2010 International Conference on Pattern Recognition

Gaussian ERP Kernel Classifier for Pulse Waveforms Classification

Dongyu Zhang[1], Wangmeng Zuo[1], David Zhang[2,1], Yanlai Li[1], Naimin Li[1]
[1]School of Computer Science and Technology
Harbin Institute of Technology
Harbin, China
cswmzuo@gmail.com

[2]Biometrics Research Centre, Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
csdzhang@comp.polyu.edu.hk

***Abstract*—While advances in sensor and signal processing techniques have provided effective tools for quantitative research on traditional Chinese pulse diagnosis (TCPD), the automatic classification of pulse waveforms is remained a difficult problem. To address this issue, this paper proposed a novel edit distance with real penalty (ERP)-based *k*-nearest neighbors (KNN) classifier by referring to recent progresses in time series matching and KNN classifier. Taking advantage of the metric property of ERP, we first develop a Gaussian ERP kernel, and then embed it into kernel difference-weighted KNN classifier. The proposed Gaussian ERP kernel classifier is evaluated on a dataset which includes 2470 pulse waveforms. Experimental results show that the proposed classifier is much more accurate than several other pulse waveform classification approaches.**

Keywords-pulse diagnosis; edit distance with real penalty; k-nearest neighbors; pulse waveform; kernel method;

I. Introduction

Traditional Chinese pulse diagnosis (TCPD) [1] is a convenient, noninvasive, and effective diagnostic method, where practitioners feel the fluctuations in radial pulse at the styloid processes of wrist classify them into distinct patterns which are related to different syndromes and diseases in traditional Chinese medicine (TCM). TCPD is a skill which requires considerable experience and training, and the diagnosis results may vary for different practitioners. Thus, techniques developed for measuring and analyzing the physiological signals are recently considered in quantitative TCPD research [2, 3, 4] as a way to improve the reliability and consistency of diagnoses.

Although much progress has been made in TCPD quantification research, the automatic classification of pulse waveforms is remained a difficult problem. TCPD recognizes more than 20 kinds of pulse patterns, which are defined according to the criteria such as shape, position, regularity, force, and rhythm [1]. Fig. 1 shows five typical pulse waveforms which differ in their shapes.

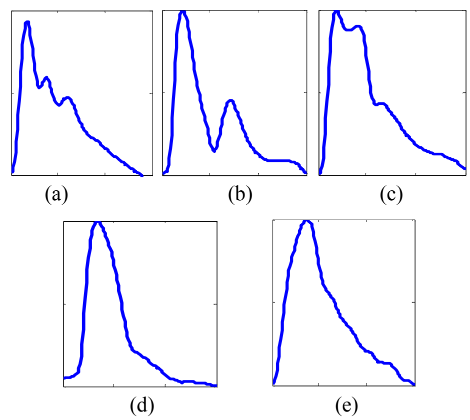

Figure 1. Five pulse patterns classified by shape: (a) moderate, (b) smooth, (c) taut, (d) hollow, and (e) unsmooth.

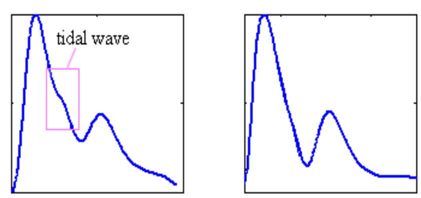

Figure 2. Two pulse patterns with similar shapes: (a) A moderate pulse with unnoticeable tidal wave is similar to (b) a smooth pulse.

Pulse waveform classification suffers from the problems of complicated intra- and inter-class variations. For example, a moderate pulse with unnoticeable tidal wave is similar to a smooth pulse (see Fig. 2). For some taut pulses, the tidal waves are very high and even merged with percussion wave (see Fig. 3). Besides, the time axis distortion and noise also have adverse influence on classification accuracy. For several classification methods which had been developed for pulse waveform classification, e.g., neural networks [5, 7], and dynamical time warping (DTW) [6], the reported accuracies are mostly below 90%, and usually are tested on small datasets.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.670

IEEE computer society

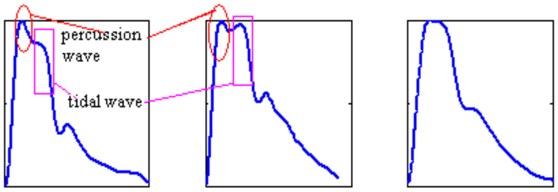

Figure 3. Intra-class variations of taut pulse waveforms: (a) typical taut pulse, (b) taut pulse with high tidal wave, (c) taut pulse with tidal wave merged with percussion wave.

