

Dariusz Jacek Jakóbczak
Katedra Podstaw Informatyki i Zarządzania
Wydział Elektroniki i Informatyki
Politechnika Koszalińska

The Method of Probabilistic Nodes Combination in Handwriting Recognition

Keywords: Hurwitz-Radon Matrices, Active Object Modeling

1. Introduction

Handwriting identification and writer verification are still the open questions in artificial intelligence and computer vision. Handwriting based author recognition offers a huge number of significant implementations which make it an important research area in pattern recognition [1]. There are so many possibilities and applications of the recognition algorithms that implemented methods have to be concerned on a single problem. Handwriting and signature identification represents such a significant problem. In the case of writer recognition, described in this paper, each person is represented by the set of modeled letters or symbols. The sketch of proposed method consists of three steps: first handwritten letter or symbol must be modeled by a curve, then compared with unknown letter and finally there is a decision of identification. Author recognition of handwriting and signature is based on the choice of key points and curve modeling. Reconstructed curve does not have to be smooth in the nodes because a writer does not think about smoothing during the handwriting. Curve interpolation in handwriting identification is not only a pure mathematical problem but important task in pattern recognition and artificial intelligence such as: biometric recognition [2-4], personalized handwriting recognition [5], automatic forensic document examination [6,7], classification of ancient manuscripts [8]. Also writer recognition in monolingual handwritten texts is an extensive area of study and the methods independent from the language are well-seen. Proposed method represents language-independent and text-independent approach because it identifies the author via a single letter or symbol from the sample. This novel method is also applicable to short handwritten text.

Writer recognition methods in the recent years are going to various directions: writer recognition using multi-script handwritten texts [9], introduction of new features [10], combining different types of features [3], studying the sensitivity of character size on writer identification [11], investigating writer identification in

multi-script environments [9], impact of ruling lines on writer identification [12], model perturbed handwriting [13], methods based on run-length features [14,3], the edge-direction and edge-hinge features [2], a combination of codebook and visual features extracted from chain code and polygonized representation of contours [15], the autoregressive coefficients [9], codebook and efficient code extraction methods [16], texture analysis with Gabor filters and extracting features [17], using Hidden Markov Model [18-20] or Gaussian Mixture Model [1]. But no method is dealing with writer identification via curve modeling or interpolation and points comparing as it is presented in this paper.

The author wants to approach a problem of curve interpolation [21-23] and shape modeling [24] by characteristic points in handwriting identification. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Nowadays methods apply mainly polynomial functions, for example Bernstein polynomials in Bezier curves, splines and NURBS [25]. But Bezier curves do not represent the interpolation method and cannot be used for example in signature and handwriting modeling with characteristic points (nodes). Numerical methods for data interpolation are based on polynomial or trigonometric functions, for example Lagrange, Newton, Aitken and Hermite methods. These methods have some weak sides [26] and are not sufficient for curve interpolation in the situations when the curve cannot be build by polynomials or trigonometric functions. Proposed 2D curve interpolation is the functional modeling via any elementary functions and it helps us to fit the curve during handwriting identification.

This paper presents novel Probabilistic Nodes Combination (PNC) method of curve interpolation and takes up PNC method of two-dimensional curve modeling via the examples using the family of Hurwitz-Radon matrices (MHR method) [27], but not only (other nodes combinations). The method of PNC requires minimal assumptions: the only information about a curve is the set of at least two nodes. Proposed PNC method is applied in handwriting identification via different coefficients: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power. Function for PNC calculations is chosen individually at each modeling and it represents probability distribution function of parameter $\alpha \in [0;1]$ for every point situated between two successive interpolation knots. PNC method uses nodes of the curve $p_i = (x_i, y_i) \in \mathbf{R}^2$, $i = 1, 2, \dots, n$:

1. PNC needs 2 knots or more ($n \geq 2$);

2. If first node and last node are the same ($p_1 = p_n$), then curve is closed (contour);
3. For more precise modeling knots ought to be settled at key points of the curve, for example local minimum or maximum and at least one node between two successive local extrema.

Condition 3 means for example the highest point of the curve in a particular orientation, convexity changing or curvature extrema. The goal of this paper is to answer the question: how to model a handwritten letter or symbol by a set of knots [28]?

