

PAPER • OPEN ACCESS

Sub-band invariants of handwritten texts fuzzy fragments

To cite this article: E G Zhilyakov *et al* 2021 *J. Phys.: Conf. Ser.* **1801** 012024

View the [article online](#) for updates and enhancements.



240th ECS Meeting ORLANDO, FL

Orange County Convention Center **Oct 10-14, 2021**

Abstract submission deadline extended: April 23rd

SUBMIT NOW

Sub-band invariants of handwritten texts fuzzy fragments

E G Zhilyakov¹, A N Zalivin², S I Matorin² and A A Chernomorets¹

¹ Belgorod State National Research University, 85, Pobedy, st., Belgorod, 308015, Russia

² Belgorod University of Cooperation, Economics and Law, 116a, Sadovaya, st., Belgorod, 308023, Russia

E-mail: zhilyakov@bsu.edu.ru

Abstract. Socio-economic processes informatization is largely implemented by using various documents electronic storage, including scanned handwritten texts or their fragments images, in particular, in the form of officials' original signatures. Among the very diverse tasks of handwritten documents scans electronic repositories computer analysis, the direction of automatic search for images fragments, that contain outlines of key words of interest, is quite relevant. The search for such fragments is significantly hampered by the variability of the letterforms, even by the same author. Therefore, it is advisable to use a fragment from the analyzed text as an initial sample. Thus, we are talking about the fragments precedent identification. In this case, a complicating factor is the impossibility of using many identical objects in training, and only one is available. Therefore, the problem of forming an artificial sample, which elements preserve some common characteristic property, arises. In this work, the authors substantiated the application adequacy of the original sub-band approach, which is based on the concept of the data segment Euclidean norm sub-band part and the mathematical apparatus of sub-band matrices obtained on this basis. Decision procedures for the handwritten text fragments sub-band identification were developed, including training on a single precedent.

1. Introduction

Nowadays, large volumes of document scan images have been accumulated, they are either entirely handwritten or contain handwritten fragments, for example, signatures. A sufficient number of reasons can be given to search in such repositories for images fragments that contain specific keyword forms and are of interest. In particular, such problems naturally arise in the process of analyzing authors' literary works manuscripts, in information and analytical security systems, in the search for signatures that were falsified, etc. Obviously, in these cases we mean searching for fragments of handwritten identical word forms created by the same author's hand. In accordance with the current terminology, we will term the collection of such fragments as a class, and their search as the identification in a given class.

The methods of object recognition in images are considered in many literary sources [1]. The main purpose of recognizing handwritten texts scans is, as the rule, to change texts into the form available for printing. Therefore, most often we are talking about the recognition of individual letters and symbols [2]. Obviously, the search for fragments, according to the keywords available in the analyzed



text, has a certain specificity, on account of the meaningful goal [2-4]. An essential feature of such fragments is the absence, even for one and the same author, of the symbols outline exact repetition within the entire text or from document to document (for example, when verifying a signature) [5,6]. Therefore, it seems natural to term them fuzzy duplicates. At the same time, for their automatic identification, it is necessary to determine some invariants that reflect the peculiarities of the authors' handwriting and, to a sufficient extent for comparisons, are preserved from fragment to fragment. It is the presence of such invariants that makes it possible to construct the corresponding decision function, the values of which are little changeable in the class of identical fragments. It is also important to require significant differences in the values of the decision functions when comparing with fragments not from the selected class.

Within the framework of this work, we consider the problem of constructing such invariants and decision functions based on them.

It is assumed that the scanned text image as a whole and its fragments are stored in digital form, that is, in the form of some rectangular table (matrix) of numbers. Therefore, without loss of generality, for some image fragment, the following representation can be used:

$$F = \{f_{ik}\}, i = 1, \dots, I; k = 1, \dots, K, \quad (1)$$

meaning that the first index denotes the lines and the second one denotes the columns of the matrix.

In accordance with the established terminology, the elements of this matrix are termed pixels. Within the framework of this work, it is assumed that the spatial discretization along both axes of the image coordinates is carried out with unit steps (equal distances between the real coordinates of pixels).

Let it now be determined that the handwritten fragment (1) was formed when drawing a word form, it is of interest within the framework of some text analysis meaningful task. The task is to search for identical fragments in the rest of the text, in the sense that they are inscribed when writing the same word form.

Thus, within the pattern recognition methodology framework the image analysis is reduced to comparing its fragments in order to test the following initial hypothesis validity.

H_0 - the compared fragments are identical.

Implementation of a decision function in favor of one or another alternative when matching with the fragment is:

$$Y = \{y_{ik}\}, i = 1, \dots, I; k = 1, \dots, K, \quad (2)$$

involves the use of some proximity measure:

$$\rho(F, Y) = \Psi(F, Y), \quad (3)$$

where Ψ is the functional that determines computational processes.

