973 *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York – London – Sydney

[3] Tou J and Gonzales R 1974 *Pattern Recognition Principles*, Addison-Wesley Publishing Co, London – Amsterdam – Sydney

[4] Kulikowski J L and Wierzbicka D 2004 *Biocybernetics and Biomedical Eng.* **24** (2) 27

[5] Kulikowski J L and Wierzbicka D 2005 *Biocybernetics and Biomedical Eng.* **25** (3) 65

[6] Kulikowski J L 2006 *Proc. Int. Multiconf. on Computer Science and Information Technology*, Wisla, Poland, PIPS, pp. 103–109 on CD

[7] Kulikowski J L, Przytulska M and Wierzbicka D 2005 *Proc. 4^{th} Int. Conf. on Computer Recognition Systems CORES'05* (Kurzynski M *et al.* Eds), Springer-Verlag, Berlin, pp. 235–242