2. Probabilistic Interpolation

The method of PNC is computing points between two successive nodes of the curve: calculated points are interpolated and parameterized for real number $\alpha \in [0;1]$ in the range of two successive nodes. PNC method uses the combinations of nodes $p_1=(x_1,y_1)$, $p_2=(x_2,y_2), \dots, p_n=(x_n,y_n)$ as $h(p_1,p_2, \dots, p_m)$ and $m = 1,2, \dots, n$ to interpolate second coordinate y for first coordinate $c = \alpha \cdot x_i + (1-\alpha) \cdot x_{i+1}$, $i = 1,2, \dots, n-1$:

$$y(c) = \gamma \cdot y_i + (1 - \gamma)y_{i+1} + \gamma(1 - \gamma) \cdot h(p_1, p_2, \dots, p_m), \quad (1)$$

$$\alpha \in [0;1], \gamma = F(\alpha) \in [0;1].$$

Here are the examples of h computed for MHR method [29]:

$$h(p_1, p_2) = \frac{y_1}{x_1} x_2 + \frac{y_2}{x_2} x_1 \quad (2)$$

or

$$h(p_1, p_2, p_3, p_4) = \frac{1}{x_1 + x_3} (x_1 x_2 y_1 + x_2 x_3 y_3 + x_3 x_4 y_1 - x_1 x_4 y_3) + \frac{1}{x_2 + x_4} (x_1 x_2 y_2 + x_1 x_4 y_4 + x_3 x_4 y_2 - x_2 x_3 y_4).$$

The examples of other nodes combinations:

$$h(p_1, p_2) = \frac{y_1 x_2}{x_1 y_2} + \frac{y_2 x_1}{x_2 y_1}$$

or

$$h(p_1, p_2) = \frac{y_1 x_2}{y_2} + \frac{y_2 x_1}{y_1}$$

or

$$h(p_1, p_2) = x_1y_1 + x_2y_2$$

or

$$h(p_1, p_2) = x_1x_2 + y_1y_2$$

or

$$h(p_1, p_2, \dots, p_m) = 0$$

or

$$h(p_1) = x_1y_1$$

or others. Nodes combination is chosen individually for each curve. Formula (1) represents curve parameterization as $\alpha \in [0;1]$:

$$x(\alpha) = \alpha \cdot x_i + (1-\alpha) \cdot x_{i+1}$$

and

$$y(\alpha) = F(\alpha) \cdot y_i + (1 - F(\alpha))y_{i+1} + F(\alpha)(1 - F(\alpha)) \cdot h(p_1, p_2, \dots, p_m),$$

$$y(\alpha) = F(\alpha) \cdot (y_i - y_{i+1} + (1 - F(\alpha)) \cdot h(p_1, p_2, \dots, p_m)) + y_{i+1}.$$

Proposed parameterization gives us the infinite number of possibilities for curve calculations (determined by choice of F and h) as there is the infinite number of human signatures, handwritten letters and symbols. Nodes combination is the individual feature of each modeled curve (for example a handwritten letter or signature). Coefficient $\gamma = F(\alpha)$ and nodes combination h are key factors in PNC curve interpolation and shape modeling.

2.1. Interpolating functions in PNC modeling

Points settled between the nodes are computed using PNC method. Each real number $c \in [a;b]$ is calculated by a convex combination $c = \alpha \cdot a + (1 - \alpha) \cdot b$ for

$$\alpha = \frac{b-c}{b-a} \in [0;1].$$

Key question is dealing with coefficient γ in (1). The simplest way of PNC calculation means $h = 0$ and $\gamma = \alpha$ (basic probability distribution). Then PNC represents a linear interpolation. MHR method [30] is not a linear interpolation. MHR [31] is the example of PNC modeling. Each interpolation requires specific distribution of parameter α and γ (1) depends on parameter $\alpha \in [0;1]$:

$$\gamma = F(\alpha), F:[0;1] \rightarrow [0;1], F(0) = 0, F(1) = 1$$

and F is strictly monotonic. Coefficient γ is calculated using different functions (polynomials, power functions, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan or arc cot, also inverse functions) and choice of function is connected with initial requirements and curve specifications. Different values of coefficient γ are connected with applied functions $F(\alpha)$. These functions