Motivated by progress in elastic metric, this paper investigates the approach to utilize time series matching for pulse waveforms classification. In time series classification, many elastic distance measures, e.g., DTW [8], have been proposed to address the time axis distortion problem, but only few of them, i.e., edit distance with real penalty (ERP) [9], are metrics. Besides, the metric property of distance measures has to date been applied mainly to the design of fast querying algorithms, and not to classifier design [9, 10]. Furthermore, simple KNN classifier usually is adopted in time series classification, where recently developed advanced KNN classifiers are rarely adopted [11, 12]. By incorporating time series matching method with appropriate classifier, we expect to develop accurate pulse waveform classification methods by simultaneously addressing the intra-class variation and the time axis distortion problems.

In this paper, we first develop a Gaussian ERP (GERP) kernel by utilizing the metric property of ERP distance. Then we further present a GERP kernel classifier (GEKC) by using the KDF-WKNN [11] classification framework. The proposed method is evaluated on a dataset which includes 2470 pulse waveforms of five common pulse patterns. Experimental results show that the proposed methods achieve an average classification accuracy of 91.74%, which is much higher than several other pulse waveform classification approaches.

The remainder of this paper is organized as follows. Section II first presents a survey on ERP and KDF-WKNN, and then proposes the GEKC method. Section III provides the experimental results on pulse waveforms classification. Finally, conclusions are drawn in Section IV.

II. METHOD

In this section, we first provide a brief introduction to ERP and KDF-WKNN. Then by defining a novel elastic kernel, GERP kernel, we propose a GERP kernel classifier (GEKC) for pulse waveform classification.

A. Related work

ERP [9] is a recently-developed elastic distance metric which can be regarded as the marriage of l_p-norm and edit distance. Given two time series $A = [a_1, a_2, \ldots, a_m]$ with m elements and $B = [b_1, b_2, \ldots, b_n]$ with n elements, the ERP distance $d_{erp}(A,B)$ of A and B is recursively defined as,

$$d_{erp}\left(A_1^m, B_1^n\right) = \begin{cases} \sum_{i=1}^{m} |a_i - g|, & \text{if } n = 0 \\ \sum_{i=1}^{n} |b_i - g|, & \text{if } m = 0 \\ \min \begin{Bmatrix} d_{erp}\left(A_2^m, B_2^n\right) + |a_1 - b_1| \\ d_{erp}\left(A_2^m, B_1^n\right) + |a_1 - g| \\ d_{erp}\left(A_1^m, B_2^n\right) + |b_1 - g| \end{Bmatrix}, & \text{otherwise} \end{cases} \quad (1)$$

where $A_i^p = \left[a_i, \ldots, a_p\right]$ denotes the subsequence of A, $|\cdot|$ denotes the l_1-norm, and g is a constant with the default value 0 [9].

Theorem 1 [9] Let Q, R, S be three time series of arbitrary length. Then it is necessary that ERP satisfies the triangle inequality, $d_{erp}(Q,S) \le d_{erp}(Q,R) + d_{erp}(R,S)$.

Corollary 2 [9] The ERP distance satisfies the triangle inequality and it is a metric.

KDF-WKNN is a kernel-based KNN classifier which could obtain classification performance comparable to or better than several state-of-the-art classification methods [11]. In KDF-WKNN, given an unclassified sample x and its k-nearest neighbors $X^n = \left\{x_1^n, x_2^n, \ldots, x_k^n\right\}$, the weights of k-nearest neighbors $W = [w_1, w_2, \ldots, w_k]^{\mathbf{T}}$ are defined as a vector corresponding to the constrained optimal reconstruction of x using X^n in the feature space $\mathbf{F}: x \rightarrow \Phi(x)$ as,

$$W = \arg\min \frac{1}{2} \left\| \Phi(x) - \Phi\left(X^n\right) W \right\|^2 \quad \text{s.t. } \sum_{i=1}^{k} w_i = 1 \quad (2)$$

The weights W can be obtained by solving the following linear equation,

$$\left[\mathbf{G} + \eta tr(\mathbf{G}) \mathbf{I}_k / k\right] W = \mathbf{1}_k, \quad (3)$$

where η is the regularization parameter, $tr(\mathbf{G})$ is the trace of $\mathbf{G}$, $\mathbf{I}_k$ is identity matrix, $\mathbf{1_k}$ is a $k\times1$ vector with all elements equal to 1. The element at the ith row and the jth column of $\mathbf{G}$ is defined as

$$\mathbf{G}_{ij} = k\left(x_i^n, x_j^n\right) + k(x,x) - k\left(x, x_i^n\right) - k\left(x, x_j^n\right), \quad (4)$$

where $k(\cdot,\cdot)$ denotes the kernel function.

