

RESEARCH

Open Access



# A novel methodology for writer (hand) identification: establishing Rigas Feraios wrote two important Greek documents discovered in Romania

Athanasios Rafail Mamatsis<sup>1</sup>, Eirini Mamatsi<sup>1</sup>, Constantinos Chalatsis<sup>1</sup>, Dimitris Arabadjis<sup>2</sup>, Pandora Kampouri<sup>1</sup> and Constantin Papaodysseus<sup>1\*</sup>

## Abstract

The main goal of the present work is to determine the hand that has written two newly discovered documents in Romania. For giving the proper answer, the authors introduced the notion of “Ideal Representative”, namely of an object that very well represents the corresponding ideal alphabet symbol that a writer had in his/her mind when writing a document by hand. Moreover, the authors have introduced a novel method, which leads to the optimal evaluation of the Ideal Representative of any alphabet symbol in association with any handwritten document. Furthermore, the authors have introduced methods for comparing these Ideal Representatives, so as a final decision about the hand that has written a document may be obtained with a highly considerable likelihood. The related analysis manifests that the two documents discovered in Romania in 1998, belong to the great personality of Rigas Feraios. The presented method of automatic handwriting Identification seems to be of general applicability.

**Keywords** Writer identification, Rigas Feraios, The Saganaki of Madness, The Tested Friendship, Ideal representative, Objects representing handwriting

## Introduction

The aim of the research presented here is to give an as objective as possible and quantitative answer to the important question, which one of the following five documents had been written by the very same hand of Rigas Feraios or not; the latter was one of the grater personalities who drastically influenced the history of Greece, but of all Balkan Nations, too, in the last three centuries (see

1.2 below). In essence, the present work deals with automatic writer identification, in association with the following five (5) handwritten texts:

1. “Compilation of problems in Physics” (“Φυσικής Απάνθισμα”) hence forth, called FYSAP, for brevity. We would like to strongly emphasize that FYSAP is the only document, which is historically certain and unambiguous that it has been written by the hand of Rigas Feraios.
2. “The Greek Constitution” and “the Thourios” (“Σύνταγμα και Θούριος”) appearing in the same document; in all the subsequent work, we shall use the common abbreviated name SYNTAG for both these texts.
3. A third document, unambiguously written by the hand of another great Greek Politician Eleftherios

\*Correspondence:

Constantin Papaodysseus  
cpapaod@cs.ntua.gr

<sup>1</sup> School of Electrical & Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou, 15780 Athens, Greece

<sup>2</sup> School of Engineering, University of West Attica, Petrou Ralli & Thivon 250 Egaleo, 12241 Athens, Greece



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Venizelos. This text is kept in National Research Foundation “Eleftherios Venizelos” [1]; we shall refer to it as ELVEN.

The other two documents which practically triggered the work in hand, are more extensively referred to in “Unpublished Documents Probably Associated with Rigas” subsections and “Scientific Dispute Concerning the Actual Writer of “The Saganaki of Madness” and “The Tested Friendship”” that follow immediately. In subsection “Scientific Dispute Concerning the Actual Writer of “The Saganaki of Madness” and “The Tested Friendship”” we shall integrate the goal of the present work.

#### Unpublished documents probably associated with Rigas

In 1998, exactly two hundred (200) years after the cowardly assassination of Rigas Feraios, the Romanian scientist Lia Brad Chisacof [2] presented a handwritten document, which, according to her opinion had been written by Rigas. This document includes two unknown and obviously unpublished literary works:

- a. The comedy “The Saganaki of Madness” (“Το σαγανάκι της τρέλας”), hence forth symbolized SAGAN; we note that “the saganaki” is a specialty of the Greek kitchen based on fried cheese. The content of SAGAN is a satire, mocking the governor of Wallachia Nikolaos Mavrogenes, whose secretary was Rigas. Mavrogenes was an aliterate, brutal, vulgar, gauche, uncultivated and psychologically unstable individual. The author of this play, i.e., of SAGAN, scoffs all these defects of Mavrogenes, sometimes following the style of the grate Moliere; however, the author bases his humor and jokes on multilingual expressions, thus achieving to go far beyond the style of Moliere.
- b. The short novel “The Tested Friendship” (“Δοκιμασμένη φιλία”), to which we will refer with the name FILIA in the following. Both this works have been published in Bucharest on behalf of the “Institute of Southeast European Studies” introduced and translated in Romanian by Lia Brad Chisacof [2].

#### Scientific dispute concerning the actual writer of “the saganaki of madness” and “the tested friendship”

There is a serious scientific dispute, concerning the hand that has written the aforementioned newly discovered documents, we have called SAGAN and FILIA ([2]). Thus, the automatic identification of the hand that has written these two documents is, substantially, the main

goal of the present work. Equivalently, we will test the hypothesis if documents SYNTAG, SAGAN and FILIA were written by the hand of Rigas Feraios or not; in attempting that, we shall employ the fact that FYSAP and ELVEN have been unambiguously written by the hand of Rigas Feraios and Eleftherios Venizelos respectively. This testing will be accomplished via an improved novel extension and proper modification of a methodology introduced by the authors in connection with the writer identification of ancient Greek (Hellenic) inscriptions and byzantine codices [3, 4].

#### A brief description of the life and the accomplishments of Rigas Feraios

Rigas Feraios-Velestinlis (Ρήγας Φεραΐος-Βελεστινλής), according to most scholars and not only, was the greater personality, intellectual, author, revolutionary in the Balkans during the late 18th and early nineteenth century. He, actually, introduced the Enlightenment in the Balkans, especially among the Greek population. Rigas was born in the Greek village Velestino, the ancient Feres (“Φεραΐ”), in 1757. His village was then occupied by the Ottoman Empire; nowadays, Velestino-Feres belongs to the modern Greek region of Thessaly.

Next, Rigas was first educated at the schools of Zagora and Ampelakia, which were famous in this era. After his graduation, he became a teacher in the village “Kissos”, where he started expressing his opposition to the Ottoman Empire, actively. At the age of twenty he killed a prominent Turk, because he had treated him as a Greek slave. Due to this action, Rigas was forced to flee in “Litoro village” on the celebrated Mount Olympus, where he joined the group of rebels led by Spiros Zeras. From there Rigas went to mount Athos and more specifically in the monastery of Vatopedion. This monastery had, and it still has, a very rich library; Rigas ensured a free, unlimited access to this library and, consequently, he drastically improved his knowledge and expertise in various disciplines. In 1785, Prince Alexandros Ypsilantis (the Ambassador of Rossa in the Ottoman Empire) invited Rigas to Constantinople (Istanbul), for further studies. Next, in 1788 Feraios went to Wallachia, where he became the secretary of the Ruler Nicholas Mavrogenes. There, he was informed about the French Revolution, and very soon, he conceived the idea that such a revolution could take place in the Balkans against the Ottoman Empire from the populous Cristian community. In the meanwhile, the Ottoman Empire has lost the war against Rossa in 1792 and considered Nicholas Mavrogenes to be responsible for the defeat of Turks in Wallachia and, as

a result, Mavrogenes were decapitated. Thus, Rigas was, once more, forced to flee in Vienna, the Austral capital and he started his very important revolutionary actions.

In Vienna, in the printing premises of the Greek brothers Pouliou, Rigas published the following books-works:

- i. “The School of Delicate Lovers” (“Σχολείο των ντελικάτων εραστών”).
- ii. “Compilation of problems in Physics”, “Φυσικής Απάνθισμα”, a collection of remarks and problems on Physics,
- iii. “Ἠθικός Τρίποδας”, consisting of three important documents fully supporting the Enlightenment; two of these texts were written by Marmodel and they had been translated by Rigas in modern Greek.
- iv. “Ανάχαρσις”, which was a translation in modern Greek of a famous book by Loukianos written in the second century B.C.; the book refers to the proper citizen education (Civics) and in particular of the young ones, as well as the role of gymnastics in Physics.
- v. The “Thourios” (“Θούριος”), a kind of march, aiming at giving courage at the enslaved Greek and Balkan Christians, in order to uprising against the Ottoman Empire. One of the more well-known, famous distiches of Thourios is the following:  
*“it’s better to live free for one hour,  
 Than forty years in prison, under slavery”.*
- vi. “The constitution of the Greek democracy”. In this book, Rigas makes the very advanced for his era proposal, that the Greeks should follow the steps of the French rebels and build a “Popular Democracy”, very similar to the French democracy.
- vii. “The Revolutionary Declaration” (“Ἡ Επαναστατική Προκήρυξη”). In this declaration, Rigas tries to persuade all Greeks and the Cristian Balkans to uprising against the Sultan’s tyranny; in this declaration, Rigas also tries to arouse the Turkish population, appealing to their democratic feeling.
- viii. “The Charta”, a map of the Balkan peninsula, which extends to the south of Danube River.

Unfortunately, the content of these publications annoyed the Ottoman Empire too much and, as a result, Turkish secret agents murdered Rigas Feraios with an abominable manner.

We should emphasize that, in connection with most of the aforementioned publications, there are handwritten documents, too. However, only for the “Compilation of problems in Physics” (“Φυσικής

Απάνθισμα”) there is sufficient historical evidence that it has been undeniably written by the hand of Rigas.

### State of the art in automatic handwriting identification

In the last years, given the great evolution of computers, identification of the hand that has written a document may be achieved via the use of computational algorithms. In fact:

Connected-component contours and edge-based features of uppercase Western script are considered for offline automatic Writer Identification by the authors of [5]. More specifically, the authors propose that, for automatic Identification of upper-case Western letters, the combination of the connected-component contour codebook and its probability-density function of shape usage, together with the edge-orientation distribution, offers really satisfactory results. In [6] a technique is proposed which divides a given handwriting into small fragments; then, each such fragment is considered to be a texture. In particular, in all these fragments-textures, three texture descriptors, namely Local Binary Patterns (LBP), Local Ternary Patterns (LTP) and Local Phase Quantization (LPQ) are considered, and each handwritten fragment is represented by the histograms of these descriptors. Comparison-classification of these descriptors is the basis of the writer identification. In [7], the authors propose quantitative methods for the classification of calligraphic style, on a statistical basis. The proposed classification and quantification methods rely on specialized local and global visual style features. [8] use texture-based schemes for Writer Identification. More specifically, the authors combine Delta encoding and oriented Basic Image Feature Columns to achieve writer identification. In [9], the Levenshtein edit distance based on Fisher-Wagner algorithm is used to calculate the cost of transforming one handwritten word into another. He et al. in [10] apply the “SOTM” method of [11] in order to generate a codebook which contains the temporal information of the handwritten patterns. In this way, one may render the historical document dating, a standard pattern recognition problem. The authors of [12] use two different features’ classes in order to identify the writer of handwritten excerpts from their binary images. In fact, the authors use two curvature-free features for writer identification based on the run-lengths of general patterns and the joint distribution of the relation between orientation and length of a set of line segments extracted from contours of ink traces. In [13], the authors try to determine the minimum number of hands that could have written a set of 16 inscriptions from the “Judahite

desert fortress of Arad”; to achieve that, they use statistical inference over the distribution of the distances between mixed features representations of letter shapes, e.g. angles between strokes and character profiles. In [14], structural and textural features, which express the hand-writing visual similarity, are extracted and used by experts to interactively cluster the documents with a manually defined feature subset. The authors of [15] use computer-vision algorithms and statistical inference methods in order to identify fragments that might originate from the same codex. The developed system consists of a subsystem that compares handwriting, as well as a subsystem that considers the appropriate cataloging data, as extracted from the images by a suitable method. Wolf et al. [16] look for eventual joins between catalogued excerpts, using a mixture of local descriptors and learning techniques. The authors developed a benchmark, the “Genizah” one, in which they have employed various vectorization methods and similarity measures, in order to reach a final statistical inference. In [17], a junction detection method is proposed, applicable in writer identification. In fact, the authors perform skeletonization of the letters, in order to determine the “forks”; at each fork, they evaluate the probability  $S(\theta)$  that direction  $\theta$  indicates one of the branches of the junction. Finally, they consider probability distribution of the junctions as a global feature for each writer based on a learned codebook. A new system for writer identification of medieval manuscripts is presented in [18]. The system is based on features from layout analysis, such as the way the scribe is distributed in each row of the text according to the ruling, intercolumnar distance, upper and lower margin, interlinear spacing, peak number, etc. In [19], Deep Learning (DL) algorithms and classical machine learning approaches are experimentally compared; the authors deduce that the DL-schemes outperform or are equivalent to the latter ones. The author of [20] proposes a new Residual Swin Transform Classifier for offline Word-Level Writer Identification. [21] investigates the effectiveness of textural measures in characterizing the writer of a handwritten document. More specifically, the authors introduce a representation of the local binary patterns (LBP column histogram) and they combine it with the oBIFs column histogram to enrich the representation. Next, they use these local patterns to characterize the writer, via the SVM classifier. Classification is carried out using the SVM classifier. [22] employs Global-context residual recurrent neural networks for writer identification. The authors combine the global-context information extracted by a convolutional neural network and local fine-grained information extracted by a recurrent neural network. The authors of [23] use Deep Learning Techniques for Writer Identification. Indeed, FAST key

points and Harris corner detector are employed to identify points of interest in the handwriting. Subsequently, small patches centered around these key points are fed into a deep convolutional neural network; hand-writer identification is based on maximum likelihood considerations. The article [24] investigates the impact on writer identification of four different techniques that entail Directional Hinge feature extraction methods.

In the system described in [25], features that capture the distribution of the strokes’ thickness are concatenated with deep features, computed by a succession of convolutional blocks; in this way, shallow softmax nets may approximate the classification of handwriting excerpts. The system has been trained with a 4/1 ratio of the training to test set and for a fixed number of writers. In [26], an analogous practice of concatenating deep global and “restricted” features is employed. The restricted features are computed locally in fragments of the input images, in order to capture graphemes, according to the method introduced in [27] and widely and effectively used thereafter ([28]). The system presented in [29], is also based on restricted deep features. There, the deep convolutional features are computed separately on words and connected components of the handwriting. Then, the optimized feature maps are concatenated and set as the input of a feed-forward network, which, in turn, is trained to be the documents’ classifier. In an analogous manner, the system of [30] restricts the deep features to be computed and optimized on the separate lines of text. In turn, the employed feed-forward classifier is not trained separately on the basis of the optimal feature maps. The authors of [31], restrict deep features on very elementary data of the handwriting, i.e. the individual letters. The developed architecture of the feature extraction network combines different deep learning mechanisms, in order to represent the characteristics of structured data, such as the letters’ contours. In [32], the idiosyncrasy evaluation that experts do on handwriting excerpts guides the training of a reinforcement learning model, designed to detect sequences of handwriting excerpts of maximal idiosyncrasy scores. In turn, these excerpts are the data upon which the convolutional features that feed the deep classifier are computed. The authors of [33] use path integral signatures and SIFT features to represent local properties of the handwritten shapes in a translation-free, scaling-free and reparameterization-free way. Then, codebook learning techniques are applied to these base features in order to extract global features vectors of as limited as possible dimension. In [34], a more diverse and writing style—specific set of graphometric features is employed to represent the handwriting word-wise. Next, principal component analysis (PCA) is used to affinely map the data in a lower dimensional space and a shallow network

is trained to perform the writer classification task. The authors of [35] employ a deep convolutional neural network, based on modified pre-trained networks, that is designed and developed to extract features from raw data hierarchically.