$\gamma = F(\alpha)$ represent the examples of probability distribution functions for random variable $\alpha \in [0;1]$ and real number $s > 0$:

$$\begin{aligned} \gamma = \alpha^s, \quad \gamma = \sin(\alpha^s \cdot \pi/2), \quad \gamma = \sin^s(\alpha \cdot \pi/2), \quad \gamma = 1 - \cos(\alpha^s \cdot \pi/2), \quad \gamma = 1 - \cos^s(\alpha \cdot \pi/2), \\ \gamma = \tan(\alpha^s \cdot \pi/4), \quad \gamma = \tan^s(\alpha \cdot \pi/4), \quad \gamma = \log_2(\alpha^s + 1), \quad \gamma = \log_2^s(\alpha + 1), \quad \gamma = (2^\alpha - 1)^s, \\ \gamma = 2/\pi \cdot \arcsin(\alpha^s), \quad \gamma = (2/\pi \cdot \arcsin \alpha)^s, \quad \gamma = 1 - 2/\pi \cdot \arccos(\alpha^s), \quad \gamma = 1 - (2/\pi \cdot \arccos \alpha)^s, \\ \gamma = 4/\pi \cdot \arctan(\alpha^s), \quad \gamma = (4/\pi \cdot \arctan \alpha)^s, \quad \gamma = \text{ctg}(\pi/2 - \alpha^s \cdot \pi/4), \quad \gamma = \text{ctg}^s(\pi/2 - \alpha \cdot \pi/4), \\ \gamma = 2 - 4/\pi \cdot \text{arcctg}(\alpha^s), \quad \gamma = (2 - 4/\pi \cdot \text{arcctg} \alpha)^s. \end{aligned}$$

Functions above, used in γ calculations, are strictly monotonic for random variable $\alpha \in [0;1]$ as $\gamma = F(\alpha)$ is probability distribution function. Also inverse functions $F^{-1}(\alpha)$ are appropriate for γ calculations. Choice of function and value s depends on curve specifications and individual requirements. Considering nowadays used probability distribution functions for random variable $\alpha \in [0;1]$ - one distribution is dealing with the range $[0;1]$: beta distribution. Probability density function f for random variable $\alpha \in [0;1]$ is:

$$f(\alpha) = c \cdot \alpha^s \cdot (1 - \alpha)^r, \quad s \geq 0, r \geq 0. \tag{3}$$

When $r = 0$ probability density function (3) represents $f(\alpha) = c \cdot \alpha^s$ and then probability distribution function F is like $f(\alpha) = 3\alpha^2$ and $\gamma = \alpha^3$. If s and r are positive integer numbers then γ is the polynomial, for example $f(\alpha) = 6\alpha(1 - \alpha)$ and $\gamma = 3\alpha^2 - 2\alpha^3$. Beta distribution gives us coefficient γ in (1) as polynomial because of interdependence between probability density f and distribution F functions:

$$f(\alpha) = F'(\alpha), \quad F(\alpha) = \int_0^\alpha f(t) dt. \tag{4}$$

For example (4): $f(\alpha) = \alpha \cdot e^\alpha$ and $\gamma = F(\alpha) = (\alpha - 1)e^\alpha + 1$.

What is very important in PNC method: two curves (for example a handwritten letter or signature) may have the same set of nodes but different h or γ results in different interpolations (Fig.6-14).

Algorithm of PNC interpolation and modeling (1) looks as follows:

Step 1: Choice of knots p_i at key points.

Step 2: Choice of nodes combination $h(p_1, p_2, \dots, p_m)$.

Step 3: Choice of distribution $\gamma = F(\alpha)$.

Step 4: Determining values of α : $\alpha = 0.1, 0.2 \dots 0.9$ (nine points) or $0.01, 0.02 \dots 0.99$ (99 points) or others.

Step 5: The computations (1).

These five steps can be treated as the algorithm of PNC method of curve modeling and interpolation (1).

Curve interpolation has to implement the coefficients γ . Each strictly monotonic function F between points (0;0) and (1;1) can be used in PNC interpolation.