Finally, the weighted KNN rule is used to assign a class label to sample x. For the detailed description of KDF-WKNN, please refer to [11].

B. The GEKC method

Our motivation for developing the GEKC method is to incorporate ERP with Gaussian function to develop a novel Gaussian ERP (GERP) kernel function. Because of the possible emergence of time axis distortion, classical kernel functions [20], such as Gaussian RBF and polynomial, generally would not be suitable for time series classification. To address the adverse influence of time axis distortion, elastic kernels, such as Gaussian DTW kernel, have been proposed for kernel-based classifiers [13, 14]. Gaussian

DTW kernel is definitely not PDS kernel [15], and it has been reported that Gaussian DTW kernel classifiers usually cannot guarantee the performance improvement over conventional Gaussian RBF kernel classifiers [15, 16]. Utilizing the metric property of ERP, we expect that the proposed GERP kernel would outperform either Gaussian RBF kernel or Gaussian DTW kernel for time series classification. Using the GERP and KDF-WKNN, we further propose a Gaussian ERP kernel classifier (GEKC) for pulse waveform classification. In the following, we will first introduce the GERP kernel and then describe the detail of GEKC.

The GERP is developed from a kind of kernel function which has the following definition.

Definition 1 Let the distance function $d(x,x')$ be symmetric, non-negative, and has zero diagonal, i.e. $d(x,x')=0$, then a kind of kernel $K(x,x')$ can be defined as follows,

$$K(x,x')=\exp\left(-\gamma d^2(x,x')\right), \forall \gamma>0. \quad (5)$$

Previous work has shown that it is possible to construct PDS kernel by using elastic distance [17, 18]. However, if the elastic distance is non-metric, the kernel function $K(x,x')$ defined in (5) is not PDS [19] and is not admissible to standard kernel machines, e.g., support vector machines. Thus we expect to utilize elastic metric rather than just elastic distance to construct kernel function.

Definition 2 Given two time series x and x', the Gaussian ERP kernel function $k_{erp}(x,x')$ is defined as,

$$k_{erp}=\exp\left(-d_{erp}^2(x,x')/\sigma^2\right). \quad (6)$$

where σ is the standard deviation of the Gaussian function.

Based on the metric property of ERP, GERP kernel may be more suitable for kernel classifiers. Actually, we also cannot guarantee the PDS property of the GERP kernel. To remedy this, we may enforce the PDS property by adding sufficiently small values of the variance [20] or forming mathematically correct kernels based on global alignments [18].

For the pulse waveform classification task, we randomly choose a set of samples and experimentally analyze the PDS property of the GERP kernel. We run the experiment 10 times and do not observe the violation of the PDS property for the GERP kernel. For Gaussian DTW kernel, however, we can observe the violation of the PDS property. In [16], Gudmundsson *et al.* claimed that several amendments to guarantee the PDS property (e.g., [18]) may make the classifier with poor generalization performance. Thus, for our pulse waveform classification task, we choose the GERP kernel without any amendment in the GEKC classifier.

We directly add the GERP kernel to KDF-WKNN to construct GEKC by defining the Gram matrix $\mathbf{G}_{\mathrm{ERP}}$,

$$\mathbf{G}_{\mathrm{ERP}}=\mathbf{K}+\mathbf{1}_{kk}-\mathbf{1}_k\mathbf{k}^{\mathrm{T}}-\mathbf{k}\mathbf{1}_k^{\mathrm{T}}, \quad (7)$$

where $\mathbf{K}$ is a $k\times k$ matrix with its element at the ith row and jth column as

$$\mathbf{K}_{ij}=\exp\left(-d_{erp}^2\left(x_i^n,x_j^n\right)/\sigma^2\right), \quad (8)$$

and $\mathbf{k}$ is a $k\times 1$ vector with its element

$$\mathbf{k}_i=\exp\left(-d_{erp}^2(x,x_i^n)/\sigma^2\right). \quad (9)$$

Once we have the Gram matrix $\mathbf{G}_{\mathrm{ERP}}$, we can use KDF-WKNN for pulse waveforms classification by solving the linear system of equations defined in (4). As a summary, the details of the GEKC algorithm are provided in Fig. 4.