Though the aforementioned approaches vary a lot in the way that graphological and/or geometrical representations of the handwriting are fused within the deep learning prototype, there is a common intrinsic restriction due to the necessity of a training procedure. The classification procedure either is independent to the identities of the selected features space or it imposes restrictions on the induced deep features. In the first class belong methods like [33, 34], and, partially, [25] and [32], that do not implicate the target classification in the feature extraction. In the second class, belong methods analogous to [26, 29] and [30], where the deep features are trained together so as to optimize the classifier's accuracy.

### A summary and the novelties of the introduced methodology

Considering the aforementioned State of the Art, we believe that both the entire methodology presented here, as well as various aspects of this approach are substantially novel. In particular:

- a. We form proper bundles of optimally fit realizations of every alphabet symbol, separately (see [Optimal Matching of Any Two Realizations of the Same Alphabet Symbol](#) Section).
- b. We render all contours of each such bundle equinumerous, in the sense that we guarantee that they all consist of the same numbers of pixels, without practically influencing their morphology (see [The “Ideal Representative” of an Alphabet Symbol, Concerning a Document of a Specific Writer](#) Section).
- c. We evaluate the curvature at the center of each pixel of all contours, via proper polynomial approximation of the contour. Next, we define a biunivocal correspondence between pixels, with pretty close value of curvature (see [The “Ideal Representative” of an Alphabet Symbol, Concerning a Document of a Specific Writer](#) Section).
- d. Among all pixels of common curvature, we evaluate the mean value of the x and y coordinates, thus obtaining an average curve, which we call “First-Version Representative” (see [The “Ideal Representative” of an Alphabet Symbol, Concerning a Document of a Specific Writer](#) Section).
- e. By repeating the aforementioned procedure in connection with all these First-Version Representatives, we obtain a final “Ideal Representative”, which we firmly feel that best represents the form of the alphabet symbol the Writer had in his mind, when writing the specific document (see [The “Ideal Representative” of an Alphabet Symbol, Concerning a Document of a Specific Writer](#) Section).
- f. For each alphabet symbol separately, we compare the Ideal Representatives that correspond to two different document parts and then, we apply a novel statistical procedure for deciding if these two documents have been written by the same hand or not (see [Identification of Handwriting Based on the Ideal Representatives](#) Section and [A Statistical Approach for Identifying Rigas Feraios’ Handwriting](#) Section).

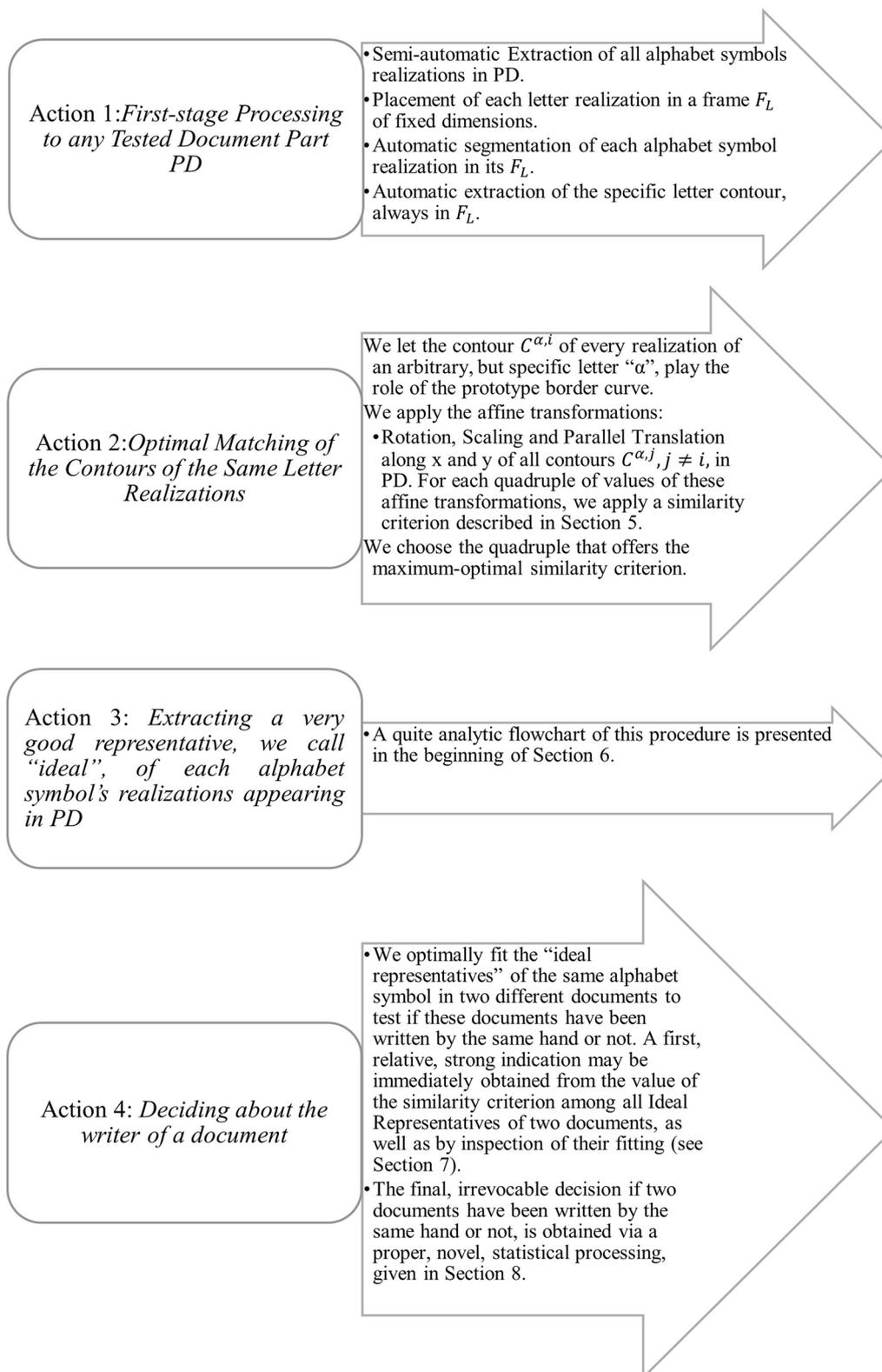
We would like to emphasize that the entire aforementioned procedure is applied to a properly selected set of eleven (11) greek alphabet symbols. The reason for and the method of this selection will be explicitly referred to in the end of present Section.

Conclusively, the aforementioned procedure allows for rigorously determining rules for the maximum likelihood classification of documents into writing hands. This formulation of the writer identification task is totally dataset-independent since the representation of the writing style via the corresponding Ideal Representatives and the statistical classification of the documents apply to any handwriting classification task without any need of training. Finally, there is no need for knowing the number of the independent writing hands that the classification seeks for, which is an intrinsic necessity of the supervised approaches based on deep learning.

A first, concise flowchart of the methodology previously described, is given below (Fig. 1).

We would like to emphasize that, although all aforementioned actions and especially Action 1, may be applied to all alphabet symbols, we have restricted their application to a rather limited number of letters, according to the following syllogism:

- i. The selected alphabet symbols must manifest a considerable complexity, so that they may sufficiently characterize the Writer. Thus, for example, the Greek “ι” (“iota”), “κ” (“kappa”), “ν” (“nu”), “ο” (“omikron”), “τ” (“tau”) etc., are rather easily reproducible and so we expect to be correspondingly similar among different writers.
- ii. At the same time, the chosen letters must have a relatively high frequency of appearance in Greek documents, so as their processing offers statistically significant results. Therefore, for instance, alphabet symbols such as “ζ” (“zeta”), “ξ” (“xi”), “φ” (“phi”), “ψ” (“psi”) etc., have a very promising complexity, but, as a rule, they are rarely encountered



**Fig. 1** A flowchart of the general methodology for Writer Identification that has been applied to the present work, in order to test if the two unknown documents have been written by Rigas Feraios

in pages of Greek documents and, consequently, we do not expect that they may offer statistically significant results.

We would like to stress, that the introduced approach is applicable to the Latin alphabet symbols, (at least) too, as we shall demonstrate in future manuscripts.

In addition, we must make clear that the method presented here, by no means gives an answer concerning the authorship of these literary works. On the contrary, the goal of the introduced research, is to test the hypothesis that two or more documents have been written by the same hand. In order to test the authorship of a certain document, it is necessary to compare each style of writing with other texts, which are known to belong to a specific writer; this requires another, substantially different, approach in the disciplines of Mathematics and Computer Engineering. We plan to deal with this problem in connection with the aforementioned documents, in a future manuscript.

**A first stage processing of the realizations of any alphabet symbol**

We consider a part  $PD$  of any document  $D$ , where  $PD$  includes a sufficient number of realizations of the alphabet symbol, which is each time treated; the term “sufficient number” will be clarified in the analysis that follows. Moreover, we stress that in the subsequent presentation, we shall use alphabet symbol “ $\alpha$ ” and its realizations as a generic representative of all alphabet symbols and their realizations. Thus, when, in the following, a statement and/or a method is reported in connection with letter “ $\alpha$ ”, this means that the corresponding approach holds true for an arbitrary alphabet symbol, too.

The steps of this processing are briefly described below, and the obtained results are manifested in corresponding images:

**Step-FSP.1**

In order to semi-automatically isolate the realization of any, but specific, alphabet symbol, we have proceeded as follows:

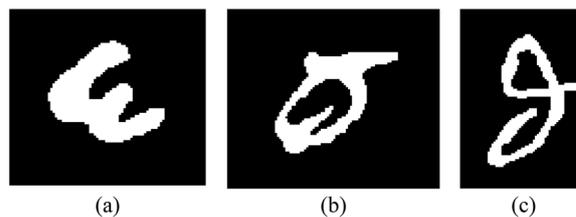
First, we have applied a method quite similar to the one introduced in [36]–[38], so as to determine the different document lines and then to divide each line into connected components. Subsequently, we have determined the points of change of letters’ realizations, following [36] again, and finally we have isolated the desired realization and we embedded it alone into a frame, say  $F_L$  (see Fig. 2).

**Step-FSP.2**

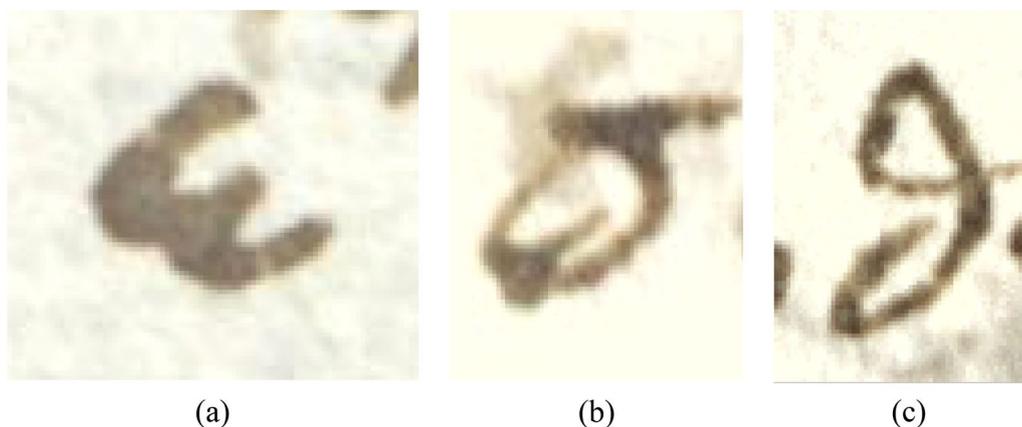
An automatic image segmentation method developed by the authors ([38]), has been applied to each one of the letters’ frames  $F_L$  obtained in Step-FSP.1.

**Step-FSP.3**

The contour of each isolated letter has been automatically extracted by means of a method specifically developed by



**Fig. 3** The result of the segmentation applied to the three letter images appearing in Fig. 2: for letter “ $e$ ” in **a**; for letter “ $\pi$ ” in **b**; for letter “ $\theta$ ” in **c**

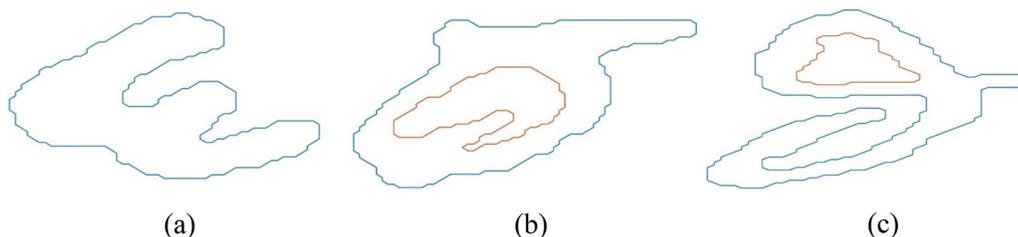


**Fig. 2** Three realizations of alphabet symbols extracted from “FYSAP”: **a** one realization of “ $e$ ”; **b** one of “ $\pi$ ”; **c** one of letter “ $\theta$ ”. The dimensions of the frames to which the realizations are embedded are kept fixed throughout the entire study

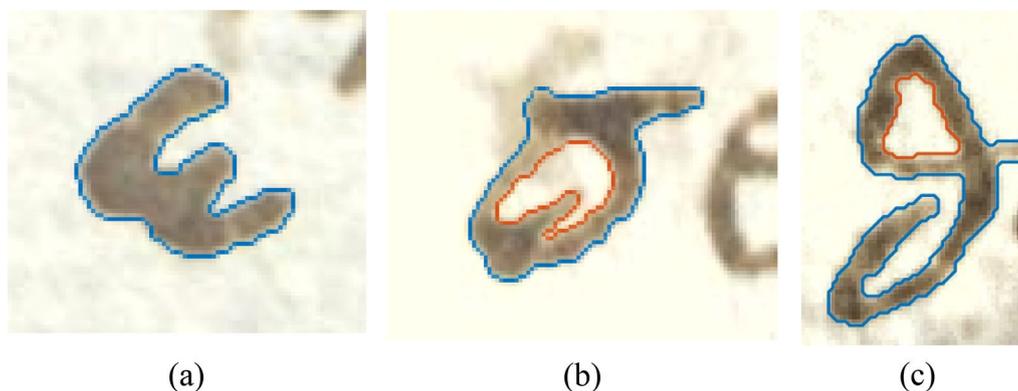
the authors (See Figs. 3, 4), always in  $F_L$ . We would like to emphasize that the obtained contours always constitute a union of closed simple Jordan curves, which separate the internal of the letter (its body) of the letter from its background. Thus, for example, most realizations of alphabet symbols “σ”, “β”, “θ”, “ρ” etc. are delimited by an external simple Jordan curve and at least an internal one, enclosed by the first borderline (e. g. see Figs. 4, 5, 6).