The initial hypothesis is rejected when the condition is fulfilled:

$$\rho(F, Y) \in G_\alpha, \quad (4)$$

where G_α is the critical area, the proximity measure values on identical fragments probability falling into (wave from above) do not exceed the desired one:

$$Ver(\rho(F, \tilde{F}) \in G_\alpha) \leq \alpha. \quad (5)$$

Following [7], the proximity measure (3) in combination with the critical area will be termed the decisive function.

In real life, the critical area is most often determined on the basis of training by using a sample of identical fragments. It is clear that in the considered problem formulation, such a sample is not available, and therefore it is necessary to develop the method for its artificial generation (augmentation), based on the invariance of some characteristics. It is easy to understand that precisely the use of invariants makes it possible to reduce first type errors probabilities when testing the hypotheses (5).

2. Materials and methods

2.1 Decision making based on sub-band analysis in the spatial frequency domain

Reasonably assuming that the number of lines in the fragments being compared is less than the number of columns, it is proposed to use a vector measure of their proximity:

$$\vec{\rho}(F, Y) = (\rho_1, \dots, \rho_N)', \quad (6)$$

its components are defined on the corresponding pairs of lines:

$$\rho_i = \rho(\vec{f}_i, \vec{y}_i), i = 1, \dots, N, \quad (7)$$

$$\vec{f}_i = (f_{i1}, \dots, f_{iM})', \quad (8)$$

$$\vec{y}_i = (y_{i1}, \dots, y_{iM})'. \quad (9)$$

Then the critical area is also a vector:

$$\vec{G}_\alpha = (G_1^\alpha, \dots, G_N^\alpha)', \quad (10)$$

its components provide admissible probabilities α_i for the values of the proximity measure corresponding components (6) to fall into them if the hypothesis about the identity of the compared fragments is true.

The final decision on the inconsistency of this hypothesis with real data is taken when an event occurs:

$$W = \bigcap (\rho_i \in G_i^\alpha), i \in D_E = \{i_1 < i_2 < \dots < i_E\}, \quad (11)$$

which consists in the simultaneous occurrence of values of no less than E proximity measure component in the corresponding critical areas.

Assuming that the proximity measure components are independent, for the probability of the first type errors we can obtain the inequality:

$$\alpha \leq (\alpha_i)^E, \quad (12)$$

where α_i is the probability of events, which intersection determines the right-hand side of the equation (11):

$$\alpha_i = \text{Ver}(\rho(\vec{f}_i, \vec{y}_i) \in G_i^\alpha). \quad (13)$$

Therefore, when determining the critical area components (10), the following equation should be used:

$$\alpha_i = (\alpha)^{1/E}, i = 1, \dots, N. \quad (14)$$

As the basis for calculating the components (7) of the vector proximity measure, it is proposed to apply sub-band analysis in the spatial frequency domain, meaning the use of sub-band parts of the Euclidean (energy) square [8]:

$$\|\vec{f}_i\|^2 = \sum_{k=1}^M f_{ik}^2 = \sum_{r=1}^R P_r(\vec{f}_i), \quad (15)$$

where:

$$P_r(\vec{f}_i) = \vec{f}_i' A_r \vec{f}_i; \quad (16)$$

$$A_r = \{a_{ik}^r\}, i, k = 1, \dots, M; a_{ik}^r = 2a_{ik}^0 \cos(w_r(i-k));$$

$$a_{ik}^0 = \sin(\Delta(i-k)) / \pi(i-k); w_r = 2\Delta(r-1), r = 1, \dots, R; \Delta = \pi / (2R-1).$$

Matrices in quadratic forms (16) are naturally termed sub-band matrices.

As a sub-band invariant of the original fragment (1), it is proposed to use the sub-band matrices set of sums (S_i is the set of indices):

$$A_{S_i} = \sum_{r \in S_i} A_r, i = 1, \dots, M, \quad (17)$$

their terms of sum satisfy the inequalities:

$$P_r(\vec{f}_i) \geq 2\Delta \|\vec{f}_i\|^2 / \pi, r \geq 2 : P_1(\vec{f}_i) \geq \Delta \|\vec{f}_i\|^2 / \pi. \quad (18)$$

Obviously, the total matrices are symmetric and positive definite, and therefore can be represented in the form [9]:

$$A_{S_i} = Q_i L_i Q_i' \quad (19)$$

where Q_i and L_i are eigenvectors matrices and numbers of sub-band matrices $A_i Q_i = Q_i L_i$,

$$L_i = \text{diag}(\lambda_1^i, \dots, \lambda_M^i); \lambda_k^i \geq \lambda_{k+1}^i \quad Q_i' Q_i = Q_i Q_i' = \text{diag}(1, \dots, 1) \quad (20)$$

As noted in [8], for the eigenvalues of the total sub-band matrices, we can quite accurately set:

$$\lambda_{J_i+k}^i \approx 0, J_i = 2[M | S_i | \Delta / \pi] + 6, \quad (21)$$

where square brackets mean the integer part of the number, and $|S_i|$ is the number of terms in the sums (15).