3. Handwriting Modeling and Recognition

PNC method enables signature and handwriting recognition. This process of recognition consists of three parts:

1. Modeling – choice of nodes combination and probabilistic distribution function (1) for known signature or handwritten letters;
2. Unknown writer - choice of characteristic points (nodes) for unknown signature or handwritten word and the coefficients of points between nodes;
3. Decision of recognition - comparing the results of PNC interpolation for known models with coordinates of unknown text.

3.1. Modeling – the basis of patterns

Known letters or symbols ought to be modeled by the choice of nodes, determining specific nodes combination and characteristic probabilistic distribution function. For example a handwritten word or signature “ rw ” may look different for persons A, B or others. How to model “ rw ” for some persons via PNC method? Each model has to be described by the set of nodes for letters “ r ” and “ w ”, nodes combination h and a function $\gamma=F(\alpha)$ for each letter. Less complicated models can take $h(p_1, p_2, \dots, p_m)=0$ and then the formula of interpolation (1) looks as follows:

$$y(c) = \gamma \cdot y_i + (1 - \gamma)y_{i+1}.$$

It is linear interpolation for basic probability distribution ($\gamma = \alpha$). How first letter “ r ” is modeled in three versions for nodes combination $h = 0$ and $\alpha=0.1, 0.2 \dots 0.9$? Of course α is a random variable and $\alpha \in [0;1]$.

Person A

Nodes (1;3), (3;1), (5;3), (7;3) and $\gamma = F(\alpha) = \alpha^2$:

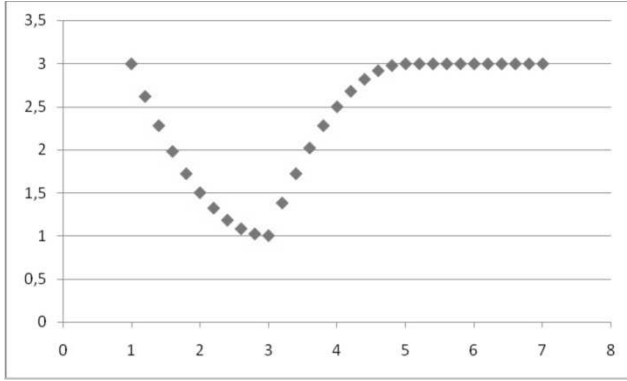


Fig. 1. PNC modeling for nine reconstructed points between nodes

Person B

Nodes (1;3), (3;1), (5;3), (7;2) and $\gamma = F(\alpha) = \alpha^2$:

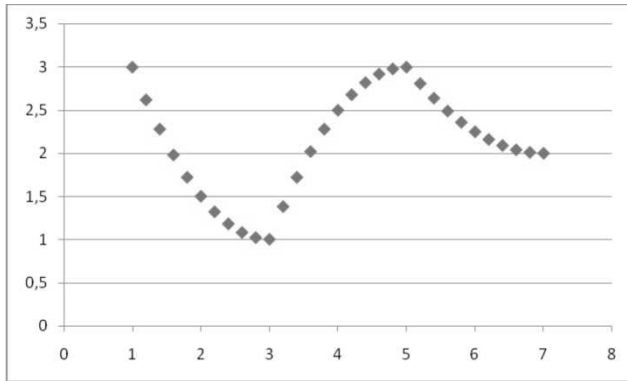


Fig. 2. PNC modeling of letter “r” with four nodes

Person C

Nodes (1;3), (3;1), (5;3), (7;4) and $\gamma = F(\alpha) = \alpha^3$:

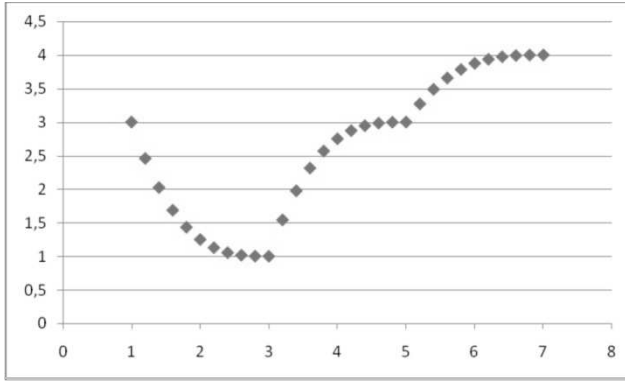


Fig. 3. PNC modeling of handwritten letter “r”