**Optimal matching of any two realizations of the same alphabet symbol**

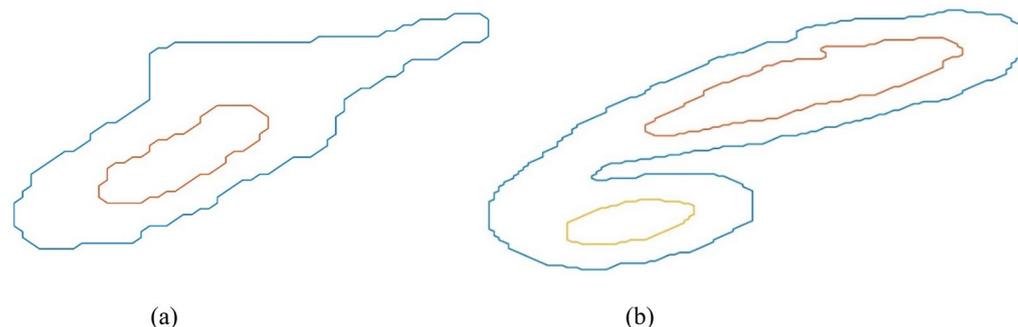
In this Section, we will state a first criterion of optimal matching (OM) of any two realizations of the same alphabet symbol, belonging either to the same document or to two different ones. In fact, this action includes the following steps:



**Fig. 4** The automatically extracted contours of the three alphabet symbol realizations of Fig. 2, standing alone



**Fig. 5** Demonstration of the fact that the extracted contours fit the corresponding alphabet symbols realizations (see Fig. 1), in a particularly good manner



**Fig. 6** **a** the border (contour) of a “σ” realization belonging to “FYSAP”: it consists of an outer simple closed Jordan curve (shown in blue) and an internal one, enclosed by the first (depicted in orange); **b** the contour of a “β” realization belonging to “SAGAN”: it consists of an outer simple closed Jordan curve (shown in blue) and two internal ones (shown in orange and yellow), which are enclosed by the external curve

**Step-OM.1**

We apply all steps referred to in the previous “A First Stage Processing of the Realizations of Any Alphabet Symbol” Section, thus obtaining the contours of all realizations of an arbitrary alphabet symbol “ $\alpha$ ” in part  $PD$  of a studied document. We randomly choose a realization of “ $\alpha$ ”, which we will call  $a_1$ , with a corresponding contour  $C_1^a$  and we temporarily let it be a prototype one.

We parallel translate  $C_1^a$  so as its center of mass coincides with the center  $(x_n, y_n)$  of the corresponding frame.

**Step-OM.2**

We let all other realizations of “ $\alpha$ ” in  $PD$  optimally fit “ $\alpha_1$ ”, by applying the subsequent affine transformations:

- A) Let  $a_i$  be an arbitrary realization in  $PD$  with contour  $C_i^a$  consisting of  $N_i$  pixels, each one having center coordinates  $(x_{i,j}, y_{i,j})$ , where  $j$  is the cardinal number of the contour pixel in hand. We parallel translate  $C_i^a$  so as its center of mass coincides with  $(x_n, y_n)$ .
- B) Next, using the well-known rotation matrix  $R = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ , we rotate  $C_i^a$  by the standard method, after translating its center of mass to the origin  $O(0, 0)$ . We symbolize the pixels of this rotated version of  $C_i^a$  as  $(x_{i,j}^{RO}, y_{i,j}^{RO})$ , where we have employed superscript  $RO$ , in order to indicate that the rotation takes place around the origin  $O$ .
- III) Subsequently, we apply scaling to  $C_i^{RO}$  by factor  $\lambda$  via the formula  $(x_{i,j}^{SRO}, y_{i,j}^{SRO}) = \lambda(x_{i,j}^{RO}, y_{i,j}^{RO})$ .
- IV) We re-evaluate the center of mass of pixels  $(x_{i,j}^{SRO}, y_{i,j}^{SRO})$  and we move it back to  $(x_n, y_n)$ , by the proper parallel translation.
- V) We parallel translate all pixels  $(x_{i,j}^{SRO}, y_{i,j}^{SRO})$  by the vector  $(dx, dy)$ .

**Step-OM.3**

For each quadruple  $(\varphi, \lambda, dx, dy)$ , restricted in a properly chosen 4-cube, we let  $C_i^{SRT}$  be the corresponding Scaled, Rotated and Translated version of  $C_i^a$ , as it has been obtained by the two aforementioned steps OM.1 and OM.2.

Next, we use the following symbolism:

- a) Let  $D(C_i^{SRT})$  be the domain enclosed by contour  $C_i^{SRT}$ , i. e. its internal.
- b) Similarly, let  $D(C_1^a)$  be the internal of the contour that played the role of the prototype-fixed one.

- c) For any closed domain, say  $D$ , in the plane we symbolize its area via  $Area(D)$ .

Then, we define the subsequent similarity criterion, abbreviated as SC, between “ $\alpha$ ” realizations  $C_1^a$  and  $C_i^a$ . Actually, this similarity criterion consists in employing  $C_i^{SRT}$  and  $C_1^a$  and evaluating quantity

$$\kappa(\varphi, \lambda, dx, dy) = \frac{Area(D(C_i^{SRT}) \cap D(C_1^a))}{Area(D(C_i^{SRT}) \cup D(C_1^a))} \tag{1}$$

Quantity  $\kappa(\varphi, \lambda, dx, dy)$  constitutes a reliable measure of similarity between contours  $C_i^{SRT}$  and  $C_1^a$ .

**Step-OM.4**

Among all computed values  $\kappa(\varphi, \lambda, dx, dy)$  in the chosen 4-cube, we select the maximum one, corresponding to the quadruple, say,  $(\varphi^{max}, \lambda^{max}, dx^{max}, dy^{max})$ . We symbolize this maximum value of  $\kappa$  as  $\kappa^{max}$ ; i.e.,  $\kappa^{max} = \kappa(\varphi^{max}, \lambda^{max}, dx^{max}, dy^{max})$ .

**Step-OM.5**

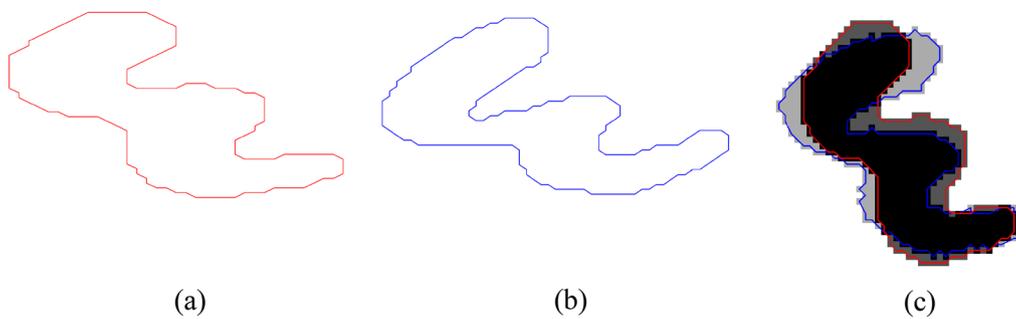
We parallel translate  $C_i^a$  by  $(dx^{max}, dy^{max})$ , we rotate it by  $\varphi^{max}$  and we scale it by  $\lambda^{max}$ , thus obtaining  $C_{i,OM}^{SRT}$ , which Optimally Matches  $C_1^a$  (see Figs. 7, 8, 9). In other words,  $\kappa^{max}$  constitutes a very reliable measure of similarity of the two original contours  $C_i^a$  and  $C_1^a$ .

**The “ideal representative” of an alphabet symbol, concerning a document of a specific writer**

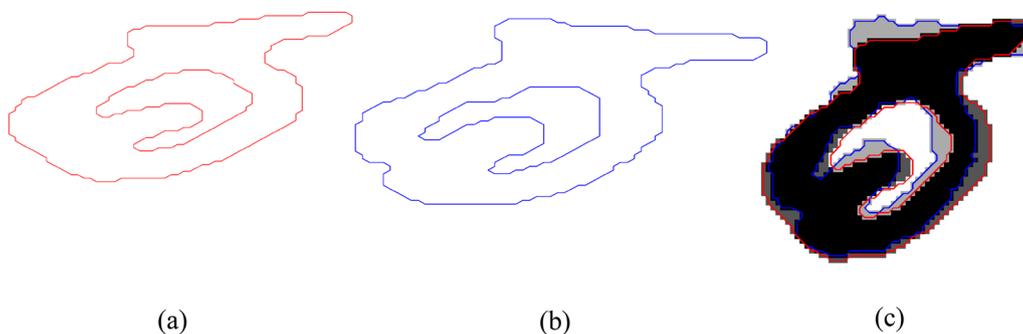
In the present Section, we shall employ the previous analysis, together with a novel approach, in order evaluate the “ideal representative” or the “platonic prototype”, of an arbitrary alphabet symbol associated with any document part  $PD$ . We assume that  $PD$  is written by the same hand; we note that we may test this hypothesis, before applying the process that follows. In the subsequent sections, we shall demonstrate that this ideal representative is a very powerful tool for writer identification. An associated flowchart for the Ideal Representative computation is given in Fig. 10 below.

We use the names “ideal” or “platonic” representatives for the following reasons:

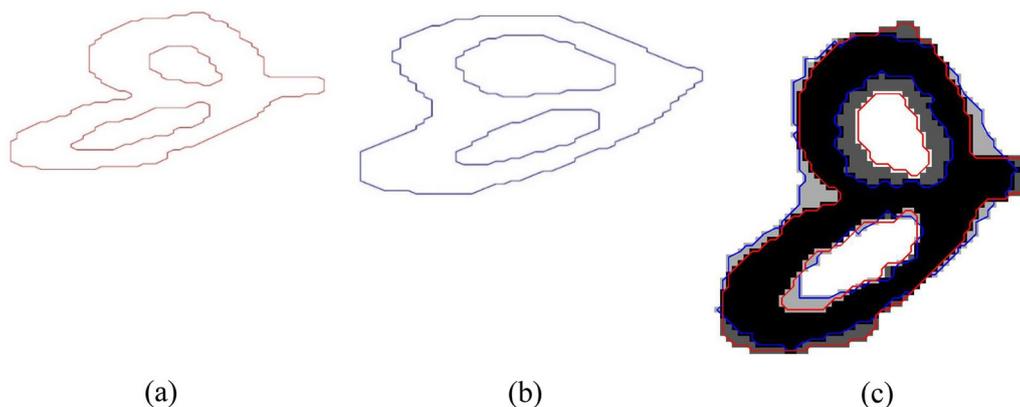
- a) We assume that each writer has one or at least a small number of ideal shapes in his/her mind, when writing a document. Evidently, this ideal shape may change during long periods of his/her life; however, we plausibly assume that, in most cases, the process of writing a specific document has a relatively small duration, as far as the letters’ ideal shapes are concerned.



**Fig. 7** **a, b**: Two distinct realizations of alphabet symbol “ $\epsilon$ ” appearing in “FYSAP”: the one in sub-figure **a** is considered to be the prototype-fixed one, while the realization in **b** is the current one; **c** Optimal matching of these two realizations, on the basis of the analysis introduced in the present sub-Sect. 3.2. We emphasize that the intersection of the bodies of the two letters is shown in black, while their union is depicted in degradations of grey; for completeness, in the figure we have added the two letters’ contours, each one with its color in **a, b**



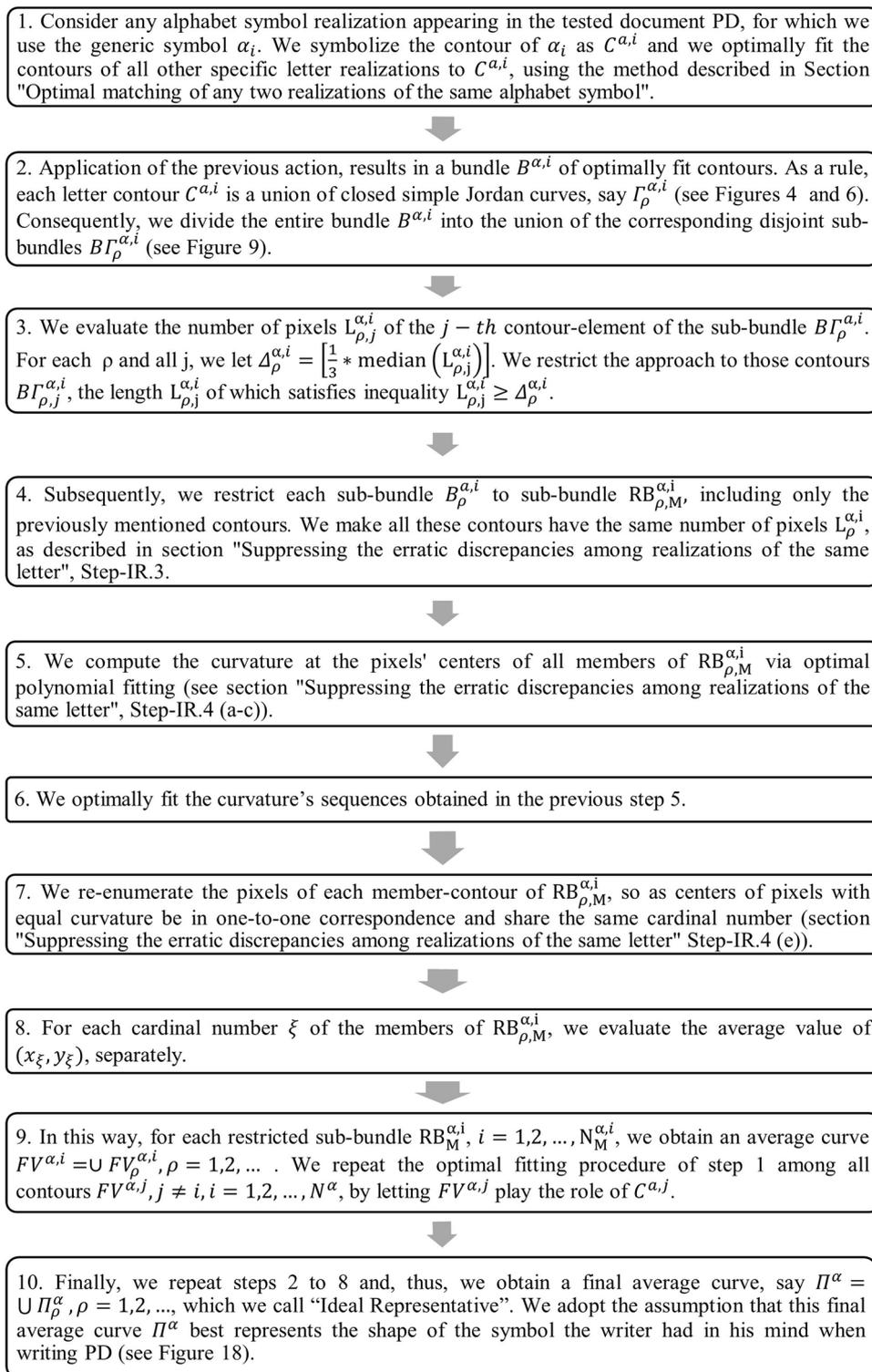
**Fig. 8** **a, b**: Two different realizations of letter “ $\pi$ ” also appearing in “FYSAP”: the one in sub-figure **a** is considered to be the prototype-fixed letter, while the realization in **b** is the current one; **c** Optimal fitting of these two realizations, according to the analysis introduced in the present sub-Sect. 3.2. Once more, the intersection of the bodies of the two letters is shown in black, while their union is depicted in degradations of grey; again, the contours of the alphabet symbols are shown colored



**Fig. 9** **a, b**: Two distinct realizations of alphabet symbol “ $\theta$ ” from “FYSAP”: the “ $\theta$ ” in sub-figure **a** plays the role of the prototype-fixed one, while the realization in **b** plays the role of the current one; **c** Optimal matching of these two realizations, using the analysis of the present sub-section; all conventions introduced in Figs. 3.6 and 3.7, have been also applied intact here

b) When a person renders a specific alphabet symbol on a writing material, the resulting letter realization is a disturbed version of the corresponding ideal shape

the writer has in his/her mind. The reasons that cause this disturbance, usually are pretty numerous: e. g. the psychological mood of the writer, her/his fatigue,



**Fig. 10** Flowchart of the method for obtaining the "Ideal Representative" of an alphabet symbol in connection with a document part PD written by a single writer

the interaction of the writing instrument with the writing material, possible defects of the writing materials and surfaces, the writer's age and perhaps many more.