Therefore [8] of the component:

$$\vec{z} = Q1_i Q1_i' \vec{x} \quad (22)$$

any vector $\vec{x} = (x_1, \dots, x_M)'$ has the property:

$$Z(\omega) = X(\omega), \omega \in \Omega_i, \quad (23)$$

where Ω_i is the combined frequency interval, consisting of sub-bands defined by inequalities (18); $Q1_i$ is the dimension matrix, consisting of the corresponding eigenvectors matrices first columns.

In relation (23) we mean the Fourier transforms of the view:

$$X(\omega) = \sum_{k=1}^M x_k \exp(-j\omega(k-1)). \quad (24)$$

In other words, in the combined frequency interval, the segments of the original vector Fourier transforms and its projection onto the subspace of eigenvectors corresponding to nonzero eigenvalues coincide.

As components of the identity measure (10), it is proposed to use the sub-band part of the compared fragments row vectors pairs differences energy:

$$G_i^\alpha = P_{S_i}(\vec{g}_i) = \vec{g}_i' A_{S_i} \vec{g}_i, \vec{g}_i = \vec{f}_i - \vec{y}_i. \quad (25)$$

It is easy to show the relation validity:

$$G_i^\alpha = P_{S_i}(\vec{t}_i), \vec{t}_i = \vec{y}_i - Q1_i Q1_i' \vec{f}_i, \quad (26)$$

that is, the value of the proximity measure component is determined by the sub-band part of the difference energy between the vector being compared and the original vector projection of the view (22) onto the subspace determined by the eigenvectors of the total sub-band matrix corresponding to nonzero eigenvalues.

It is clear that this projection also serves as an invariant when comparing fragment lines.

2.2. Identification of critical areas based on training

The presence of only one fragment of a given class leads to the need to construct an artificial sample for training in order to determine the decision function components critical areas (25). When it is necessary to use relations of the view (15), where α and E are assumed to be pre-determined.

The model for generating artificial samples is based on the use of their representation through invariants:

$$\tilde{f}_i^m = Q1_i Q1_i' \vec{f}_i + c_m \vec{u}_m, m = 1, \dots, D, \quad (27)$$

where, in accordance with (13), the number of samples is determined by the inequality:

$$D \geq 1/(\alpha)^{1/E} \quad (28)$$

the vector components $\vec{u}_m = (u_{1m}, \dots, u_{Mm})'$ are Gaussian pseudo-random numbers with zero mean and unit variance.

Positive coefficients $c_m > 0$ are determined by the condition that all components of the generated vector are positive:

$$\tilde{f}_{ik}^m = 0, \forall k = 1, \dots, M \quad (29)$$

The purpose of training is to determine a positive threshold, its frequency of exceeding does not exceed the right side (13).

3. Conclusion

The urgency of the problem of developing a procedure for automatic search for identical fragments in the handwritten texts scans images and the vector decision function based on the proposed sub-band invariants of the fragments lines in the area of spatial frequencies are substantiated. The scheme for generating an artificial training sample was developed, it allows to construct the critical area of the vector decision function.

4. Acknowledgments

The research was carried out with the support of the RFBR grant No. 20-07-00241a

References

- [1] Gonzalez R and Woods R 2006 *Digital image processing* (Moscow: Technosphere) 1072
- [2] Gorskiy N 1997 *Handwriting recognition: from theory to practice* (SPb.: Polytechnic) 126
- [3] Merkov A B 2004 *The main methods used for handwriting recognition* (Moscow: Pattern Recognition Laboratory MCNMO) 31
- [4] Sorokin A I and Zapryagaev S A 2010 Segmentation of handwritten and typewritten texts by the method of Voronoi diagrams *Voronezh State University 2010 VSU bulletin, series: system analysis and information technologies* **1** 160-165
- [5] Zhilyakov E G, Efimov N O, Chernomorets A A and Trubitsyna D I 2019 About subband analysis of handwriting images *Electronic information systems* **20(1)** 73-86
- [6] Efimov N O 2018 The procedure for segmentation of handwritten text images into words in the tasks of precedent identification *Electronic information systems* **18(3)** 69-80
- [7] Abraham Wald and Sons 1950 *Statistical Decision Functions* *New York: John Wiley and Sons, Inc.; London: Chapman and Hall, Ltd* 179
- [8] Zhilyakov E G 2017 Construction of trends in time series segments *Autom. Remote Control* **78(3)** 450-462
- [9] Horn R and Johnson H 1989 *Matrix Analysis*. (Moscow: Mir) 665