These three versions of letter “r” (Fig.1-3) with nodes combination $h = 0$ differ at fourth node and probability distribution functions $\gamma = F(\alpha)$. Much more possibilities of modeling are connected with a choice of nodes combination $h(p_1, p_2, \dots, p_m)$. MHR method [32] uses the combination (2) with good features because of orthogonal rows and columns at Hurwitz-Radon family of matrices:

$$h(p_i, p_{i+1}) = \frac{y_i}{x_i} x_{i+1} + \frac{y_{i+1}}{x_{i+1}} x_i$$

and then (1)

$$y(c) = \gamma \cdot y_i + (1 - \gamma)y_{i+1} + \gamma(1 - \gamma) \cdot h(p_i, p_{i+1}).$$

Here are two examples of PNC modeling with MHR combination (2).

Person D

Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^2$:

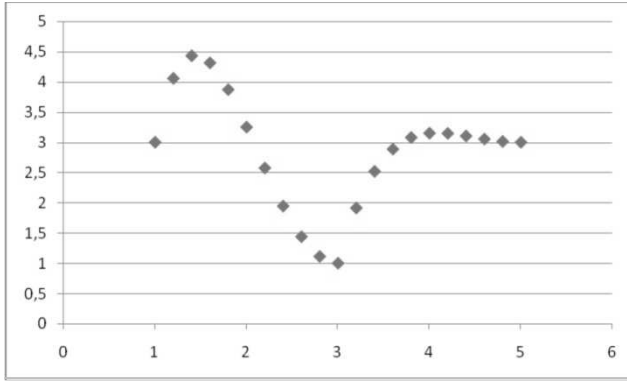


Fig. 4. PNC modeling of letter “r” with three nodes

Person E

Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^{1.5}$:

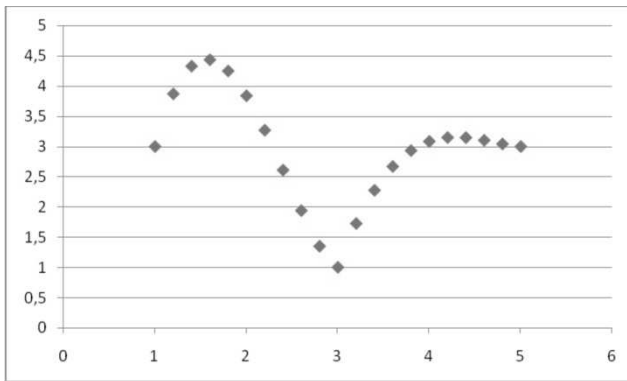


Fig. 5. PNC modeling of handwritten letter “r”

Fig.1-5 show modeling of letter “r”. Now let us consider a letter “w” with nodes combination $h = 0$.

Person A

Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = (5^\alpha - 1)/4$:

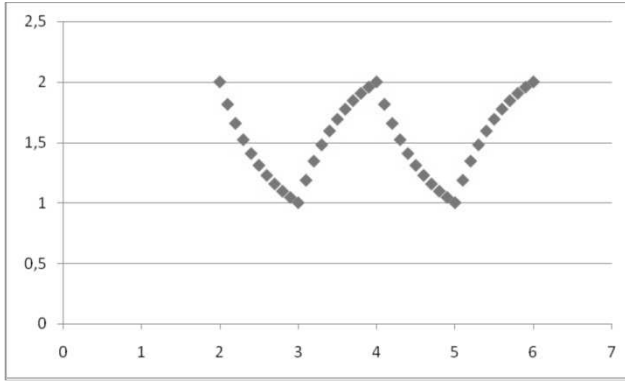


Fig. 6. PNC modeling for nine reconstructed points between nodes

Person B

Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = \sin(\alpha \cdot \pi/2)$:

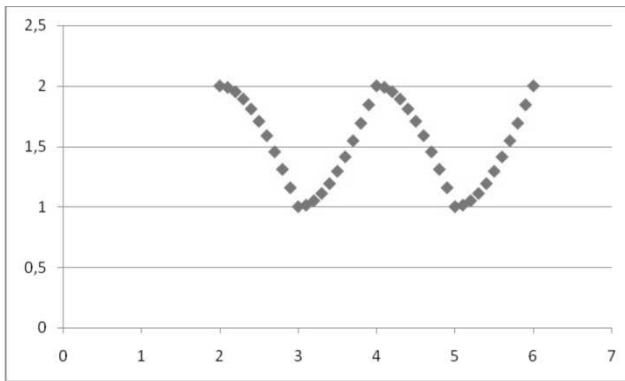


Fig. 7. PNC modeling of letter “w” with five nodes

Person C

Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = \sin^{3.5}(\alpha \cdot \pi/2)$:

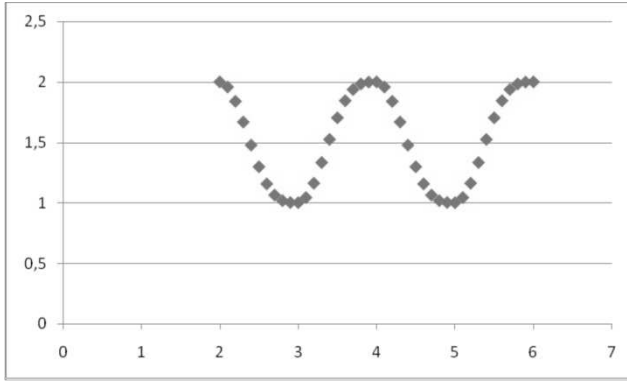


Fig. 8. PNC modeling of handwritten letter “w”

These three versions of letter “w” (Fig.6-8) with nodes combination $h = 0$ and the same nodes differ only at probability distribution functions $\gamma = F(\alpha)$. Fig.9 is the example of nodes combination $h(2)$ from MHR method:

Person D

Nodes (2;2), (3;1), (4;1), (5;1), (6;2) and $\gamma = F(\alpha) = 2^\alpha - 1$:

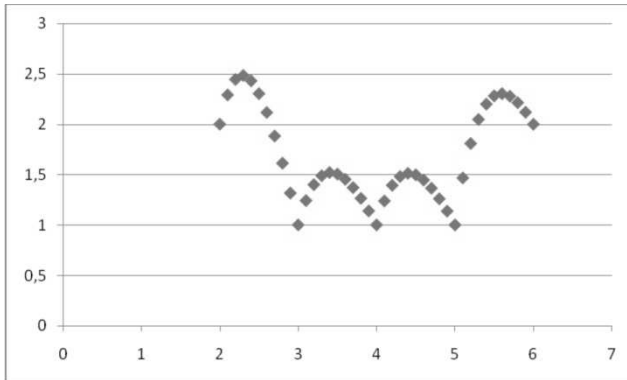


Fig. 9. PNC modeling for nine reconstructed points between nodes

Examples above have one function $\gamma = F(\alpha)$ and one combination h for all ranges between nodes. But it is possible to create a model with functions $\gamma_i = F_i(\alpha)$ and combinations h_i individually for a range of nodes $(p_i; p_{i+1})$. It enables very precise modeling of handwritten symbol between each successive pair of nodes.

Each person has its own characteristic and individual handwritten letters, numbers or other marks. The range of coefficients x has to be the same for all models because of comparing appropriate coordinates y . Every letter is modeled by PNC via three factors: the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h . These three factors are chosen individually for each letter, therefore this information about modeled letters seems to be enough for specific PNC curve interpolation, comparing and handwriting identification. Function γ is selected via the analysis of points between nodes and we may assume $h = 0$ at the beginning. What is very important - PNC modeling is independent of the language or a kind of symbol (letters, numbers or others). One person may have several patterns for one handwritten letter. Summarize: every person has the basis of patterns for each handwritten letter or symbol, described by the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h . Whole basis of patterns consists of models S_j for $j = 0, 1, 2, 3 \dots K$.