- c) However, we may classify the discrepancies among the realizations of the same alphabet symbol generated by the same writer in two classes: i) the causal ones and ii) the erratic ones. The class of the causal discrepancies includes the size and the orientation of the realization as well as its different position in PD each time. The erratic discrepancies are mainly due to the reasons described in b) above.
- d) In order to account for the causal discrepancies, we applied the optimal fitting process described in "Optimal Matching of Any Two Realizations of the Same Alphabet Symbol" Section.

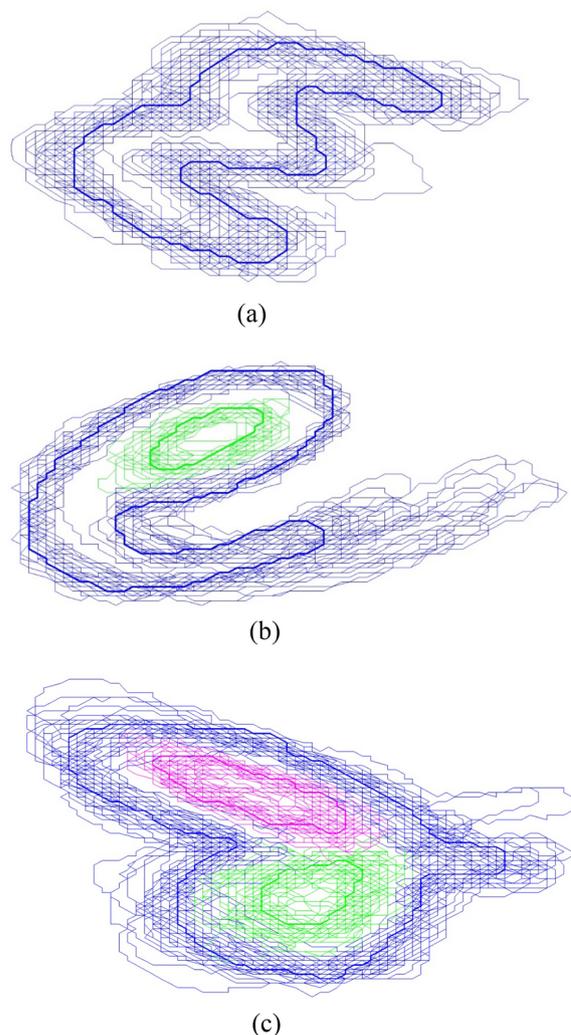
The process described in the present Section, aims at reducing the erratic discrepancies. This is achieved by a proper averaging process, given that averaging of  $N$  optimally fit curves reduces the average distance of these curves from the prototype one that has "generated" them, by a factor of  $\sqrt{N}$ , each time.

We shall apply both methods introduced in the present "The "Ideal Representative" of an Alphabet Symbol, Concerning a Document of a Specific Writer" Section, as well as in "Identification of Handwriting Based on the Ideal Representatives" Section, to a selected subset of eleven (11) greek alphabet symbols and more specifically to letters "α", "β", "γ", "ε", "θ", "λ", "μ", "π", "ρ", "σ", "ω", following the analysis made in the end of "A Summary and the Novelties of the Introduced Methodology" Section. We repeat here the two criteria that guided our choice:

- a) We have selected treating alphabet symbols, for which we a priori know/guess that might better convey the writing idiosyncrasies of a hand,
- b) We have focused our methodology on letters that have a significant frequency of appearance in all tested documents.

#### Suppressing the erratic discrepancies among realizations of the same letter

Consider an arbitrary alphabet symbol, symbolized as "α" and its realizations in PD, where we know a priori and/or we have verified that it has been written by a single hand. We optimally suppress both the causal and the erratic differences among the realizations of "α" in PD, so as to evaluate the "Ideal Representative" (IR) of "α" in this document. In order to achieve that, we apply the procedure consisting of the following steps:



**Fig. 11** **a** A bundle of optimally fit "ε" realizations according to the previous analysis and especially the one of "Optimal Matching of Any Two Realizations of the Same Alphabet Symbol" Section, all obtained from "FYSAP". **b** An analogous tuft of optimally "ρ" contours all extracted from "FYSAP". **c** A similar cluster of "θ" realizations, again belonging to "FYSAP"

#### Step-IR.1

We arbitrarily choose a first realization of "α" say "α<sub>1</sub>" with contour  $C^{\alpha,i}$  and we make all the pairwise optimal matching of "α<sub>1</sub>" with the other realizations α<sub>i</sub> of "α", with border line  $C^{\alpha,i}$ , in PD, where  $i = 2, 3, \dots, N^\alpha$ . Each such optimal pairwise fitting is accomplished by means of the method presented in "Optimal Matching of Any Two Realizations of the Same Alphabet Symbol" Section, where "α<sub>1</sub>" plays the role of a "prototype"-fixed realization, while α<sub>i</sub> of the current one. Each such pairwise fitting gives rise to a corresponding scaling factor  $\lambda_1^1$ , where superscript 1 stands for the cardinal number of the fixed

letter “ $\alpha_1$ ” and subscript  $i$  indicates the cardinal number of the current letter “ $\alpha_i$ ” optimally fit to “ $\alpha_1$ ” via the method presented in “Optimal Matching of Any Two Realizations of the Same Alphabet Symbol” Section.

**Step-IR.2**

We repeat this procedure by successively letting all other  $\alpha_i$  realizations in PD,  $i = 2, 3, \dots, N^\alpha$ , play the role of the fixed realization; in this way, we obtain a double matrix  $\Lambda_j^i$ , having as entries the scaling factors  $\lambda_j^i$ , which correspond to the optimal fitting of the letter  $\alpha_j$  to  $\alpha_i$ , where  $j = 1, 2, 3, \dots, N^\alpha$ , with  $j \neq i$ .

We confine ourselves to those entries of matrix  $\Lambda_j^i$ , which correspond to scaling factors  $\lambda_j^i$  belonging in the interval [0.55, 1.8]; we would like to emphasize that all alphabet symbol realizations appearing in the documents treated in the present work, as well as in Byzantine codices [4], do belong in this class. All the same, if a larger scaling factor interval is necessary to be considered, we firmly believe that the introduced approach remains totally applicable. Eventually, for each realization  $\alpha_i$ , when it plays the role of the prototype-fixed one, we obtain a bundle  $B^{\alpha,i}$ , consisting of all letter realizations  $\alpha_j$  optimally fit to  $\alpha_i$  (see Fig. 11). We repeat with emphasis that the aforementioned optimal matching is achieved using the contours  $C^{\alpha,i}$  of each realization  $\alpha_i$ , as described in “Optimal Matching of Any Two Realizations of the Same Alphabet Symbol” Section.

**Step-IR.3**

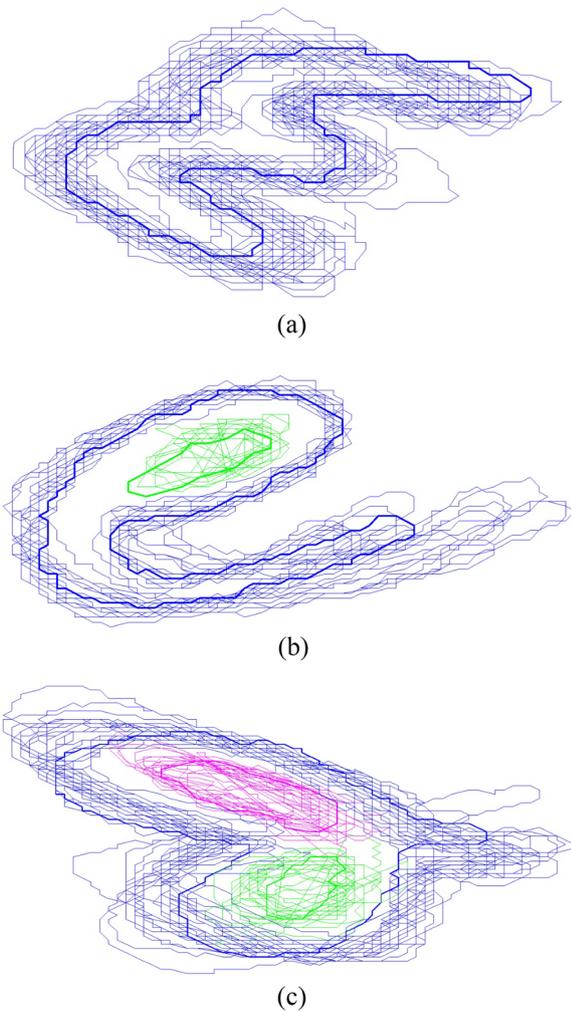
We consider that the border  $C^{\alpha,i}$  of an arbitrary realization  $\alpha_i$ , is a union of closed simple Jordan curves, enumerated in exactly the same way that will be immediately described:  $\Gamma_1^{\alpha,i}$  is the closed outer contour of  $\alpha_i$ ,  $\Gamma_2^{\alpha,i}$  is an eventual first internal contour, if exists,  $\Gamma_3^{\alpha,i}$  an eventual third internal contour, etc. We stress that, for a specific alphabet symbol, all  $\Gamma_\rho^{\alpha,i}$ ,  $\rho = 1, 2, 3, \dots$  curves are topologically analogous.

Consider the arbitrary realization  $\alpha_i$ , with contour  $C^{\alpha,i}$ , which momentarily plays the role of the prototype-fixed curve; for example, in Fig. 11c, the bundle  $B^{\theta,i}$  is presented, where the prototype-fixed contour of  $\theta_i$  is shown with a wider line. The  $C^{\theta,i}$  contour consists of three (3) closed simple curves  $\Gamma_\rho^{\theta,i}$ , where  $\rho = 1, 2, 3$ , presented with wider blue, magenta, and green lines respectively. As a consequence, bundle  $B^{\theta,i}$  includes three (3) sub-bundles  $B\Gamma_{\rho,j}^{\theta,i}$ ,  $\rho = 1, 2, 3$ , where subscript  $j$  expresses the cardinal number of the arbitrary  $j$ -th contour  $C^{\theta,i}$ . Evidently, for each letter  $\alpha$  and momentarily prototype realization  $\alpha_i$ , bundle  $B^{\alpha,i}$  is defined in an analogous manner, and it is the union of sub-bundles  $B\Gamma_{\rho,j}^{\alpha,i}$ ,  $j = 1, 2, \dots, N^\alpha$ .

For all elements of bundle  $B^{\alpha,i}$  and for each contour curve, say  $B\Gamma_{\rho,j}^{\alpha,i}$ ,  $\rho = 1, 2, \dots$ , we evaluate the number of pixels  $L_{\rho,j}^{\alpha,i}$ , where subscript  $j$  expresses the cardinal number of each “current” realization  $\alpha_j$  which is optimally fit to  $\alpha_i$ . Among all curves of the considered sub-bundle  $B\Gamma_{\rho}^{\alpha,i}$ , in connection with each realization  $\alpha_j \in B^{\alpha,i}$ , we choose the one which is nearest to the integer part  $\left[ \frac{1}{3} * \text{median} \left( L_{\rho,j}^{\alpha,i} \right) \right]$ ,  $j = 1, 2, \dots, N^\alpha$ ,  $i \neq j$ ; we symbolize this letter realization as  $\alpha_{i,M}$ , consisting of  $L_{\rho,M}^{\alpha,i}$  pixels (see Fig. 11). Subsequently, and always in connection with each  $\Gamma_\rho^{\alpha,i}$  separately, we proceed as follows:

- i. we keep those realizations of sub-bundle  $B\Gamma_\rho^{\alpha,i}$ , the number of pixels of which is equal to or greater than  $L_{\rho,M}^{\alpha,i}$ ; for the collection of these optimally fit sub-contours, we use the symbol  $RB_{\rho,M}^{\alpha,i}$ , including, say  $N_{\rho,M}^{\alpha,i}$  elements, where symbol RB stands for Restricted Bundle.
- ii. For each element (contour)  $\Gamma_{\rho,j}^{\alpha,i} \in RB_{\rho,M}^{\alpha,i}$ ,  $j > M$ , we compute the integer part of  $p_{\rho,j}^{\alpha,i}$ , were  $P_{\rho,j}^{\alpha,i} = \frac{L_{\rho,j}^{\alpha,i}}{L_{\rho,j}^{\alpha,i} - L_{\rho,M}^{\alpha,i}}$ .
- iii. In connection with each  $\Gamma_{\rho,j}^{\alpha,i} \in RB_{\rho,M}^{\alpha,i}$ , we remove the pixel with cardinal number  $\left( L_{\rho,j}^{\alpha,i} - p_{\rho,j}^{\alpha,i} \right)$  from the considered contour, bringing its two adjacent pixels in contact, so as the final digital curve “has no holes”. Next, we remove the pixel with cardinal number  $\left( L_{\rho,j}^{\alpha,i} - 2p_{\rho,j}^{\alpha,i} \right)$ , from the very same contour  $\Gamma_{\rho,j}^{\alpha,i}$ , bridging the resulting gap once more and so on.
- iv. We continue removing pixels  $\left( L_{\rho,j}^{\alpha,i} - m * p_{\rho,j}^{\alpha,i} \right)$ ,  $m = 3, 4, \dots$  from  $\Gamma_{\rho,j}^{\alpha,i}$ , following the procedure of step iii. The process is terminated at the  $r$ -th removal, provided that  $\left( L_{\rho,j}^{\alpha,i} - (r + 1) * p_{\rho,j}^{\alpha,i} \right) < 0$  holds.
- v. We close the digital curve, say  $RI_{\rho,j}^{\alpha,i}$ , obtained by this method, by connecting the last and first pixel of  $RI_{\rho,j}^{\alpha,i}$ .