Input: Unclassified sample x, training sample $\{x_1,\ldots,x_n\}$ with the corresponding class labels $\{y_1,\ldots,y_n\}$, parameter, η, the standard deviation, σ, and the number of nearest neighbors, k.
Output: Predicted class label $\omega_{j_{\max}}$ of the sample x.

1. Use ERP metric to obtain the k nearest neighbors $\{x_1^n,\cdots,x_k^n\}$ of sample x, and their corresponding class labels $\{y_1^n,\cdots,y_k^n\}$.
2. Calculate the GERP-induced inner product of the nearest neighbors of the sample $\boldsymbol{x}$ by (8).
3. Calculate the GERP-induced inner product of sample x and each nearest neighbors by (9).
4. Calculate Gram matrix $\mathrm{G}_{\mathrm{ERP}}$ using (7).
5. Obtain W by solving $\left[\mathrm{G}_{\mathrm{ERP}}+\eta tr(\mathrm{G}_{\mathrm{ERP}})\mathrm{I}_k/k\right]W=\mathbf{1}_k$.
6. Assign class label $\omega_{j_{\max}}$ to sample x using following rule,

$$\omega_{j_{\max}}=\arg\max_{\omega_j}\left(\sum_{y_i^n=\omega_j}w_i\right).$$

Figure 4. The proposed GEKC algorithm.

III. EXPERIMENTAL RESULTS

In order to evaluate the performance of GEKC, we construct a dataset which includes 2470 pulse waveforms of five pulse patterns, moderate (M), smooth (S), taut (T), hollow (H), and unsmooth (U) pulse. All of the data are collected at Harbin Binghua Hospital from people of 15 to 75 years old. The pulse waveform datasets used in literature [6, 7] only contain less than or about 1000 waveforms, and this dataset may be the largest data set yet used in pulse waveforms classification. Table I lists the number of pulse waveforms for each pulse pattern. We made use of only one period from each pulse and normalized them all to the same length of 150 points by using bilinear interpolation.

TABLE I. THE DATASET USED IN THE EXPERIMENT

Pulse Patterns	M	S	T	U	H	Total
Numbers	800	550	800	160	160	2470

We randomly split the dataset into three parts and applied 10 runs of 3-fold cross-validation to evaluate the classification performance of GEKC method. The hyper-parameters, k, η, and σ of GEKC, are chosen as k = 31, η = 0.01, σ = 16 using the stepwise selection strategy described in [11]. The specificity and the sensitivity obtained using GEKC are 78.76% and 91.74%, respectively.

To provide an objective comparison, we independently implemented two pulse waveform classification methods, i.e., IDTW [6] and wavelet network [7], and evaluate their performance on our dataset. The average classification rates of these two methods are listed in Table II. Besides, we also compare the proposed method with several related classification methods, i.e., 1NN-Euclidean (1NN-ED), 1NN-DTW, and 1NN-ERP. These results are also listed in Table II. From Table II, one can see that, in term of overall average classification accuracy, our method outperforms all the other methods.

TABLE II. AVERAGE CLASSIFICATION RATES (%)

Classification Methods	1NN-DTW	1NN-ERP	1NN-ED	WN [7]	IDTW [6]	GEKC
Moderate	82.44	88.31	86.11	87.23	87.31	**91.25**
Smooth	81.16	86.31	85.02	85.36	80.38	**87.09**
Taut	87.95	95.10	95.76	89.63	93.15	**96.88**
Hollow	82.44	87.56	86.75	85.63	80.44	**89.38**
Unsmooth	70.81	84.75	84.06	80.63	**89.50**	86.88
Average accuracy	83.19	89.79	87.36	87.08	88.90	**91.74**

IV. CONCLUSION

In this paper, we propose a Gaussian ERP kernel and a novel kernel classifier, GEKC, for pulse waveform classification. To evaluate the classification performance of GEKC, we built a dataset of 2470 pulse waveforms. Experimental results show that the proposed method achieves an average recognition accuracy of 91.74%, which is much higher than several other pulse waveforms classification methods.

One potential advantage of the proposed methods is to utilize the low bounds and metric property of ERP for fast pulse waveform classification and indexing [9]. In our future work, we will further investigate accurate and computationally efficient ERP-based classifiers for pulse waveform classification.