3.2. Unknown author – points of handwritten character

Choice of characteristic points (nodes) for unknown letter or handwritten symbol is a crucial factor in object recognition. The range of coefficients x has to be the same like the x range in the basis of patterns. Knots of the curve (opened or closed) ought to be settled at key points, for example local minimum or maximum (the highest point of the curve in a particular orientation), convexity changing or curvature maximum and at least one node between two successive key points. When the nodes are fixed, each coordinate of every chosen point on the curve $(x_0^c, y_0^c), (x_1^c, y_1^c), \dots, (x_M^c, y_M^c)$ is accessible to be used for comparing with the models. Then probability distribution function $\gamma = F(\alpha)$ and nodes combination h have to be taken from the basis of modeled letters to calculate appropriate second coordinates $y_i^{(j)}$ of the pattern S_j for first coordinates $x^c, i = 0, 1, \dots, M$. After interpolation it is possible to compare given handwritten symbol with a letter in the basis of patterns.

3.3. Recognition – the writer

Comparing the results of PNC interpolation for required second coordinates of a model in the basis of patterns with points on the curve $(x_0^c, y_0^c), (x_1^c, y_1^c), \dots, (x_M^c, y_M^c)$, we can say if the letter or symbol is written by person A, B or another. The comparison and decision of recognition [33] is done via minimal distance criterion. Curve points of unknown handwritten symbol are: $(x_0^c, y_0^c), (x_1^c, y_1^c), \dots, (x_M^c, y_M^c)$. The criterion of recognition for models $S_j = \{(x_0^c, y_0^{(j)}), (x_1^c, y_1^{(j)}), \dots, (x_M^c, y_M^{(j)})\}$, $j=0, 1, 2, 3 \dots K$ is given as:

$$\sum_{i=0}^M |y_i^c - y_i^{(j)}| \rightarrow \min .$$

Minimal distance criterion helps us to fix a candidate for unknown writer as a person from the model S_j .

4. Conclusions

The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves [34] using nodes combinations and different coefficients γ : polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting or signature identification via discrete set of fixed knots. PNC makes possible the combination of two important problems: interpolation and modeling in a matter of writer identification. Main features of PNC method are:

- a) the smaller distance between knots the better;
- b) calculations for coordinates close to zero and near by extremum require more attention because of importance of these points;
- c) PNC interpolation develops a linear interpolation into other functions as probability distribution functions;
- d) PNC is a generalization of MHR method via different nodes combinations;
- e) interpolation of L points is connected with the computational cost of rank $O(L)$ as in MHR method;
- f) nodes combination and coefficient γ are crucial in the process of curve probabilistic parameterization and interpolation: they are computed individually for a single curve.

Future works are going to: application of PNC method in signature and handwriting recognition, choice and features of nodes combinations and coefficient γ , implementation of PNC in computer vision and artificial intelligence: shape geometry, contour modelling, object recognition and curve parameterization.

References

1. Schlappbach, A., Bunke, H.: Off-line writer identification using Gaussian mixture models. In: International Conference on Pattern Recognition, pp. 992–995 (2006)
2. Bulacu, M., Schomaker, L.: Text-independent writer identification and verification using textural and allographic features. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (4), 701–717 (2007)

3. Djeddi, C., Souici-Meslati, L.: A texture based approach for Arabic writer identification and verification. In: International Conference on Machine and Web Intelligence, pp. 115–120 (2010)
4. Djeddi, C., Souici-Meslati, L.: Artificial immune recognition system for Arabic writer identification. In: International Symposium on Innovation in Information and Communication Technology, pp. 159–165 (2011)
5. Nosary, A., Heutte, L., Paquet, T.: Unsupervised writer adaption applied to handwritten text recognition. *Pattern Recogn. Lett.* 37 (2), 385–388 (2004)
6. Van, E.M., Vuurpijl, L., Franke, K., Schomaker, L.: The WANDA measurement tool for forensic document examination. *J. Forensic Doc. Exam.* 16, 103–118 (2005)
7. Schomaker, L., Franke, K., Bulacu, M.: Using codebooks of fragmented connected-component contours in forensic and historic writer identification. *Pattern Recogn. Lett.* 28 (6), 719–727 (2007)
8. Siddiqi, I., Cloppet, F., Vincent, N.: Contour based features for the classification of ancient manuscripts. In: Conference of the International Graphonomics Society, pp. 226–229 (2009)
9. Garain, U., Paquet, T.: Off-line multi-script writer identification using AR coefficients. In: International Conference on Document Analysis and Recognition, pp. 991–995 (2009)
10. Bulacu, M., Schomaker, L., Brink, A.: Text-independent writer identification and verification on off-line Arabic handwriting. In: International Conference on Document Analysis and Recognition, pp. 769–773 (2007)
11. Ozaki, M., Adachi, Y., Ishii, N.: Examination of effects of character size on accuracy of writer recognition by new local arc method. In: International Conference on Knowledge-Based Intelligent Information and Engineering Systems, pp. 1170–1175 (2006)
12. Chen, J., Lopresti, D., Kavallieratou, E.: The impact of ruling lines on writer identification. In: International Conference on Frontiers in Handwriting Recognition, pp. 439–444 (2010)
13. Chen, J., Cheng, W., Lopresti, D.: Using perturbed handwriting to support writer identification in the presence of severe data constraints. In: Document Recognition and Retrieval, pp. 1–10 (2011)
14. Galloway, M.M.: Texture analysis using gray level run lengths. *Comput. Graphics Image Process.* 4 (2), 172–179 (1975)
15. Siddiqi, I., Vincent, N.: Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recogn. Lett.* 43 (11), 3853–3865 (2010)