We emphatically repeat that by the end of this procedure, all reduced members of  $RB_{\rho,M}^{\alpha,i}$  have all contours  $RI_{\rho,j}^{\alpha,i}$  with the same number of pixels  $L_{\rho}^{\alpha,i}$ ; needless to say, that, as a rule, the number of pixels  $L_{\rho}^{\alpha,i}$  concerning a specific contour  $RI_{\rho,j}^{\alpha,i}$  depends on  $\Gamma_{\rho,M}^{\alpha,i}$  (see Fig. 12). For this new version of the optimally fit realizations’ contours, we shall employ the symbol  $RB_{\rho,M}^{\alpha,i}$ , were the additional letter E in this symbolism, expresses the fact that all members of bundle  $RB_{\rho,M}^{\alpha,i}$  consist of digital curves with the same number of pixels. In a straightforward extension, we shall employ the symbol  $\alpha_{\rho,j}^{RE,i}$  for each member of this restricted bundle.

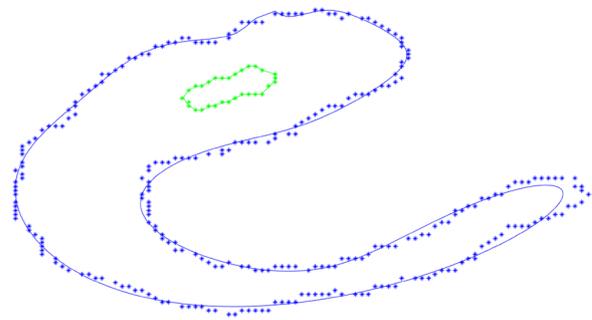


**Fig. 12** **a** A number of optimally fit contours of “ $\epsilon$ ” realizations, forming a bundle  $RB_{1,M}^{\epsilon,i}$ , where now, all involved curves consist of the same number of pixels, while they remain closed. **b** An ensemble of optimally fit contours of “ $\rho$ ” realizations. They form two (2) sub-bundles: first, the  $RB_{1,M}^{\rho,i}$  one, shown in blue, that corresponds to the outer contour of each letter realization, all consisting of the same number of pixels, namely  $L_1^{\rho,i}$ . Similarly, the sub-bundle  $RB_{2,M}^{\rho,i}$ , shown in green, includes the internal contours  $\Gamma_{2,M}^{\rho,i}$ , having the same number of pixels  $L_2^{\rho,i}$ . **c** An analogous bundle of optimally fit “ $\theta$ ” contours. The external closed contours  $RT_{1,M}^{\theta,i}$ , shown in blue, share the same number of pixels  $L_1^{\theta,i}$ ; similarly, the same holds true for the internal contours corresponding to  $\rho = 2$ , which are shown in green, and  $\rho = 3$ , shown in magenta

**Step-IR.4**

For each realization  $R\Gamma_{\rho,j}^{\alpha,i} \in RB_{\rho,j}^{\alpha,i}$ , we accomplish the following:

- a) We evaluate the arc length  $s_{\rho,j}^{\alpha,i}(k)$  for all pixels  $k$  of  $R\Gamma_{\rho,j}^{\alpha,i}$  in connection with the arbitrary-generic alphabet symbol “ $\alpha$ ”; we emphasize that length  $s_{\rho,j}^{\alpha,i}(k)$



**Fig. 13** Optimal approximation of the outer and internal contours of a “ $\rho$ ” realization obtained from “SAGAN”, by a polynomial of twenty-one (21) degree (for the outer contour) and a polynomial of eleven (11) degree for the internal one. The actual border line is depicted with asterisks (\*), while the polynomial approximation of both sub-contours is shown with a continuous line of the same colour.

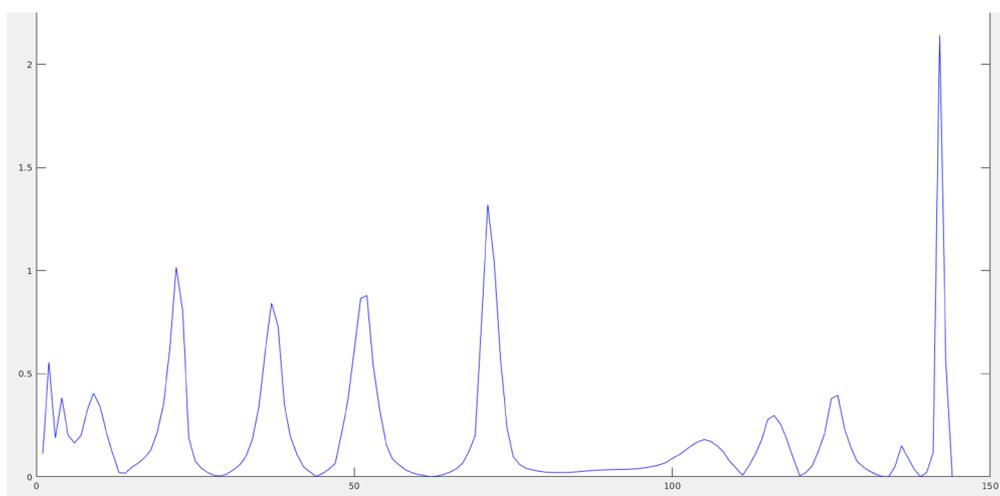
is evaluated recursively, starting from the origin of this discrete curve, i.e. from  $k = 1$ . We recall that, so far, the origin has been arbitrarily generated from the process of image segmentation and contour extraction of each symbol realization.

- b) We optimally approximate each contour  $R\Gamma_{\rho,j}^{\alpha,i}$  with two polynomials for the  $x$  and  $y$  coordinates of each pixel  $k$  of it, as a function of the arc lengths  $s_{\rho,j}^{\alpha,i}(k)$ . The degree of the polynomials has been chosen to be twenty-one (21) for the external (or the entire) contour  $R\Gamma_{1,j}^{\alpha,i}$  and of eleven (11) degree, for the internal contours  $R\Gamma_{\rho,j}^{\alpha,i}$ ,  $\rho = 2, 3, \dots$ . These degrees have proved to be very efficient and have offered pretty small error-distance and fluctuations, in connection with all studied alphabet symbols (see Fig. 13).

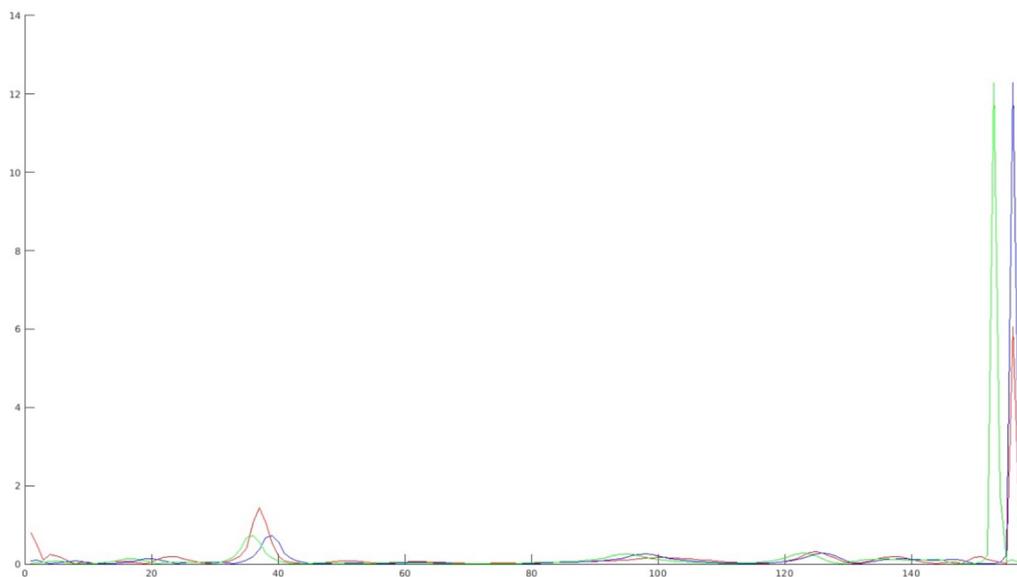
- c) We have computed the curvature of the aforementioned approximating polynomials and we have considered it to be the curvature of the actual contour at the specific point (see Fig. 14a).

- d) We have optimally fit the obtained sequences of curvatures for all  $R\Gamma_{\rho,j}^{\alpha,i} \in RB_{\rho,j}^{\alpha,i}$ , for a specific  $\rho$  and  $j = 1, 2, \dots, N_{\rho,M}^{\alpha,i}$ . This optimal fit has been realized by the quite standard method of shifting the curvature sequence of each  $R\Gamma_{\rho,j}^{\alpha,i}$  realization, evaluating each distance from the prototype one and keeping the position that offers the minimum error distance (see Fig. 14b).

- e) Finally, we have re-enumerated all pixels  $k$  of every considered contour  $R\Gamma_{\rho,j}^{\alpha,i}$ , so as pixels with the same cardinal number belong to the same class (see Figs. 15, 16).



(a)



(b)

**Fig. 14** **a** The curvature sequence obtained after approximating the entire contour of an “ $\varepsilon$ ” realization from “SAGAN”, shown in blue. **b** Optimal matching between two curvature sequences of the outer contours of two “ $\theta$ ” realizations belonging to  $RB_{\rho, M}^{\rho, i}$ , extracted from “FYSAP”: the prototype-fixed contour curvature is shown in red, the initial position of the curvature of the transformed “ $\theta$ ” realization is depicted in blue, while its translated version, so as it optimally fits the curve of the prototype curvature, is shown in green

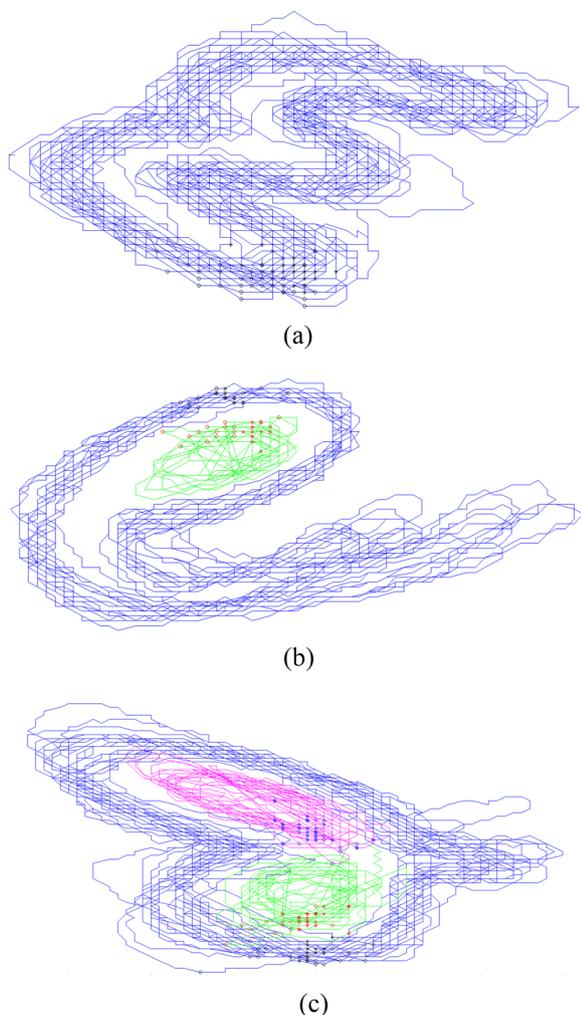
**Step-IR.5**

After accomplishing re-enumeration of all pixels for each contour of sub-bundle  $RB_{\rho, M}^{\alpha, i}$  and for each  $\rho$  separately, we have evaluated the mean value of the coordinates  $(x_{\rho, j}^{\alpha, i}, y_{\rho, j}^{\alpha, i})$ , in connection with all pixels sharing the same cardinal number  $k$  (see Fig. 17). Namely for each  $\rho$ -sub-bundle consisting of equinumerous contours, we compute the average value of the coordinates of all pixels that

share the same cardinal number. In this way, one obtains an average curve for each contours’ sub-bundle  $RB_{\rho, M}^{\alpha, i}$ , which we symbolize as  $FV_{\rho}^{\alpha, i}$  (FV stands for “First Version of the ideal representative”) (see Fig. 17).

**Step-IR.6**

We repeat steps IR.1 to IR.5 above, by letting all other realizations  $\alpha_j, j \neq i$  appearing in PD, play the role of

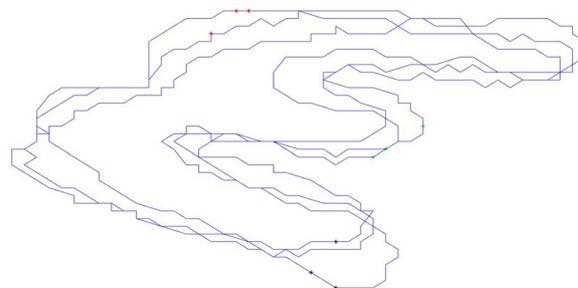


**Fig. 15** Re-enumeration of all pixels belonging to a bundle of optimally fit realizations of the same alphabet symbol ( $RB_{\rho_j}^{\alpha,i}$ ), appearing in the same text. **a** Concerning letter “e”; for clarity of presentation, we have indicated the ensemble of #1 pixels, only, in all matched realizations, as follows: the small circles (o) depict the initial/random position of the #1 pixels in each realization separately, while the asterisks (\*) indicate the position of #1 pixel, after re-enumeration. **b** Manifesting the re-enumeration of all optimally fit realizations of “e”, in connection with all curves  $R\Gamma_{1j}^{\rho_j,i}, R\Gamma_{2j}^{\rho_j,i}$ , fully respecting the conventions refer to in **a** above. **c** A completely analogous figure, manifesting re-enumeration of the contours of optimally fit “e” realizations, for all  $R\Gamma_{1j}^{\theta_j,i}, R\Gamma_{2j}^{\theta_j,i}, R\Gamma_{3j}^{\theta_j,i}$

the prototype letter. In this way, we end up with a number  $N^\alpha$ , of average curves  $FV^{\alpha,i} = \cup FV_{\rho}^{\alpha,i}, \rho = 1, 2, \dots$ , which we have already called “first versions of ideal representatives”.