ACKNOWLEDGMENT

This work is partially supported by the CERG fund from the HKSAR Government, the central fund from Hong Kong Polytechnic University, and the NSFC/SZHK-innovation funds of China under Contract Nos. 60871033, 60902099, 60872099, and SG200810100003A.

REFERENCES

[1] S. Z. Li, Pulse Diagnosis, Paradigm Press, 1985.

[2] Z. F. Fei, Contemporary Sphygmology in Traditional Chinese Medicine, People’s Medical Publishing House, Beijing, China, 2003 (in Chinese).

[3] C. C. Tyan, S. H. Liu, J. Y. Chen et al., "A novel noninvasive measurement technique for analyzing the pressure pulse waveform of the radial artery," IEEE Trans. Biomed. Eng., vol. 55, no. 1, 2008, pp. 288-297.

[4] L. S. Xu, D. Zhang, and K. Q. Wang, "Wavelet-based cascaded adaptive filter for removing baseline drift in pulse waveforms," IEEE Trans. BME, vol. 52, no. 11, 2005, pp. 1973-1975.

[5] R. Folland, A. Das, R. Dutta, E. L. Hines, N. G. Stocks, and D. Morgan, "Pulse waveform classification using neural networks with cross-validation techniques: Some signal processing considerations," Proc. IASTED Int'l Conf. Biomed. Eng., 2004, pp. 337-342.

[6] L. Wang, K. Q. Wang, and L. S. Xu, "Recognizing wrist pulse waveforms with improved dynamic time warping algorithm," Proc. Third ICMLC, 2004, pp. 3644-3649.

[7] L. S. Xu, K. Q. Wang, and L. Wang, "Pulse waveforms classification based on wavelet network," Proc. 27th IEEE EMBC, 2005, pp. 4596-4599.

[8] B.-K. Yi, H. V. Jagadish and C. Faloutsos, "Efficient retrieval of similar time sequences under time warping," Proc. 14th ICDE, 1998, pp. 201-208.

[9] L. Chen, and R. Ng, "On the marriage of Lp-norm and edit distance," Proc. 30th VLDB, 2004, pp.792-801.

[10] P.F. Marteau, "Time warp edit distance with stiffness adjustment for time series matching," IEEE Trans. PAMI, vol. 31, no. 2, 2009, pp. 306-318.

[11] W. M. Zuo, D. Zhang, and K. Q. Wang, "On kernel difference-weighted *k*-nearest neighbor classification," Pattern Analysis & Applications, vol. 11, 2008, pp. 247-257.

[12] M. R. Gupta, R. M. Gray, and R. A. Olshen, "Nonparametric supervised learning by linear interpolation and maximum entropy," IEEE Trans. PAMI, vol. 28, no.5, 2006, pp. 766-781.

[13] C. Bahlmann, B. Haasdonk, and H. Burkhardt, "Online handwriting recognition with support vector machines -a kernel approach," Proc. 8th IWFHR, 2002, pp. 49-54.

[14] W. A. Chaovalitongse, and P. M. Pardalos, "On the time series support vector machine using dynamic time warping kernel for brain activity classification," Cybernetics and Systems Analysis, vol. 44, no. 1, 2008, pp. 125-138.

[15] H. S. Lei, and B. Y. Sun, "A Study on the dynamic time warping in kernel machines," Proc. Third Int’l IEEE Conf. SITIS, 2007, pp. 839-845.

[16] S. Gudmundsson, T.P. Runarsson, and S. Sigurdsson, "Support vector machines and dynamic time warping for time series", IEEE International Joint Conference on Neural Networks, 2008, pp. 2772-2776.

[17] C. Cortes, P. Haffner, and M. Mohri, "Rational kernels: theory and algorithms", Journal of Machine Learning Research, vol. 5,2004, pp. 1035-1062.

[18] M. Cuturi, J. P. Vert, O. Birkenes, and T. Matsui, "A kernel for time series based on global alignments", Proc. ICASSP, 2007, pp. 413-416.

[19] B. Schölkopf and A.J. Smola, Learning with Kernels, MIT Press, 2002, pp. 49 - 53.

[20] V. Sulimova, V. Mottl, and A. Tatarchuk, "Multi-kernel approach to on-line signature verification", Proceedings of the 8th IASTED Internal Conference on Signal and Image Processing, Hawaii, USA, August, 2006.