16. Ghiasi, G., Safabakhsh, R.: Offline text-independent writer identification using codebook and efficient code extraction methods. *Image and Vision Computing* 31, 379–391 (2013)
17. Shahabinejad, F., Rahmati, M.: A new method for writer identification and verification based on Farsi/Arabic handwritten texts, Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), pp. 829–833 (2007)
18. Schlapbach, A., Bunke, H.: A writer identification and verification system using HMM based recognizers, *Pattern Anal. Appl.* 10, 33–43 (2007)
19. Schlapbach, A., Bunke, H.: Using HMM based recognizers for writer identification and verification, 9th Int. Workshop on Frontiers in Handwriting Recognition, pp. 167–172 (2004)
20. Marti, U.-V., Bunke, H.: The IAM-database: an English sentence database for offline handwriting recognition, *Int. J. Doc. Anal. Recognit.* 5, 39–46 (2002)
21. Collins II, G.W.: *Fundamental Numerical Methods and Data Analysis*. Case Western Reserve University (2003)
22. Chapra, S.C.: *Applied Numerical Methods*. McGraw-Hill (2012)
23. Ralston, A., Rabinowitz, P.: *A First Course in Numerical Analysis – Second Edition*. Dover Publications, New York (2001)
24. Zhang, D., Lu, G.: Review of Shape Representation and Description Techniques. *Pattern Recognition* 1(37), 1-19 (2004)
25. Schumaker, L.L.: *Spline Functions: Basic Theory*. Cambridge Mathematical Library (2007)
26. Dahlquist, G., Bjoerck, A.: *Numerical Methods*. Prentice Hall, New York (1974)
27. Jakóbczak, D.: 2D and 3D Image Modeling Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies* 4A(16), 104-107 (2007)
28. Jakóbczak, D.: Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science* 6374 (Computer Vision and Graphics: Proc. ICCVG 2010, Part I), Springer-Verlag Berlin Heidelberg, 411-419 (2010)
29. Jakóbczak, D.: Curve Interpolation Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies* 3B(18), 126-130 (2009)

Abstract

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve modeling and handwriting identification by using the set of key points. Nodes are treated as characteristic points of signature or handwriting for modeling and writer recognition. Identification of handwritten letters or symbols need modeling and the model of each individual symbol or character is built by a choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables curve parameterization and interpolation for each specific letter or symbol. Two-dimensional curve is modeled and interpolated via nodes combination and different functions as continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

Streszczenie

Zaproponowana autorska metoda Probabilistycznej Kombinacji Węzłów-Probabilistic Nodes Combination (PNC) jest sposobem modelowania krzywej 2D oraz identyfikacji i rozpoznania pisma odręcznego na podstawie punktów kluczowych (węzłów). Węzły traktowane są jako punkty charakterystyczne podpisu odręcznego lub pisma przed modelowaniem i rozpoznaniem. Dwuwymiarowe dane są interpolowane z wykorzystaniem różnych funkcji rozkładu prawdopodobieństwa: potęgowych, wielomianowych, wykładniczych, logarytmicznych, trygonometrycznych, cyklometrycznych. W pracy pokazano propozycję algorytmu modelowania i rozpoznania pisma odręcznego.