**Step-IR.7**

We repeat steps IR.1 to IR.5 by letting  $FV^{\alpha,i}$  evaluated in step IR.6, play the role of the contours  $C^{\alpha,i}$  of



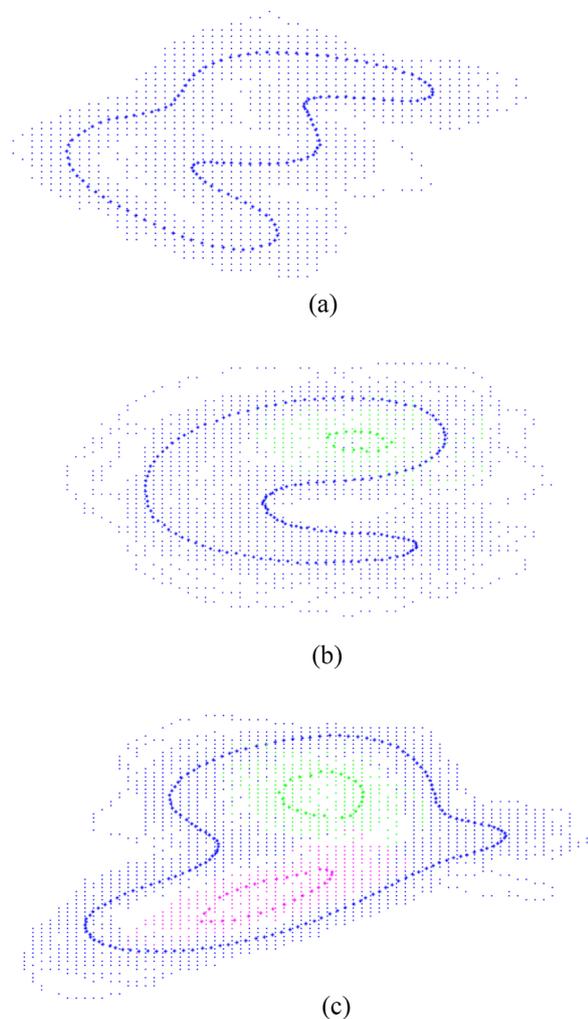
**Fig. 16** Further demonstration of the re-enumeration of the pixels of the contours of optimally fit “e” realizations: the ensemble of pixels bearing the cardinal number #1 is presented with black asterisks (\*), the pixels of cardinal number #40 are depicted with green asterisks (\*), while those bearing cardinal number #100 are depicted with red asterisks (\*). We stress, again, that in the initial re-enumeration of the contours’ pixels, all pixels that shared the same cardinal number were completely randomly placed

all realizations. Among all the resulting bundles of optimally fit digital curves  $FV^{\alpha,i}$  we choose the one with the minimum overall fitting error. Let the corresponding average curve be  $\Pi^\alpha = \cup \Pi_{\rho}^\alpha, \rho = 1, 2, \dots$ , where superscript  $\alpha$  refers to the letter in hand and subscript  $\rho$  to its contours (the external one and the internal, if any). We assume that this “final” union of curves  $\Pi^\alpha$  is the “ideal” or “platonic” representative of alphabet symbol  $\alpha$ , for the specific document part PD (see Fig. 18).

**Identification of handwriting based on the ideal representatives**

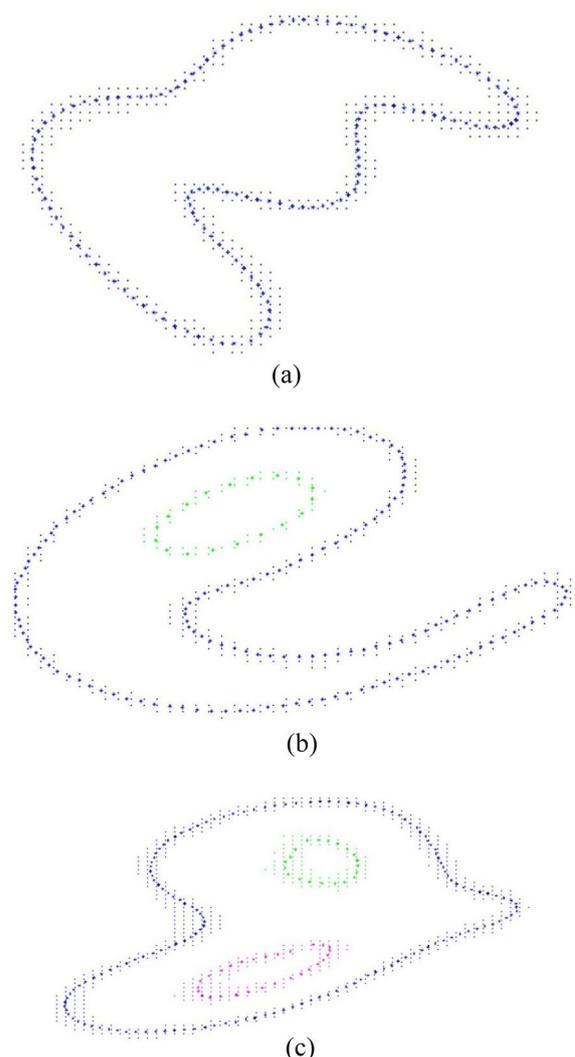
The method previously introduced has been applied to all selected alphabet symbols, namely, “α”, “β”, “γ”, “ε”, “θ”, “λ”, “μ”, “π”, “ρ”, “σ”, “ω”, for the reasons described in “A Summary and the Novelties of the Introduced Methodology” Section. In connection with each one of these symbols, we have extracted the ideal representatives associated with FYSAP, SAGAN, FILIA, SYNTAG and ELVEN. As a next step, we have applied the similarity criterion analytically discussed in “Optimal Matching of Any Two Realizations of the Same Alphabet Symbol” Section, to all pairs of “platonic prototypes”, corresponding to the aforementioned alphabet symbols (see Figs. 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29). The obtained comparison results are excellent as Figs. 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 manifest. We emphasize that in these Figures, we have adopted the following *convention*:

- A) The contour of letters of FYSAP are always depicted in blue, while the border lines of SAGAN in cyan, the contour of FILIA in green and those of SYNTAG in red.



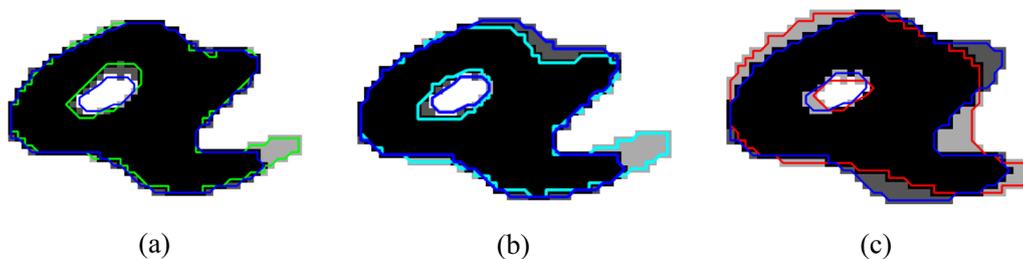
**Fig. 17** Demonstration of the method with which we obtain a first approximation of the ideal representative of “ε”, “ρ” and “θ”, all extracted from “FYSAP”. We have obtained this first “ideal approximation”, by averaging the coordinates of the re-enumerated contour pixels bearing the same cardinal number (as depicted in Figs. 13, 14); for each cardinal number, the corresponding mean values give rise to a single point shown with a colored asterisk in all three sub-figures. More specifically, the ideal representative’s first approximation is shown: **a** with deep blue asterisks, for “ε”; **b** with deep blue asterisks for the external contour of “ρ” and with green asterisks for the internal one; **c** with deep blue asterisks, for the external contour of “θ”, with green asterisks for the first internal contour and with magenta asterisks for the second internal one

B) The pixels of the union of the internals of two optimally fit ideal representatives are shown in degradations of grey, while the points (pixels) of their intersection is always shown in black.

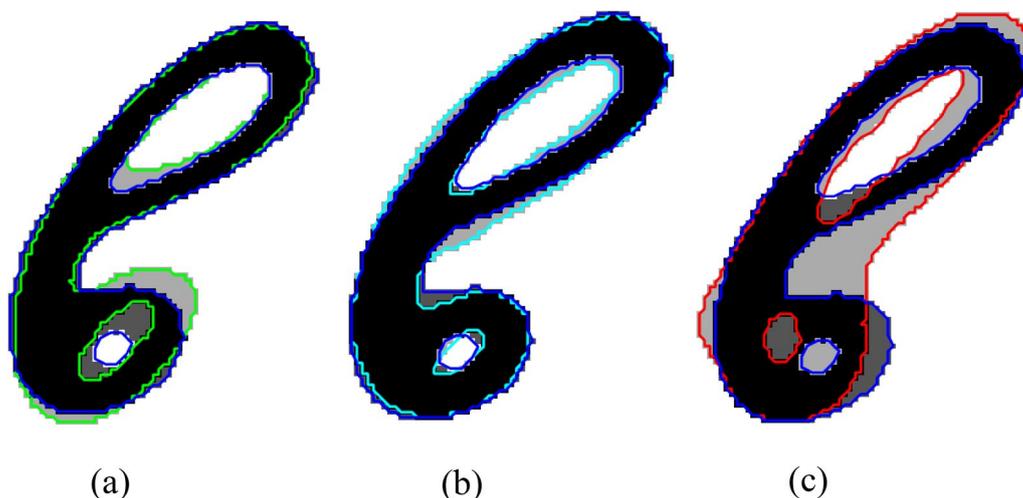


**Fig. 18** The finally obtained ideal representatives of all realizations of alphabet symbols “ε”, “ρ”, “θ”, belonging to the same text. These ideal representatives have been evaluated as it has been previously, analytically described. Evidently, in **a** the platonic representative of letter “ε” is presented, in **b** of “ρ” and in **c** of alphabet symbol “θ”, all extracted from FYSAP

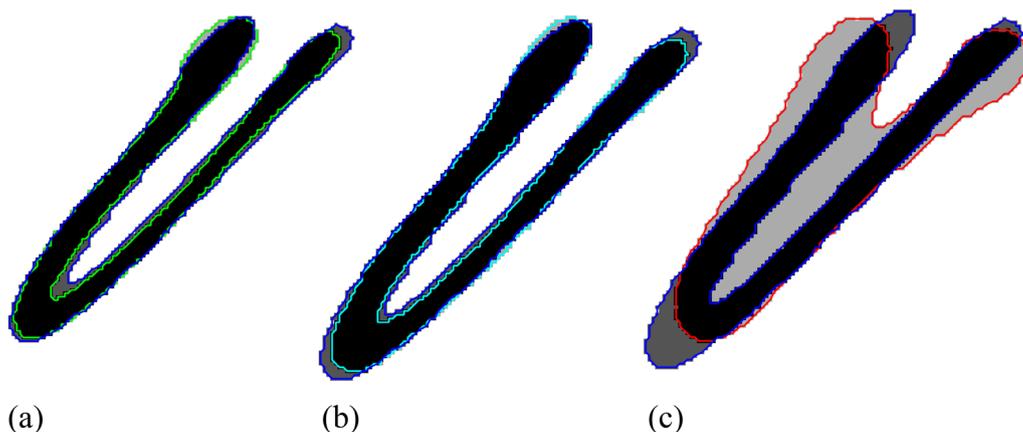
From Figs. 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, we feel that the optical representation of the matching of two ideal representatives of the same alphabet symbol, offers a very good indication if the letter considered each time comes from the same hand or not. However, in addition, the similarity criterion introduced in “Optimal Matching of Any Two Realizations of the Same Alphabet Symbol” Section, applied to any two optimally fit platonic prototypes will be given as a percentage, so as to achieve a clearer, quantitative understanding of the comparison results.



**Fig. 19** The ideal representative of “α” in connection with FYSAP, optimally fit to the corresponding platonic prototype **a** of FILIA (Similarity Criterion-SC: 89.65%), **b** of SAGAN (SC: 84.44%) and **c** of SYNTAG (SC: 78.77%)



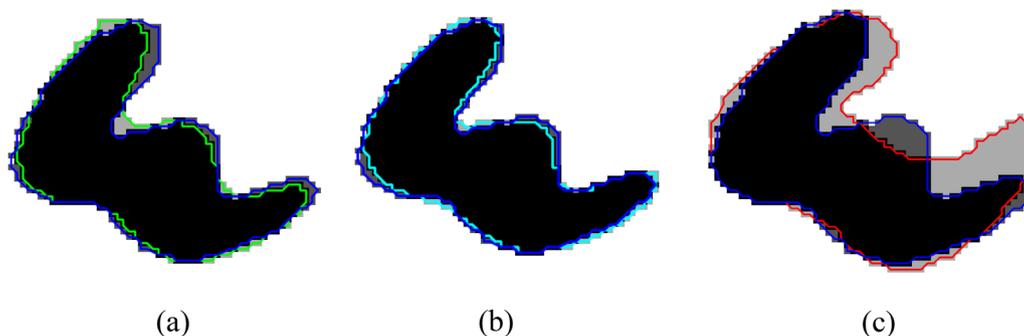
**Fig. 20** The platonic prototype of “β” of FYSAP, optimally matched to the associated ideal representative **a** of FILIA (SC: 84.3%), **b** of SAGAN (SC: 86.14%) and **c** of SYNTAG (SC: 65.85%)



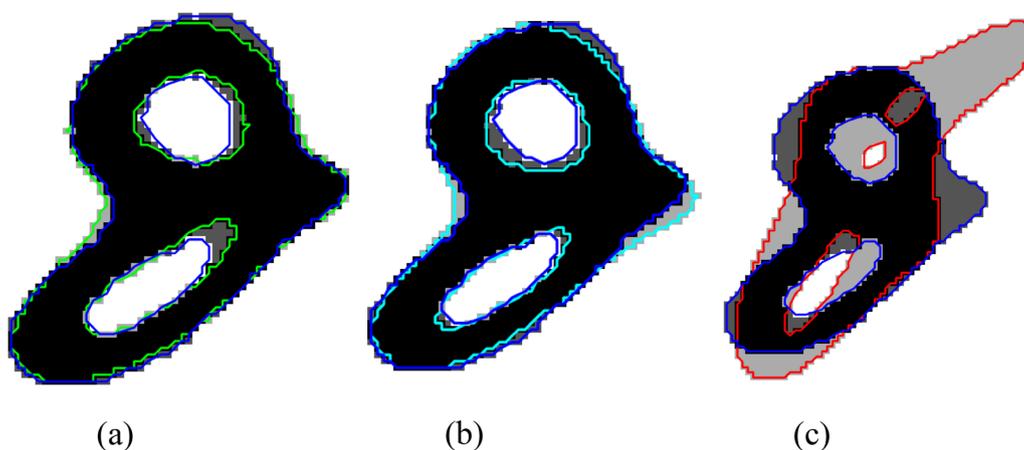
**Fig. 21** The ideal representative of letter “γ” in connection with FYSAP, optimally fit to the corresponding platonic prototype **a** of FILIA (SC: 82.27%), **b** of SAGAN (SC: 86%) and **c** of SYNTAG (SC: 56.68%)

Figures 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 above, together with the corresponding similarity criteria, constitute very strong evidence that documents FYSAP,

SAGAN and FILIA have been written by the same hand and in particular by that of Rigas Feraios, while document SYNTAG has not. Given that the content



**Fig. 22** The ideal representative of letter “ε” in connection with FYSAP, optimally matched to the associated platonic prototype **a** of FILIA (SC: 88.4%), **b** of SAGAN (SC: 91.92%) and **c** of SYNTAG (SC: 74.72%)



**Fig. 23** The platonic prototype of the letter “θ” in connection with FYSAP, optimally fit to the associated ideal representative **a** of FILIA (SC: 88.64%), **b** of SAGAN (SC: 87.02%) and **c** of SYNTAG (SC: 55.08%)

of SYNTAG undoubtedly belongs to Rigas Feraios, it results that the treated document of SYNTAG is a copy made by another hand.

We would like to emphasize that an analogous clear-cut, visual demonstration also holds true in connection with the platonic prototypes’ comparison of FYSAP with ELVEN, as well as SAGAN and FILIA with ELVEN. In fact, all these optimal matches clearly, visually manifest that ELVEN has not been written by Rigas Feraios; actually, the discrepancies among the ideal representatives of FYSAP, SAGAN and FILIA in one hand and ELVEN on the other, are more evident than the corresponding discrepancies concerning SYNTAG.

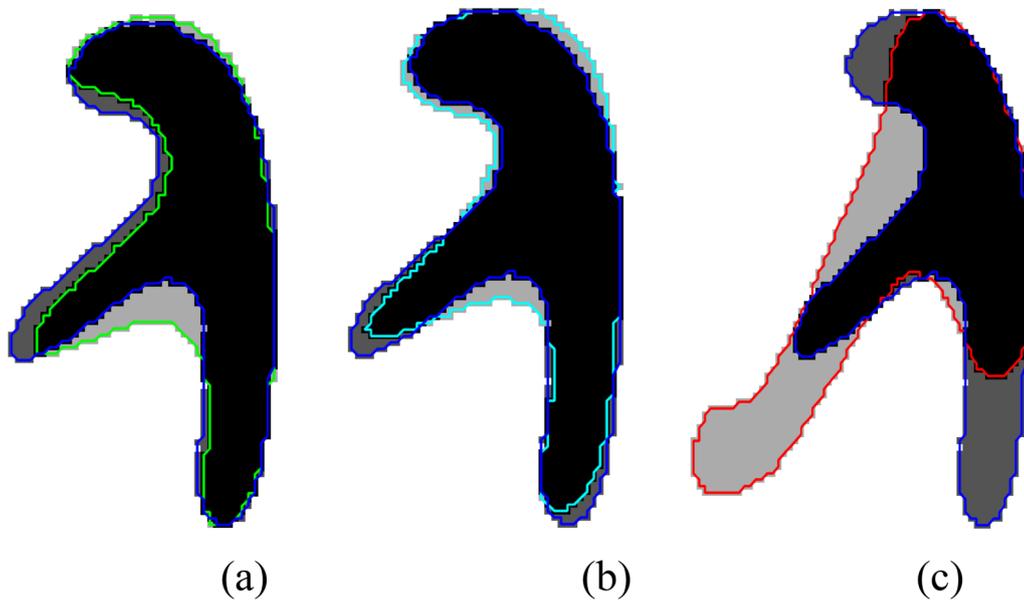
### A statistical approach for identifying Rigas Feraios’ handwriting

The serious problem concerning the verification if a number of texts belong to Rigas Feraios or not is the fact that there is only one document for which we are historically

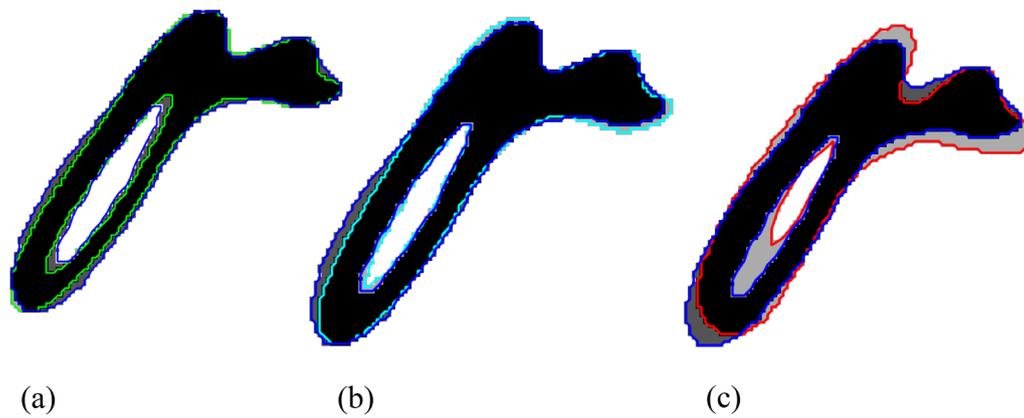
certain that it has been written by the hand of this great personality. Indeed, the only document that has been definitely written by Rigas’ hand is FYSAP (“Φυσικής Απάνθισμα”, “Compilation of Physics”). As a consequence, we cannot obtain statistical measure of similarity, so reliable as it would have been resulted after the comparison of SAGAN and of FILIA with two or even more documents undoubtedly written by Rigas’ hand.

To circumvent this difficulty as much as possible, we proceeded as follows:

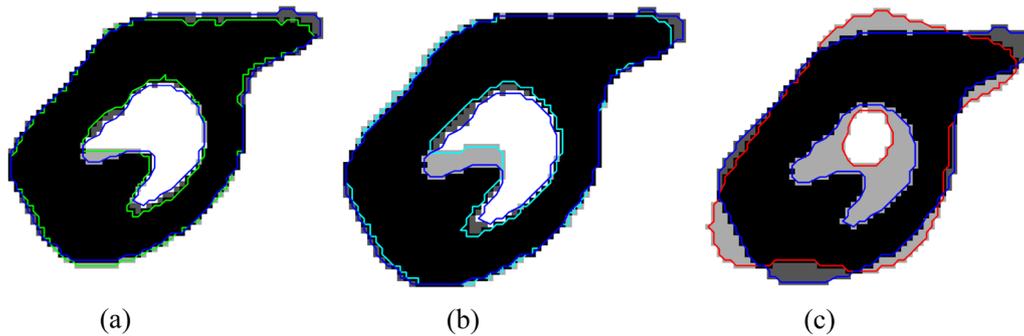
- i. consider any alphabet symbol, say “ε”, a statistically sufficient number of realizations of which, say  $N^\epsilon$ , are found in FYSAP. At a first step, we divide these  $N^\epsilon$  realizations into groups of nine (9) “ε” realizations randomly chosen from FYSAP. For each one of the selected 9-tuples we apply the method introduced in “The “Ideal Representative” of an Alphabet Symbol, Concerning a Document of a Specific Writer” Section and we extract an ideal



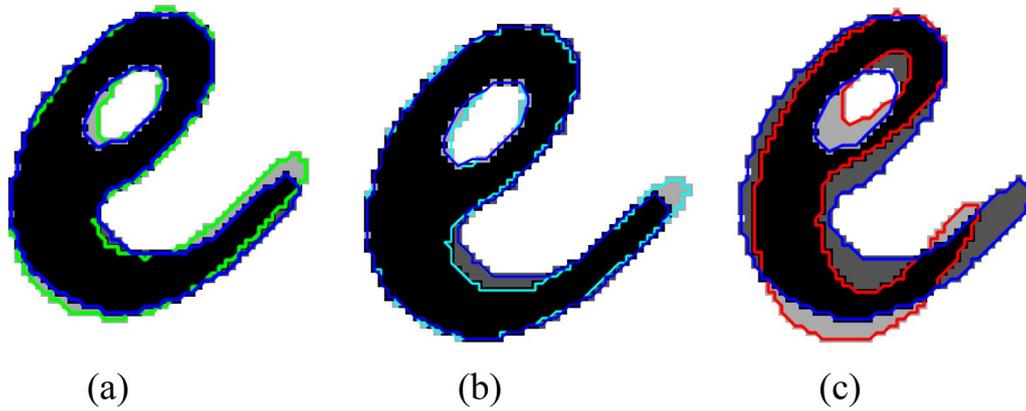
**Fig. 24** The platonic prototype of the alphabet symbol “λ” of FYSAP, optimally matched to the corresponding ideal representative **a** of FILIA (SC: 82.31%), **b** of SAGAN (SC: 85.07%) and **c** of SYNTAG (SC: 54.85%)



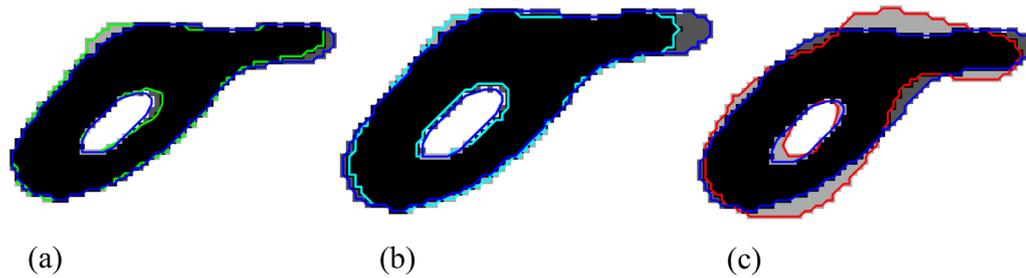
**Fig. 25** The ideal representative of letter “μ” concerning FYSAP, optimally fit to the corresponding platonic prototype **a** of FILIA (SC: 84.74%), **b** of SAGAN (SC: 87.86%) and **c** of SYNTAG (SC: 73.95%)



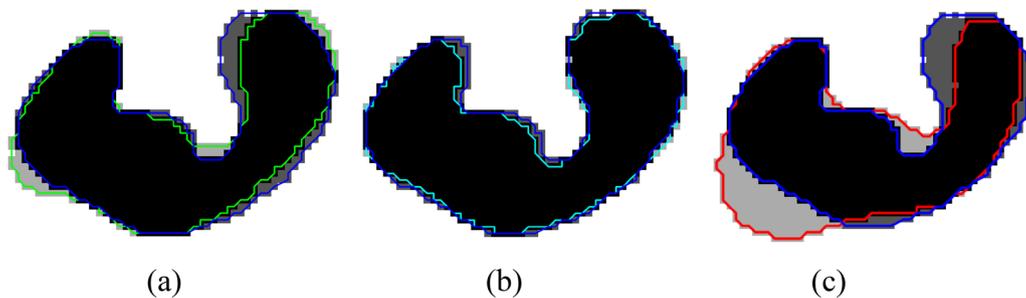
**Fig. 26** The ideal representative of letter “π” concerning FYSAP, optimally fit to the corresponding platonic prototype **a** of FILIA (SC: 89.18%), **b** of SAGAN (SC: 88.16%) and **c** of SYNTAG (SC: 73.97%)



**Fig. 27** The platonic prototype of letter “p” in connection with FYSAP, optimally matched to the associated ideal representative **a** of FILIA (SC: 86.42%), **b** of SAGAN (SC: 89.98%) and **c** of SYNTAG (SC: 61.9%)



**Fig. 28** The ideal representative of letter “o” regarding with FYSAP, optimally matched to the associated platonic prototype **a** of FILIA (SC: 91.2%), **b** of SAGAN (SC: 88.11%) and **c** of SYNTAG (SC: 75.77%)



**Fig. 29** The platonic prototype of letter “w” in connection with FYSAP, optimally matched to the associated ideal representative **a** of FILIA (SC: 85.5%), **b** of SAGAN (SC: 93.22%) and **c** of SYNTAG (SC: 74.46%)

representative of the realizations belonging to the first 9-tuple, which we symbolize as “ $\epsilon_{1,i}^9$ ”, where subscript 1 symbolizes the fact that it is the first division in 9-tuples, while the second subscript “1” stands for the cardinal number of the ideal representative of the ( $i - th$ ) 9-tuple of the current division. We emphasize again that the entire ensemble of “ $\epsilon$ ” realizations in FYSAP give rise to various, random divisions of them in 9-tuples.

- ii. we optimally match all pairs of these platonic prototypes  $(\epsilon_{1,i}^9, \epsilon_{1,j}^9)$ , where  $i, j = 2, \dots, \lfloor \frac{N^e}{9} \rfloor$  with  $i \neq j$ . Let  $m = \lfloor \frac{N^e}{9} \rfloor$ , i. e. the integer part of  $\frac{N^e}{9}$ ; then, the aforementioned optimal matches give rise to  $\frac{m(m-1)}{2}$  similarity criteria  $SC_1^{\epsilon^9}(i, j)$ , with  $i \neq j$ , as evaluated in “A Summary and the Novelities of the Introduced Methodology” Section.
- iii. We repeat steps 1 and 2 above for three (3) more randomly chosen divisions of  $N^e$  “ $\epsilon$ ” realizations

**Table 1** Manifestation of the way the values of the similarity criterion, between ideal representatives of n-tuples, depend on n

n: the number of same letter realizations that generate a platonic prototype	Average value of the similarity criterion, in connection with all selected n-tuples (%)
9	80.27
10	80.29
12	82.89
half	85.39
all	94.90

into 9-tuples, thus obtaining three new sets of similarity criteria  $SC_2^{\epsilon 9}(i, j), SC_3^{\epsilon 9}(i, j), SC_4^{\epsilon 9}(i, j)$ .

- iv. We reapply the three previous steps, dividing all  $N^\epsilon$  realizations of “ε” into 10-tuples; in this way, we obtain four (4) sets of similarity criteria,  $SC_1^{\epsilon 10}(i, j), SC_2^{\epsilon 10}(i, j), SC_3^{\epsilon 10}(i, j), SC_4^{\epsilon 10}(i, j)$ , with  $i, j = 1, \dots, \lfloor \frac{N^\epsilon}{10} \rfloor, i \neq j$ .
- v. We repeat step iv after dividing the “ε” realizations into 12-tuples, thus obtaining the sets of similarity criteria  $SC_1^{\epsilon 12}(i, j), SC_2^{\epsilon 12}(i, j), SC_3^{\epsilon 12}(i, j), SC_4^{\epsilon 12}(i, j)$ , with  $i, j = 1, \dots, \lfloor \frac{N^\epsilon}{12} \rfloor$ , always with  $i \neq j$ .
- vi. We, once more, apply the aforementioned step, for four (4) random divisions of  $N^\epsilon$  realizations into two (2) practically equinumerous sets, having  $\lfloor \frac{N^\epsilon}{2} \rfloor$  distinct elements in number. In other words, we randomly choose practically half the realizations of “ε”, we estimate their ideal representative, we let it play the role of the prototype letter and we repeat this process for the other half, which we consider to be the current letter. We apply this approach for three (3) additional random divisions of  $N^\epsilon$  into two (2) halves, in practice; consequently, we obtain four (4) similarity criteria  $SC_k^{\epsilon H^\epsilon}, k = 1, \dots, 4$ . It is logical to assume that there is an intimate relation between the number of elements of a n-tuple each time employed in one hand and of the obtained similarity measures on the other. Indeed, for example, one may expect that nine (9) realizations of the same alphabet symbol, say “ε”, convey less information of the overall handwriting style. The 10-tuples convey “a bit greater amount of related information”, the 12-tuples a little greater too, while half of the “ε” realizations appearing in the same document convey even greater amount of information; evidently, the maximum information is obtained, when all realizations of the tested alphabet symbol in the studied document are taken into consideration. Thus, for example, in connection with letter “λ” that has 66 appearances in various pages of

FYSAP, a typical sequence of corresponding similarity criteria is presented in Table 1.

- vii. We reapply steps i to vi above for all selected alphabet symbols, presented in “A Summary and the Novelties of the Introduced Methodology” Section.

Following this approach, for each selected Greek alphabet symbol, separately, we end up with a minimum  $SC_{min}^\alpha(FYSAP)$  and a maximum value  $SC_{max}^\alpha(FYSAP)$  of all the aforementioned similarity criteria associated with the document “Φυσικής Απάνθισμα”, which is historically certain that it has been written by Rigas’ hand. We assume that if the similarity criterion of FYSAP with a handwritten document, say PD, lies inside interval  $[SC_{min}^\alpha(FYSAP), SC_{max}^\alpha(FYSAP)]$ , then this is a measure indicating that the unknown document PD has been written by Rigas as far as the considered alphabet symbol is concerned.

If for all selected alphabet symbols, it holds that  $SC^\alpha(FYSAP, PD) \in [SC_{min}^\alpha(FYSAP), SC_{max}^\alpha(FYSAP)]$ , then we deduce that PD has been written by Rigas Feraios. Otherwise, we may make a statistical estimation of the likelihood that a document, say PD, has not been written by the hand of Rigas by first evaluating the similarity criterion of each letter’s platonic prototype from the corresponding one of FYSAP and at a second step, by computing the distance of this similarity criterion from the aforementioned intervals (see Table 2).

### Conclusion

In the present work, the notion and a method for evaluation of the “Ideal Representative” (or “Platonic Prototype”) are introduced, in connection with any alphabet symbol appearing in an arbitrary document part, say PD. We have shown that this ensemble of curves very well represents the associated ideal alphabet symbol that a writer had in his mind when was writing PD by hand. The stimulus for this approach was to give an answer to the important (for many nations) question if two handwritten documents discovered in Romania in 1998 by Chisacof and in particular the “Saganaki of Madness” (SAGAN) and the “Tested Friendship” (FILIA) have been written by the hand of Rigas Feraios.

Thus, we have explicitly evaluated and compared the Ideal Representatives of eleven (11), properly selected, alphabet symbols, associated with the aforementioned two documents, SAGAN and FILIA, but also with “Compilation of Physics” (FYSAP), which from the historical point of view, has been undoubtedly written by the hand of Rigas Feraios, “Constitution of the Greek State” (SYNTAG) and a document (ELVEN) written by another important Greek Politician, namely Eleftherios Venizelos.

**Table 2** Similarity Criteria of the Ideal Representatives

Alphabet symbol	Interval of similarity measures among n-tuples of FYSAP	Degree of similarity		
		FYSAP-FILIA	FYSAP-SAGAN	FYSAP-SYNTAG
“α”	[0.8371, 0.9584]	0.8965	0.8444	0.7877
“β”	[0.8185, 0.925]	0.843	0.8614	0.6585
“γ”	[0.765, 0.882]	0.8227	0.86	0.5668
“ε”	[0.805, 0.958]	0.8841	0.9192	0.7472
“θ”	[0.7729, 0.934]	0.8864	0.8702	0.5508
“λ”	[0.802, 0.95]	0.8231	0.8507	0.5485
“μ”	[0.7963, 0.9349]	0.8474	0.8786	0.7395
“π”	[0.826, 0.953]	0.8918	0.8816	0.7397
“ρ”	[0.788, 0.9313]	0.8642	0.8998	0.619
“σ”	[0.8162, 0.9455]	0.912	0.8811	0.7577
“ω”	[0.805, 0.962]	0.855	0.9322	0.7446

For each alphabet symbol separately, in each row, the interval of similarity criteria between the selected n-tuples is presented in the second column. The third column shows the similarity criteria between the three couples of ideal representatives: (i) FYSAP-FILIA, (ii) FYSAP-SAGAN, (iii) FYSAP-SYNTAG. The present Table manifests that documents SAGAN and FILIA have been written by the hand of Rigas Feraios while SYNTAG has not

Therefore, we have reached the conclusion that the two newly discovered documents SAGAN and FILIA have indeed been written by the hand of Rigas Feraios, while the celebrated “Constitution of the Greek State” has not, although its content undoubtedly belongs to Rigas. Moreover, the Ideal Representatives of the eleven (11) selected letters obtained from document ELVEN manifest serious and statistically important differences from the three documents belonging to Rigas.

#### Abbreviation

SC Similarity criterion

#### Acknowledgements

Not applicable.

#### Author contributions

All authors read and approved the final manuscript.

#### Funding

Not applicable.

#### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

#### Declarations

#### Competing interests

The authors declare that they have no competing interests.

Received: 14 September 2022 Accepted: 27 January 2023

Published online: 23 February 2023

#### References

1. “Home,” *National Research Foundation “Eleftherios K. Venizelos,”* 2022. <https://www.venizelos-foundation.gr/en/>
2. Mackridge P, Lia Brad Chisacof (ed.) Πήγας, ανέκδοτα κείμενα. athens: panepistimiakes ekdoseis kyprou and ekdoseis gutenber, 2011. Pp. 364. *Byzantine Mod Greek Stud.* 2016;37(2):295–7. <https://doi.org/10.1017/S030701310000687X>.
3. Panagopoulos M, Papaodysseus C, Rousopoulos P, Dafi D, Tracy S. Automatic writer identification of ancient greek inscriptions. *IEEE Trans Pattern Anal Mach Intell.* 2009;31(8):1404–14. <https://doi.org/10.1109/TPAMI.2008.201>.
4. Arabadjis Dd, Giannopoulos F, Panagopoulos M, Exarchos M, Blackwell C, Papaodysseus C. A general methodology for identifying the writer of codices. application to the celebrated ‘twins.’ *J Cult Herit.* 2019;39:186–201. <https://doi.org/10.1016/j.culher.2019.04.002>.
5. Schomaker L, Bulacu M. Automatic writer identification using connected-component contours and edge-based features of uppercase Western script. *IEEE Trans Pattern Anal Mach Intell.* 2004;26(6):787–98. <https://doi.org/10.1109/TPAMI.2004.18>.
6. Hannad Y, Siddiqi I, El Kettani MEY. Writer identification using texture descriptors of handwritten fragments. *Expert Syst Appl.* 2016;47:14–22. <https://doi.org/10.1016/j.eswa.2015.11.002>.
7. Zhang X, Nagy G. Computational method for calligraphic style representation and classification. *J Electron Imaging.* 2015;24(5):053003. <https://doi.org/10.1117/1.JEI.24.5.053003>.
8. Newell AJ, Griffin LD. Writer identification using oriented basic image features and the Delta encoding. *Pattern Recognit.* 2014;47(6):2255–65. <https://doi.org/10.1016/j.patcog.2013.11.029>.
9. Bensefia A, Paquet T. Writer verification based on a single handwriting word samples. *EURASIP J Image Video Process.* 2016;2016(1):34. <https://doi.org/10.1186/s13640-016-0139-0>.
10. He S, Samara P, Burgers J, Schomaker L. Historical manuscript dating based on temporal pattern codebook. *Comput Vis Image Underst.* 2016;152:167–75. <https://doi.org/10.1016/j.cviu.2016.08.008>.
11. Sarlin P. Self-organizing time map: an abstraction of temporal multivariate patterns. *Neurocomputing.* 2013;99:496–508. <https://doi.org/10.1016/j.neucom.2012.07.011>.
12. He S, Schomaker L. Writer identification using curvature-free features. *Pattern Recognit.* 2017;63:451–64. <https://doi.org/10.1016/j.patcog.2016.09.044>.
13. Faigenbaum-Golovin S, et al. Algorithmic handwriting analysis of Judah’s military correspondence sheds light on composition of biblical texts. *Proc Natl Acad Sci.* 2016;113(17):4664–9. <https://doi.org/10.1073/pnas.1522200113>.

14. Diem M, Kleber F, Fiel S, Sablatnig R. Semi-automated document image clustering and retrieval. *Document Recognit Retr.* 2014;9(21):206–15. <https://doi.org/10.1117/12.2043010>.
15. Shweka R, Choueka Y, Wolf L, Dershowitz N. Automatic extraction of catalog data from digital images of historical manuscripts. *Lit Linguist Comput.* 2013;28(2):315–30. <https://doi.org/10.1093/lc/fqt007>.
16. Wolf L, et al. Identifying join candidates in the cairo genizah. *Int J Comput Vis.* 2011;94(1):118–35. <https://doi.org/10.1007/s11263-010-0389-8>.
17. He S, Wiering M, Schomaker L. Junction detection in handwritten documents and its application to writer identification. *Pattern Recognit.* 2015;48(12):4036–48. <https://doi.org/10.1016/j.patcog.2015.05.022>.
18. De Stefano C, Maniaci M, Fontanella F, Scotto di Freca A. Reliable writer identification in medieval manuscripts through page layout features: the 'avila' bible case. *Eng Appl Artif Intell.* 2018;72:99–110. <https://doi.org/10.1016/j.engappai.2018.03.023>.
19. Cilia ND, De Stefano C, Fontanella F, Marrocco C, Molinara M, di Freca AS. An experimental comparison between deep learning and classical machine learning approaches for writer identification in medieval documents. *J Imaging.* 2020;6(9):9. <https://doi.org/10.3390/jimaging6090089>.
20. Zhang P. RSTC: a new residual swin transformer for offline word-level writer identification. *IEEE Access.* 2022;10:57452–60. <https://doi.org/10.1109/ACCESS.2022.3178597>.
21. Abbas F, Gattal A, Djeddi C, Siddiqi I, Bensefia A, Saoudi K. Texture feature column scheme for single- and multi-script writer identification. *IET Biom.* 2021;10(2):179–93. <https://doi.org/10.1049/bme2.12010>.
22. GR-RNN. Global-context residual recurrent neural networks for writer identification. *Pattern Recognit.* 2021;117:107975. <https://doi.org/10.1016/j.patcog.2021.107975>.
23. Semma A, Hannad Y, Siddiqi I, Djeddi C, El El Youssfi Kettani M. Writer identification using deep learning with fast keypoints and harris corner detector. *Expert Syst Appl.* 2021;184:115473. <https://doi.org/10.1016/j.eswa.2021.115473>.
24. Diamantatos P, Kavallieratou E, Gritzalis S. Directional hinge features for writer identification: the importance of the skeleton and the effects of character size and pixel intensity. *SN Comput Sci.* 2021;3(1):56. <https://doi.org/10.1007/s42979-021-00950-9>.
25. Javidi M, Jampour M. A deep learning framework for text-independent writer identification. *Eng Appl Artif Intell.* 2020;95:103912. <https://doi.org/10.1016/j.engappai.2020.103912>.
26. He S, Schomaker L. FragNet: writer identification using deep fragment networks. *IEEE Trans Inf Forensics Secur.* 2020;15:3013–22. <https://doi.org/10.1109/TIFS.2020.2981236>.
27. Bulacu M, Schomaker L. Text-independent writer identification and verification using textural and allographic features. *IEEE Trans Pattern Anal Mach Intell.* 2007;29(4):701–17. <https://doi.org/10.1109/TPAMI.2007.1009>.
28. A. Bensefia and C. Djeddi, "Relevance of Grapheme's Shape Complexity in Writer Verification Task," in *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, 2020, <https://doi.org/10.1109/IRI49571.2020.00016>
29. Chahi A, El Merabet Y, Ruichek Y, Touahni R. WriterNet: a multi-path deep CNN for offline text-independent writer identification. *Int J Doc Anal Recognit IJDAR.* 2022. <https://doi.org/10.1007/s10032-022-00418-3>.
30. Cilia ND, De Stefano C, Fontanella F, Marrocco C, Molinara M, Scotto Di Freca A. An end-to-end deep learning system for medieval writer identification. *Pattern Recognit Lett.* 2020;129:137–43. <https://doi.org/10.1016/j.patrec.2019.11.025>.
31. Chen Z, Yu H-X, Wu A, Zheng W-S. Letter-level online writer identification. *Int J Comput Vis.* 2021;129(5):1394–409. <https://doi.org/10.1007/s11263-020-01414-y>.
32. Adak C, Chaudhuri BB, Lin C-T, Blumenstein M. Intra-variable handwriting inspection reinforced with idiosyncrasy analysis. *IEEE Trans Inf Forensics Secur.* 2020;15:3567–79. <https://doi.org/10.1109/TIFS.2020.2991833>.
33. Lai S, Zhu Y, Jin L. Encoding pathlet and SIFT features with bagged VLAD for historical writer identification. *IEEE Trans Inf Forensics Secur.* 2020;15:3553–66. <https://doi.org/10.1109/TIFS.2020.2991880>.
34. Vásquez JL, Ravelo-García AG, Alonso JB, Dutta MK, Travieso CM. Writer identification approach by holistic graphometric features using off-line handwritten words. *Neural Comput Appl.* 2020;32(20):15733–46. <https://doi.org/10.1007/s00521-018-3461-x>.
35. Khosroshahi SNM, Razavi SN, Sangar AB, Majidzadeh K. Deep neural networks-based offline writer identification using heterogeneous handwriting data: an evaluation via a novel standard dataset. *J Ambient Intell Humaniz Comput.* 2022;13(5):2685–704. <https://doi.org/10.1007/s12652-021-03253-2>.
36. Bar-Yosef I, Beckman I, Kedem K, Dinstein I. Binarization, character extraction, and writer identification of historical Hebrew calligraphy documents. *Int J Doc Anal Recognit IJDAR.* 2007;9(2):89–99. <https://doi.org/10.1007/s10032-007-0041-5>.
37. Nomura S, Yamanaka K, Katai O, Kawakami H, Shiose T. A novel adaptive morphological approach for degraded character image segmentation. *Pattern Recognit.* 2005;38(11):1961–75. <https://doi.org/10.1016/j.patcog.2005.01.026>.
38. Papaodysseus C, Exarhos M, Panagopoulos M, Rousopoulos P, Triantafyllou C, Panagopoulos T. Image and pattern analysis of 1650 B.C. wall paintings and reconstruction. *IEEE Trans Syst Man Cybern Part Syst Hum.* 2008;38(4):4. <https://doi.org/10.1109/TSMCA.2008.923078>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)